STATISTICAL ANALYSIS PLAN

Study: PS0013

Product: Bimekizumab

A PHASE 3, MULTICENTER, DOUBLE-BLIND, PLACEBO-CONTROLLED STUDY WITH AN INITIAL TREATMENT PERIOD FOLLOWED BY A RANDOMIZED-WITHDRAWAL 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	30 Aug 2018
Amendment 1	16 Sep 2019

REDACTED COPY
This document cannot be used to support any marketing authorization application or any extensions or variations thereof.

Confidentiality Statement

Confidential

This document is the property of UCB and may not – in full or in part – be passed on, reproduced, published, or otherwise used without the express permission of UCB.

TABLE OF CONTENTS

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

REDACTED COPY
This document cannot be used to support any marketing authorisation application and any extensions or variations thereof.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

LIST OF ABBREVIATIONS

ADAb	anti-bimekizumab antibodies
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
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

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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
IDC	Infectious Disease Committee
IGA	Investigator's Global Assessment
IGRA	interferon-gamma release assay
IL	interleukin
IMP	investigational medicinal product
LLOQ	Lower level of quantification
LOCF	last observation carried forward
LFT	liver function test
mNAPSI	Modified Nail Psoriasis Severity Index Score
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
MNAR	missing not at random
nADAb	Neutralizing anti-bimekizumab antibodies
NRI	nonresponder 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	Physician's Global Assessment

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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
PT	preferred term
Q4W	every 4 weeks
Q8W	every 8 weeks
QOL	quality of life
RS	Randomized Set
SAE	serious adverse event
SIB	suicidal ideation and behavior
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
SMQ	Standard MedDRA query
SOC	system organ class
SS	Safety Set
TB	tuberculosis
TEAE	treatment-emergent adverse event
TNF	tumor necrosis factor
ULN	upper limit of normal
VAS	visual analog scale

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

WBC	white blood cell
WHO-DD	World Health Organization Drug Dictionary
WK16ResS	Week 16 Responder Set

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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 3, 21 May 2019.

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 subcutaneously (sc) for 16 weeks versus placebo 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 placebo at achieving complete clearance (Psoriasis Area and Severity Index [PASI]100) after 16 weeks of treatment
- Evaluate the efficacy of bimekizumab compared to placebo after 4 weeks of treatment
- Evaluate the change in itch, pain, and scaling of bimekizumab compared to placebo after 16 weeks of treatment as reported by subjects using the Patient Symptom Diary
- Evaluate the change in psoriatic scalp disease of bimekizumab compared to placebo after 16 weeks of treatment in subjects with scalp PSO at Baseline
- Evaluate the efficacy of continuous treatment with bimekizumab versus treatment withdrawal (placebo) as defined by PASI90 at Week 56 for subjects who responded to bimekizumab treatment at Week 16
- 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 time to relapse among subjects who responded to bimekizumab treatment at Week 16
- Assess the efficacy of 12 weeks of escape treatment with bimekizumab in subjects who relapse (defined as not achieving a PASI75 response) during the Randomized-Withdrawal Period
- Assess the efficacy of 12 weeks of escape treatment with bimekizumab for subjects who do not achieve a PASI90 response (defined as a subject that does not achieve 90% reduction in the PASI score) at Week 16

REDACTED COPY
This document cannot be used to support any marketing authorisation application and any extensions or variations thereof.

- Assess the efficacy of bimekizumab over time
- Assess the efficacy of bimekizumab dosing Q4W versus Q8W during the Randomized-Withdrawal Period
- 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 16
- Assess depression
- Assess the pharmacokinetics (PK) of bimekizumab
- Assess the immunogenicity of bimekizumab
- 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 variables

The secondary efficacy variables are:

- PASI100 response at Week 16
- IGA response (defined as Clear with at least a 2-category improvement relative to Baseline) at Week 16
- PASI75 response at Week 4

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- Patient Symptom Diary responses for itch, pain, and scaling at Week 16
- Scalp IGA response (Clear or Almost Clear with at least a two-category improvement from Baseline) at Week 16 for subjects with scalp PSO at Baseline
- PASI90 response at Week 56 among Week 16 PASI90 responders

2.2.1.3 Other efficacy variables

Change from Baseline variables evaluated during the Initial Treatment Period are relative to the Baseline (first dose) Visit. For subjects who are re-randomized at the Week 16 Visit, change from Baseline variables during the Randomized-Withdrawal Period may be evaluated relative to both the Baseline (first dose) Visit and the Week 16 Visit. Similarly, for subjects who initiate escape treatment with bimekizumab (either due to lack of response at Week 16 or loss of response [relapse] during the Randomized-Withdrawal Period), these variables during the 12-week escape treatment period may be evaluated relative to both the Baseline Visit (first dose) and the visit at which escape treatment was initiated.

For simplicity, "change from Baseline" is used below for all such variables and applies to all scheduled visits where the assessments were performed (whether during the Initial Period, Randomized-Withdrawal Period, or Escape Treatment Period). 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 listed below and will be summarized for all time points specified in the protocol schedule of assessments, with the exception of the previously defined primary and secondary efficacy variables (which will be summarized separately):

- PASI50, PASI75, PASI90, and PASI100 response
- Time to PASI50, PASI75, PASI90, and PASI100 response during the Initial Treatment Period
- Time to relapse during the Randomized-Withdrawal Period
- Percentage of subjects who relapse during the Randomized-Withdrawal Period
- Percentage of subjects who rebound (defined as a ≥25% increase from Baseline in PASI score occurring within 2 months of stopping therapy) during the Randomized-Withdrawal 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 2-category improvement from Baseline)
- IGA response (Clear or Almost Clear with at least 2-category improvement from Baseline)
- Shift from Baseline in IGA score
- Absolute and percent change from Baseline in the body surface area (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)

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- Change from Baseline in Dermatology Life Quality Index (DLQI)
- Percentage 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 the Patient's Global Assessment of Disease Activity (PGADA) for the arthritis visual analog scale (VAS) in the subgroups of subjects with PsA as defined by PASE score at Baseline
- Change from Baseline in Patient Global Assessment (PGA) of PSO score
- Patient Symptom Diary responses
- Scalp-specific Investigator's Global Assessment (scalp IGA) response (Clear or Almost Clear with at least 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 response (defined as a subject that achieves at least a 75% reduction from Baseline in the mNAPSI score)
- 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 at least a 100% reduction from Baseline in the mNAPSI score)
- Palmoplantar Investigator's Global Assessment (pp-IGA) response (Clear or Almost Clear with at least 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

2.2.2 Pharmacokinetic variable

The PK variable is the plasma concentration of bimekizumab.

2.2.3 Safety variables

2.2.3.1 Secondary safety variables

- TEAEs adjusted by duration of subject exposure to study treatment

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- 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

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

2.2.4 Pharmacogenomic variables

Additional blood samples will be collected from subjects who consent to participate in the substudy at specific time points and stored at -80°C for up to 20 years.

Genomic, genetic, epigenetic, proteins, and metabolite biomarkers may be measured to evaluate the relationship with response to treatment with bimekizumab, PSO disease biology, and inflammatory and immune response processes. The nature and format of these tentative substudy analyses will be determined when the results of the main study are made available.

The candidate exploratory variables are the blood or blood derivative (e.g., serum) concentrations of cytokines and chemokines of relevance to IL-17A/F signaling pathway and PSO biology. Additional variables may include but will not be limited to serum complement concentrations.

2.2.5 Immunological variable

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

PS0013 is a Phase 3, multicenter, randomized, double-blind placebo-controlled study consisting of a 16-week Initial Treatment Period followed by a 40-week Randomized-Withdrawal Period to evaluate the efficacy and safety of bimekizumab in adult subjects with moderate to severe chronic plaque PSO. To be eligible to participate in this study, subjects must be adults with a diagnosis of moderate to severe PSO (Baseline PASI ≥12 and BSA affected by PSO ≥10% and IGA score ≥3 [on a 5-point scale]) who are a candidate for systemic PSO therapy and/or phototherapy. Subjects may have been previously exposed to a biologic therapy.

2.3.2 Study periods

This study will include 4 periods, a Screening Period (2 to 5 weeks), a placebo-controlled Initial Treatment Period (16 weeks), a placebo-controlled Randomized-Withdrawal Period (40 weeks), and a Safety Follow-Up (SFU) Period (20 weeks after the last dose of IMP). After completing the study, eligible subjects will be allowed to enroll in an open-label study. Subjects enrolling into the open-label study will not have the PS0013 SFU Visit.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Subjects who do not achieve a PASI90 response at Week 16 of the Initial Treatment Period will be allocated to escape treatment (see Section 2.3.2.5). All subjects who relapse at Week 20 or later during the Randomized-Withdrawal Period (up to Week 56) will be allocated to escape treatment (see Section 2.3.2.5). Relapse is defined as not achieving a PASI75 response. Subjects who achieve a PASI50 response at Week 12 of the open-label escape arm will be allowed to enroll in the open-label study.

Subjects who complete the Randomized-Withdrawal Period (after a Week 16 response [≥PASI90] on bimekizumab during the Initial Treatment Period) will be allowed to enroll in an open-label study. Placebo-randomized subjects, who receive placebo during the entire Initial and Randomized-Withdrawal periods, that complete Week 56 without relapse into the open-label escape arm will not be eligible for the open-label study.

2.3.2.1 Screening Period

The Screening Period will last 2 weeks, but can be extended up to a total of 5 weeks in cases where a laboratory assessment needs to be repeated or to allow washout of prohibited medications. 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 Initial Treatment Period

During the placebo-controlled Initial Treatment Period, approximately 400 subjects will be randomized 4:1 to receive the following blinded IMP regimens:

- Bimekizumab 320mg administered Q4W (320 subjects)
- Placebo administered Q4W (80 subjects)

Subjects will be followed in a double-blind fashion. Subjects may receive placebo injections at certain visits in order to blind the IMP.

Subjects withdrawing early from the study will undergo the Premature End of Treatment (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.

2.3.2.3 Randomized-Withdrawal Period

At the Week 16 study visit, bimekizumab subjects who have achieved a PASI90 response will enter into a double-blind, placebo-controlled Randomized-Withdrawal Period lasting 40 weeks. Subjects who do not achieve a PASI90 response at the Week 16 study visit will be allocated to the escape arm (see Section 2.3.2.5).

During the Randomized-Withdrawal Period, dosing will be based on treatment group assignment at initial randomization as follows:

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- Subjects initially randomized to bimekizumab 320mg Q4W will be re-randomized 1:1:1 to either bimekizumab 320mg Q4W or bimekizumab 320mg Q8W or placebo (i.e., treatment withdrawal)
- All subjects initially randomized to placebo who achieve a PASI90 response at Week 16 will continue to receive placebo (Q4W)

Investigational medicinal product will be administered in the clinic at Week 16 and Q4W thereafter through Week 52.

All subjects who relapse at Week 20 or later during the Randomized-Withdrawal Period (up to Week 56) will be allocated to the escape arm (see Section 2.3.2.5). Relapse is defined as not achieving a PASI75 response.

At the end of the Randomized-Withdrawal Period, all subjects enrolling in the open-label study (after signing a new ICF) will undergo the Week 56 study assessments before receiving their first dose of IMP 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 Escape Treatment Period

As described above, subjects who do not achieve a PASI90 response at Week 16 of the Initial Treatment Period and all subjects who relapse at Week 20 or later during the Randomized-Withdrawal Period (up to Week 56) will receive open-label bimekizumab 320mg Q4W for 12 weeks (i.e. escape treatment).

For subjects who relapse during the Randomized-Withdrawal Period, the subject should complete the 12-week escape treatment period even if this extends to a time point beyond what would have been the Week 56 visit (if the subject had not escaped).

Subjects who achieve a PASI50 response by the end of the 12 weeks of escape treatment are eligible to enroll in the open-label study.

2.3.2.6 Premature End of Treatment

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

2.3.3 Study duration per subject

For each subject (of those not entering Escape treatment), the study will last a maximum of up to 77 weeks, as follows:

- Screening Period: 2 to 5 weeks
- Double-blind, placebo-controlled Initial Treatment Period: 16 weeks
- Double-blind, placebo-controlled Randomized-Withdrawal Period: 40 weeks
- Escape treatment (if required): 12 weeks

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- 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 Randomized-Withdrawal 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. Subjects who achieve a PASI50 response in the open-label escape arm will also be allowed to enroll in the open-label study. The SFU Visit will not be required for subjects who enroll in the open-label study. Subjects who complete the Randomized-Withdrawal Period (after a Week 16 response [≥PASI90] on bimekizumab during the Initial Treatment Period) will be allowed to enroll in an open-label study. Placebo-randomized subjects who receive placebo during the entire Initial and Randomized-Withdrawal periods, that complete Week 56 without relapse into the open-label escape arm will not be eligible for 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 sites

Approximately 535 subjects will be screened in order to have 400 subjects randomized in the study. There will be approximately 320 subjects in the bimekizumab treatment arm and 80 subjects in the placebo treatment arm. The planned number of study sites is approximately 100. Every eligible subject who signs an ICF will 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 400 subjects will be randomly assigned in a 4:1 ratio to the following treatment groups:

- Bimekizumab 320mg (320 subjects)
- Placebo (80 subjects)

The primary efficacy analysis is based on the comparison of bimekizumab to placebo for the co-primary efficacy variables of PASI90 response and IGA response at Week 16. The assumed responder rates for PASI90 at Week 16 are 75% and 2% for bimekizumab and placebo, respectively. Additionally, the assumed responder rates for IGA are 85% and 5% for bimekizumab and placebo, respectively. The assumed responder rates for bimekizumab are based on the Phase 2b PS0010 data. The power to show statistical superiority of bimekizumab relative to placebo under these assumptions is >99% for the co-primary endpoints.

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.

REDACTED COPY
This document cannot be used to support any marketing authorisation application and any extensions or variations thereof.

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 will be accounted for using 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 will 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.

Per protocol, visit windows of ±3 days from the first dose to Week 24 and ±7 days from Week 28 to Week 52 are permissible. For the Week 56 visit, the visit window is±3 days. 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

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

assessments that 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.

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

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 visit measurement is missing, and a Screening visit measurement is available, the Screening value will be utilized as Baseline.

For subjects who are re-randomized at the Week 16 Visit, change from Baseline variables during the Randomized-Withdrawal Period may be evaluated relative to both the Baseline (first dose) Visit and the Week 16 Visit. For the latter, the Baseline is defined as the latest measurement on or prior to the first dose in the Randomized-Withdrawal Period (intended to be Week 16).

Similarly, for subjects who initiate escape treatment with bimekizumab (either due to lack of response at Week 16 or loss of response [relapse] during the Randomized-Withdrawal Period), these variables during the 12-week escape treatment period may be evaluated relative to both the Baseline Visit (first dose) and the visit at which escape treatment was initiated. For the latter, the Baseline is defined as the latest measurement on or prior to the first dose in the Escape Period (for those who escape due to lack of response at Week 16) or the latest measurement on or prior to the visit at which escape treatment is initiated.

An additional Baseline for Week 16 and the visit at which escape treatment was initiated (as described above) are defined for the following variables:

- PASI score
- DLQI
- Laboratory data

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.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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 will not be summarized in by-visit tables (though it will 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 antibodies 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 identifier.

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 received at least 1 dose of the IMP and have a valid measurement for each of the co-primary efficacy variables at Baseline.

3.5.5 Escape Subject Set

The Escape Subject Set (ESS) will consist of all subjects who receive at least 1 dose of escape bimekizumab 320mg treatment either due to not achieving a PASI90 response at Week 16 or

REDACTED COPY
This document cannot be used to support any marketing authorisation application and any extensions or variations thereof.

experiencing a relapse after entering the Randomized-Withdrawal Period. Summaries based on the ESS will be split between subjects who enter the escape arm: 1) due to PASI90 non-response at Week 16 or 2) due to relapse during the Randomized-Withdrawal Period (after achieving PASI90 response at Week 16).

3.5.6 Week 16 Responder Set

The Week 16 Responder Set (WK16ResS) will consist of all subjects who achieve a PASI90 response at Week 16 and receive at least 1 dose of the IMP during the Randomized-Withdrawal Period at Week 16 or later.

3.5.7 Active Medication Set

The Active Medication Set (AMS) will consist of all subjects who have received at least 1 dose of active IMP (bimekizumab). The AMS will be used for summaries of safety that include all data from the Initial Treatment Period and/or Randomized-Withdrawal Period.

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 co-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) consists 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 placebo received bimekizumab at any time, then for safety analyses these subjects will be reallocated to the appropriate bimekizumab treatment group, unless otherwise specified. Subjects randomized to bimekizumab will only be reallocated to the placebo treatment group if they never received bimekizumab. Efficacy analyses will be according to randomized treatment and not actual treatment received.

3.6.1 Subjects Incorrectly Entering Randomized-Withdrawal or Escape Period

Subjects who enter the Escape Period without meeting the escape criteria, or subjects who meet the criteria but don't enter the Escape Period, will be summarized as follows for efficacy analyses:

1. If a subject meets escape criterion at Week 16 (<PASI90) but does not enter into the Escape Period at Week 16, then their data during the Randomized-Withdrawal Period will not be summarized in any table as they did not fulfill the definition for inclusion in the WK16ResS. Selected data from the Randomized-Withdrawal Period will be listed only.
2. If a subject meets escape criterion during the Randomized-Withdrawal Period (<PASI75) but does not enter into the Escape Period at the corresponding visit, then their data following the

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

visit at which they meet the escape criterion will be treated as missing in summaries based on the WK16ResS and will be subject to imputation as appropriate.

3. If a subject enters the Escape Period without meeting escape criterion at the corresponding visit, they will not be in the ESS. These subjects also do not qualify for inclusion in the WK16ResS. Selected data corresponding to the Randomized Withdrawal Period and Escape Period for these subjects will be listed only.
4. For subjects incorrectly entered into the Randomized-Withdrawal or Escape Period, a listing for some efficacy variables (e.g. PASI and IGA) and listing for all TEAE will be provided.

3.7 Center pooling strategy

The subject treatment assignment is stratified by prior biologic exposure and region. Geographic regions have been 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 and corresponding countries

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

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 region, then the two regions will be combined as a geographic region stratum for all efficacy modeling, so that there are no modeling convergence issues across efficacy variables.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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 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 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.

3.10 Changes to protocol-defined analyses

The following variable was added as a secondary efficacy variable and added to the sequential testing procedure:

- IGA response (defined as Clear with at least a 2-category improvement relative to Baseline) at Week 16

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

- Palmoplantar Investigator's Global Assessment (pp-IGA) response (clear or almost clear with at least a two-category improvement from Baseline) for subjects with palmoplantar PSO at Baseline
- Change from Baseline in Patient Global Assessment (PGA) of PSO score

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

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- Percentage of subjects who rebound (defined as a ≥25% increase from Baseline in PASI score occurring within 2 months of stopping therapy) during the Randomized-Withdrawal Period
- 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%
- Change from Baseline in clinical laboratory values (chemistry and hematology) will be summarized, urinalysis results will be listed

The following additional changes were made:

- The calculation of nominal p-values has been added for selected of the 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.
- Subgroup analyses will be performed for PASI75/90/100 and IGA variables only and not all co-primary and secondary efficacy variables for the Initial Treatment Period that are part of the fixed sequence testing procedure.
- Prior primary failure to biologic (yes/no) was removed as a subgroup.
- Selected TEAE tables were added to include calculation of risk differences.
- Selected summaries of safety during the Escape Period will also be presented and will be based on the ESS.

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 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 nonresponse. Achieving the clinical response and making it through 16 weeks of

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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 continued with scheduled assessments, all efficacy data after discontinuation of study treatment 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 the 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, MI-MCMC/reference-based methods, last observation carried forward (LOCF) and observed case (OC) methods 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 - Section 8.1.3.3.

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, LOCF will be used. The OC method will be performed as a 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 the primary method.

For other continuous efficacy variables, the MI-MCMC/monotone regression method will be used to impute missing data as the primary method except for analysis based on the ES where the OC method will be used due to the short duration of the Escape Period and the impact of missing data in the period. If the imputation model cannot converge, LOCF will be used.

For other ordinal variables (shift in IGA score, EQ-5D-3L and Patient Global Assessment of PSO), the 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:

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- 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.
- 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 placebo (reference) data.
- 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]	S[a]	S	S
Secondary	Responder	P	S[a]			S
	Continuous		P		B	S
Other	Responder	P				S[b, c]
	Continuous[e]		P		B	
	Ordinal					P[d]

P=Primary method, S=Sensitivity method, B=Backup method, NRI-Nonresponder 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 responses, Scalp IGA, PASI75, PASI90, PASI100 and PSD of PSO for pain, itch and scaling.
[d] Includes Patient Global Assessment of PSO, IGA score and EQ-5D-3L responses.
[e] For PASE and analysis of the Escape Period, OC is the primary analysis method.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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 (i.e., 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 (i.e., where all subject data is missing after a given time point), monotone regression will then be used to impute missing data. This will be based on the100 sets of imputations already created using the MCMC method such that there will be 100 imputation 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 numbers of randomized subjects is less than 10% in either of the Asia/Australia or Western Europe region, 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,

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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 for values for secondary and 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
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 (e.g., 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 key points to consider relative to the calculation of the odds ratios and corresponding confidence intervals are noted below:

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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.

Further details about imputation for different analysis sets are 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 2 randomized treatment groups (placebo, Bimekizumab 320mg Q4W) including all scheduled assessment visits from Baseline to Week 16.
- Wk16ResS: When programming multiple imputation based on the Wk16ResS, PROC MI will be used with a separate data set for each of the 3 re-randomized treatment groups (Bimekizumab 320mg Q4W/Placebo, Bimekizumab 320mg Q4W/Q8W, Bimekizumab 320mg Q4W/Q4W) including all scheduled assessment visits from Week 16 to Week 56. Note: The Placebo/Placebo responders will not be included in these tables or in the MI procedure.

MI – MCMC / Reference-based imputation

In this case, placebo will be described as the reference arm.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

This procedure will use an imputation model developed based on data from the placebo group only (Mallinckrodt, 2013). Reference-based multiple imputation assumes that the statistical behavior of the bimekizumab and placebo-treated subjects after discontinuing study medication becomes that of the placebo-treated subjects. All time points after discontinuation of the double-blind study treatment for both the bimekizumab and placebo groups will be considered missing. Multiple imputations are used to replace missing outcomes for bimekizumab- and placebo-treated subjects who discontinued using multiple draws from the posterior predictive distribution estimated from the placebo arm. For binary efficacy variables (e.g., PASI90 at Week 16), imputation will be done on the raw PASI before assessing the imputed results for PASI90 response.

The steps for the procedure are as follows:

1. For non-monotone (intermittent) missing data, MCMC will be used to impute PASI, with biologic exposure, geographic region, and PASI values at Baseline and at each post-Baseline visit (in chronological order) being included in the imputation model. This will be done only once for each subject in order to provide a dataset with monotone missing data.
2. Data will be processed sequentially by repeatedly calling SAS® PROC MI to impute missing outcome data at visits t=1, … T, where T is Week 16 for PASI90.
    a. *Initialization*. Set t=1 (Baseline visit)
    b. *Iteration*. Set t=t+1. Create a data set combining records from bimekizumab- and placebo-treated subjects with columns for covariates (prior biologic exposure and geographic region) and outcomes at visits 1 to t. Outcomes for all bimekizumab-treated subjects are set to missing at visit t and set to observed or previously imputed values at visits 1 to t-1. Outcomes for placebo-treated subjects are set to observed at visit t or observed or previously imputed values at visits 1 to t-1.
    c. *Imputation*. Run MCMC to impute missing values for visit t using previous outcomes for visits 1 to t-1, prior biologic exposure, and geographic region. Note that only placebo data will be used to estimate the imputation model since no outcome is available for bimekizumab-treated subjects at visit t.
    d. Repeat steps 2a-2c, 100 times with different seed values (seeds ranging from 853 to 952) to create 100 imputed complete data sets.
    e. *Analysis*. For each completed data set, the outcome (response or change from Baseline) will be calculated using the complete datasets (with no missing data).
3. Each complete imputed data set will then be analyzed based on the statistical model specified in this study (stratified CMH test). The Week 16 results from stratified CMH test 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.

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).

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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, 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 on all available data at that time point (including all subjects who are in the Escape Period when the interim data cut is taken) and a corresponding interim clinical study report (CSR) will 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. Further details related to the DMC will be outlined in a separate analysis plan.

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

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 8.1.3.6) and adding a center-by-treatment interaction term. In the model, center will be based on the original centers prior to pooling (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 is still not achieved a pooling by region will be applied.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

4.5 Multiple comparisons/multiplicity

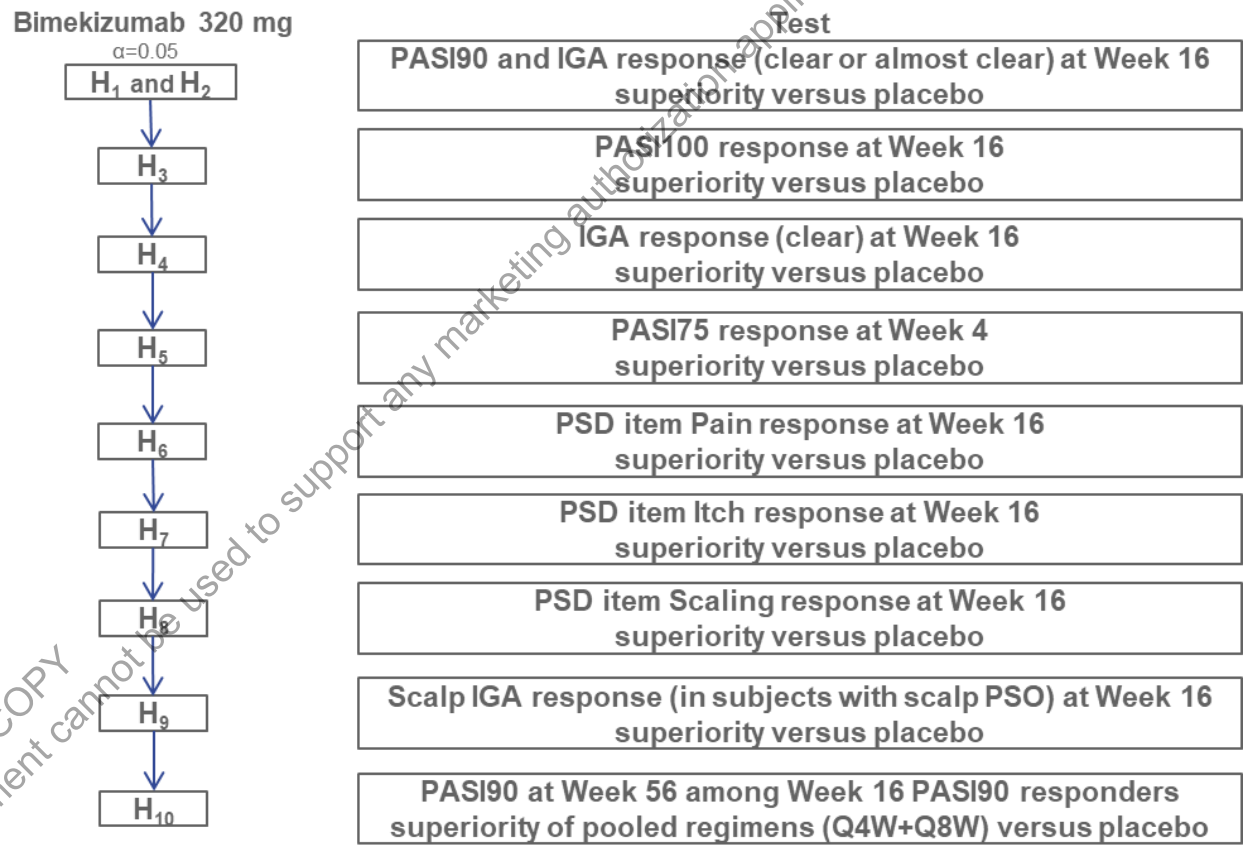
The statistical analysis of the co-primary efficacy variables and selected secondary efficacy variables will account for multiplicity and control the familywise Type I error rate at a 2-sided alpha level of 0.05 by using a fixed sequence testing procedure.

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

The first 2 hypotheses (H_1 and H_2) test whether bimekizumab is superior to placebo for PASI90 response and IGA response at Week 16. These are the hypothesis tests corresponding to the co-primary endpoints. If both are rejected at a 2-sided alpha level of 0.05, that alpha will be passed to the next test in the sequence, allowing the testing procedure to proceed.

The hypotheses associated with the subsequent tests are for secondary efficacy endpoints and are based on testing for superiority relative to placebo. The pairwise comparisons will follow the sequential testing sequence in Figure 4-1 and the formal evaluation of statistical significance of each comparison is dependent upon the previous comparison achieving statistical significance at the 2-sided alpha level of 0.05.

Figure 4-1: Sequence of testing



BL=Baseline; IGA=Investigator's Global Assessment; PASI=Psoriasis Area Severity Index, PSD=patient symptom diary; PSO=psoriasis; Q4W=every 4 weeks; Q8W=every 8 weeks.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

4.6 Use of an efficacy subset of subjects

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

4.7 Active-control studies intended to show equivalence

Not applicable.

4.8 Examination of subgroups

Subgroup analyses will be performed on PASI75/90/100 and IGA. 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, United Kingdom], Central/Eastern Europe [Hungary, Poland, Russian Federation], Asia/Australia [Australia, 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 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 chemotherapy or phototherapy will be classified as not receiving prior systemic treatment for psoriasis.

In addition, a subgroup analysis will be performed on PASI90/100 and IGA to assess predictability of future response at Week 4 and Week 16 using the following subgroups:

- PASI75 responders (OC) at Week 4 (yes, no) to predict PASI90/100 and IGA (NRI) through Week 56
- PASI90 responders (OC) at Week 16 (yes, no) to predict PASI100 (NRI) through Week 56

For responder analysis, the definition of subgroups of the PASI75/90/100 will be based on observed values. All summaries will be based on imputed data as appropriate and will include descriptive statistics only.

REDACTED COPY
This document cannot be used to support any marketing authorisation application and any extensions or variations thereof.

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 and ESS), disposition and discontinuation reasons in the Initial Treatment period (for RS) and in the Randomized-Withdrawal period (for WK16ResS) and Escape Treatment period (for ESS), as well as the subjects who discontinued due to AEs in the Initial Treatment period (for RS) and in the Randomized-Withdrawal period (for WK16ResS) and Escape Treatment period (for ESS) 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, ESS, WK16ResS, AMS, PPS, and PK-PPS) overall and by site.

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, using the RS, displaying the number and percentage of subjects with an important protocol deviation (including a summary of subjects excluded from the PPS or PK-PPS due to important protocol deviations) by treatment group in the Initial Treatment Period, the Randomized-Withdrawal Period, and in the Escape Treatment Period will be provided.

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 on SS, WK16ResS and ESS. If the RS and SS analysis sets are identical, the SS summary will not be produced.

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:

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

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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 (≤100 kg, >100 kg)
- BMI (<25 kg/m2, 25 to <30 kg/m2, ≥30 kg/m2)
- Gender
- Race
- Ethnicity

By-subject listings of demographics for all subjects screened 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:

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

[1]If the date of onset of plaque psoriasis is partial, it should be imputed to the most recent feasible date (i.e., 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, United Kingdom], Central/Eastern Europe [Hungary, Poland, Russian Federation], Asia/Australia [Australia, Korea])

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- Country
- Duration of disease (<median, ≥median)
- Investigator's Global Assessment (IGA) score
- Baseline disease severity (PASI<20, PASI≥20)
- Nail involvement (yes, no)
- Scalp involvement (yes, no)
- Palmoplantar involvement (yes, no)
- Prior biologic therapy (yes, no)
- Prior anti-TNF therapy (yes, no)
- Prior primary failure to biologic (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 received prior 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 groups, 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),

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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:

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

where the total number of expected injections is derived relative to when the subject finishes treatment. In this study, Baseline visit and Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52 are dosing visits. Two injections are administered at each given visit either with active dosing or placebo injection. It is expected that a subject should complete total 28 injections by the end of study, if they do not enter the open-label escape arm. If a subject enters and completes the open label escape arm, they will receive an additional 6 open-label injections during the Escape Treatment Period beyond what they have already received. 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 12 visit, the total number of expected injections will be 6.

A summary of percent treatment compliance categorized as <75% and ≥75% will be provided by treatment group and study periods (Initial Treatment Period for the RS, Randomized-Withdrawal Period for the WK16ResS, Escape Treatment Period for the ESS, and the Initial and Randomized-Withdrawal Period for the AMS).

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

8 EFFICACY ANALYSES

All efficacy analyses of primary and secondary variables will be performed on the RS unless otherwise specified. All efficacy analyses of other efficacy variables will be performed on the RS, WK16ReS and ESS unless otherwise specified. All efficacy summary tables will be displayed by treatment group unless otherwise specified.

8.1 Statistical analysis of the primary efficacy variables

8.1.1 Derivations of co-primary efficacy variables

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%.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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, T, S) 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.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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. Pairwise treatment comparisons will be made 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 provided. If one of the treatment groups has zero or very low response where the CMH odds ratio can no longer be calculated, the logit method will be applied instead.

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.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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. If one of the treatment groups has zero or very low response where the CMH odds ratio can no longer be calculated, the logit method will be applied instead as for the primary analyses. 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

Deviations from the MAR pattern assumed above will be evaluated. A sensitivity analysis will be performed in which, as with the above analysis, intermittent missing data will be imputed using MI based on the MCMC method. However, the remaining monotone missing data will be assumed to follow a missing not at random (MNAR) pattern. These data will be imputed using reference-based imputation in which the imputation model is based on data from the placebo group (see Section 4.2.1.6), thereby assuming that monotone missing data follow a trajectory similar to the placebo group (Mallinckrodt et al, 2012). As specified in the previous procedure, actual PASI/IGA scores will be imputed and then dichotomized to get 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. If one of the treatment groups has zero or very low response where the CMH odds ratio can no longer be calculated, the logit method will be applied instead as for the primary analyses. The estimand in this procedure is the difference in outcome improvement in all randomized subjects at the planned endpoint of the study attributable to the initially randomized medication (Mallinckrodt et al, 2012). This is an estimand of effectiveness to evaluate the de facto hypothesis.

REDACTED COPY
This document contains data to support any marketing authorization application and extensions or variations thereof.

8.1.3.3 Sensitivity Analysis #3

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 4.2.1.5). The same stratified CMH test as in the primary efficacy analysis will be used.

8.1.3.4 Sensitivity Analysis #4

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

8.1.3.5 Sensitivity Analysis #5

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

8.1.3.6 Sensitivity Analysis #6

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.

8.1.3.7 Sensitivity Analysis #7

The center-by-treatment interaction will be tested by replacing region with center in the logistic regression model described above 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 applied 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 4:1 randomization allocation scheme, this should provide a minimum of about 12 subjects in the bimekizumab arm and 3 subjects in the placebo arm. 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, prior biologic exposure may be dropped from the model. If convergence still cannot be achieved, this analysis will not be performed.

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 the 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. The 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

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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

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 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.

8.2.1.2 IGA Response at Week 16

A categorical response variable, IGA response (Clear with at least 2-category improvement from Baseline) is defined as an IGA score of zero (Table 8-1) with at least 2-category improvement from Baseline.

8.2.1.3 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.

8.2.1.4 Patient Symptom Diary (PSD)

A patient reported outcome (PRO) measure, the PSD, will be used to assess key symptoms relevant to subjects with moderate to severe chronic plaque PSO. The ePRO diary will be administered on a daily basis from Screening to the Week 16 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 16. A weekly average is the sum of the scored item over the course of the study week divided by the number of days on which the item was completed. The weekly averages should be relative to the respective visit date except for baseline which should follow the usual convention and should be based on at least 4 non-missing values (do not need to be consecutive). Otherwise, the PSD will be set to missing. The Baseline value will be computed in the same manner.

The PSD will be computed based on the responder definition. Each of the 3 PSD response scores included in the statistical testing procedure (itch, pain, and scaling) will be characterized in terms of the cumulative percent of subjects demonstrating a pre-specified point improvement at Week 16. The threshold for the PSD response score of itch, pain, and scaling is 2.39, 1.98, and 2.86, respectively. The responder analysis will be limited to the subjects with a Baseline PSD response

REDACTED CSP
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

score at or above the applicable threshold score (i.e., at least 2.39, 1.98, and 2.86 for itch, pain, and scaling, respectively).

The derivations for absolute change from Baseline is as follows:

- Change from Baseline in PSD responses for itch, pain, and scaling at Week 16 is defined as Week 16 PSD response minus Baseline PSD response.

8.2.1.5 Scalp IGA Response (Clear or Almost Clear) at Week 16

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

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 schedule of assessments. 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 the Table 8-2 below. This is the same scale with corresponding descriptors used for the overall IGA, but this will be specific to the scalp.

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

8.2.1.6 PASI90 Response at Week 56 among Week 16 PASI90 Responders

A categorical response variable, PASI90 at Week 56 is defined to be equal to 1 if the percentage improvement from Baseline to Week 56 in the PASI scores is 90% or greater and 0 if the percentage improvement from Baseline to Week 56 is less than 90%. This definition is introduced for the purpose of identifying subjects who respond to the treatment (1 = responder, 0

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

= non-responder). The definition of percentage improvement from Baseline is given in Section 8.1.1.1.

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), unless otherwise specified.

8.2.2.1 PASI responses

The secondary efficacy variables for time points during the initial Treatment Period will be analyzed for all subjects in the RS.

For binary secondary variables, including PASI100 at Week 16 and PASI75 at Week 4, the stratified CMH test as specified for the primary analysis will be implemented to test for superiority. As in the primary analysis, NRI will be used to account for missing data.

PASI90 response at Week 56 for subjects in the WK16ResS will be evaluated using a stratified CMH test where region and prior biologic exposure will be used as stratification variables, similar to the approach specified for the primary analysis. In this analysis, both dose regimens (Q4W and Q8W) within a given bimekizumab dose will be pooled such that each bimekizumab dose (regardless of regimen) will be compared to placebo (based on subjects who have switched to placebo from bimekizumab at Week 16, to evaluate the efficacy of continuous treatment with bimekizumab versus treatment withdrawal (placebo)). Missing data will be handled via NRI. Specifically, if a subject has missing data at Week 56, has discontinued from study treatment, or has met the criterion for relapse during the Randomized-Withdrawal Period, they will be treated as a non-responder at Week 56 for this analysis.

While not part of the fixed sequence testing procedure, it is a secondary objective to assess the maintenance of efficacy of bimekizumab dosing Q4W versus Q8W at Week 56. While multiple variables and analyses will be used to assess this objective, the secondary efficacy variable that will be used for this assessment is also PASI90 response at Week 56 for subjects in the WK16ResS. This analysis will be based on simple percentages of PASI90 responders at Week 56 for each dosing regimen (Q4W and Q8W). Subjects with missing data at Week 56 or who meet the criterion for relapse will be counted as non-responders.

8.2.2.2 IGA Response

For IGA response (clear) at Week 16, the stratified CMH test similar to the primary analysis will be applied.

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

8.2.2.3 Scalp IGA response

For scalp IGA (Clear or Almost Clear) at Week 16, the stratified CMH test similar to the primary analysis will be applied.

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

REDACTED COPY
This document cannot be used to support any marketing authorisation application and any extensions or variations thereof.

8.2.2.4 PSD

For each of the three PSD response scores of itch, pain, and scaling, the following analyses will be performed for the purpose of supporting each item as individual key secondary variables.

The PSD response variables will be characterized in terms of the cumulative percent of subjects demonstrating a pre-specified point improvement at Week 16. The threshold for the PSD response score of itch, pain, and scaling is 2.39, 1.98, and 2.86, respectively. This responder analysis will be limited to the subjects with a Baseline PSD response score at or above the applicable threshold score (i.e., at least 2.39, 1.98, and 2.86 for itch, pain and scaling respectively). The stratified CMH test, similar to the primary efficacy analysis, will be applied to each of these responder analyses. These are the inferential analyses that will be used in the multiplicity-controlled testing procedure for itch, pain, and scaling.

Additionally, the entire distribution of responses for each of PSD response scores of itch, pain and scaling will be evaluated via continuous plots showing the absolute and percent change from Baseline on the X-axis and the percent of subjects experiencing that change on the Y-axis. This distribution curve will reveal the extent to which overall results are driven by outliers who improve or worsen more than others. The cumulative distribution function (CDF) is the probability that the variable takes a value less than or equal to x. That is:

$$x \mapsto F_X(x) = \mathrm{P}(X \leq x)$$

where the right-hand side represents the probability that the random variable X takes on a value less than or equal to x. The probability that X lies in the interval $[a, b]$ is therefore:

$$F_X(b) - F_X(a) \text{ if } a < b$$

Note that cumulative distribution plots will be provided for absolute and percent change from Baseline PSD to Week 16 for itch, pain, and scaling.

8.2.3 Sensitivity analyses of the secondary efficacy variables

Sensitivity analyses for the handling of missing data will be performed for the binary and continuous secondary efficacy variables that are part of the fixed sequence testing procedures in Figure 4-1.

For binary response variables (PASI100 at Week 16, PASI75 at Week 4, Scalp IGA at Week 16, and PSD response for itch, pain and scaling), sensitivity analyses #1 and #3 (Section 8.1.3.1, Section 8.1.3.3) will be performed.

Missing data for these PSD variables will use the MI procedure similar to the sensitivity analysis #1 (Section 8.1.3.1) for the co-primary efficacy variables with the following differences:

1. The data included in the imputation model will be limited only to those subjects with sufficiently high Baseline values to achieve the response for a given item (based on observed cases).
2. The PSD item value will be imputed. Once imputation has been performed across the 100 iterations specified, any values below 0 and above 10 will be truncated to 0 and 10, respectively.
3. The change from Baseline values will be computed based on the complete data sets.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

The analysis model will be based on the CMH test.There may be cases where the multiple imputation model fails to converge. In such situations, the LOCF approach will instead be used to impute the missing data. If LOCF is used instead of multiple imputation for this reason, this will be clearly specified in the corresponding table summary. Note that for LOCF imputation, any missing data or data collected following discontinuation of IMP will be imputed with the most recent previous value.

Missing data for binary secondary efficacy variables during the Initial Period will be imputed using NRI.

Missing data for the binary secondary efficacy variable during the Randomized-Withdrawal Period (PASI90 at Week 56 among Week 16 PASI90 responders) will be handled via NRI. Specifically, if a subject has missing data at Week 56, has discontinued from study treatment, or has met the criterion for relapse during the Randomized-Withdrawal Period, they will be treated as a non-responder at Week 56 for this analysis.

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 other efficacy variables will be summarized based on imputed data (NRI and MCMC/Monotone Regression for binary and continuous variables, respectively), unless otherwise specified.

Missing data for continuous secondary efficacy variables will be imputed using MI. The MI procedure for continuous variables will be similar to sensitivity analysis #1 described in Section 8.1.3.1 for the co-primary efficacy endpoints with the following difference – no dichotomization will be necessary. Also, no requirement of meeting minimum threshold at baseline is needed to be included in imputation for a continuous variable

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 LOCF approach will instead be used to impute the missing data. If LOCF is used instead of multiple imputation for this reason, this will be clearly specified in the corresponding table summary. Should there be no missing data for a study variable, then only observed case data will be produced. Note that for LOCF imputation, any missing data or data collected following discontinuation of IMP will be imputed with the most recent previous value.

Note that in addition, for subjects who change treatment at Week 16, change from two Baselines (i.e Initial Treatment Period Baseline and Randomized Withdrawal Period Baseline) will be summarized in the WK16ResS for PASI and DLQI.

For subjects that enter the escape arm at Week 16 or later, change from two Baselines (i.e. Initial Period Baseline and Baseline at which escape treatment was initiated) will be summarized in the ESS for PASI and DLQI.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

8.3.1 PASI

8.3.1.1 PASI50, PASI75, PASI90, and PASI100 response

Binary 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 percentage improvement from Baseline is given in Section 8.1.1.1.

A line plot of the PASI responder (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

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 including unscheduled visits are considered.

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 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 the Week 16 Visit. Subjects will be censored at Baseline if there is no Baseline PASI assessment or no Post Baseline PASI assessment.

Time to PASI50, PASI75, PASI90, and PASI100 response during Initial Treatment Period 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 be presented by treatment group. In these Kaplan-Meier plot, 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 bimekizumab vs placebo will be analyzed using a log-rank test stratified by region and prior biologic exposure.

8.3.1.3 Time to relapse

Relapse is defined as not achieving a PASI75 response at Week 20 or later during the Randomized-Withdrawal Period for subjects in the WK16ResS.

Time to relapse (in days) is defined as: Date of relapse - Date of first re-randomized IMP treatment + 1.

Time to relapse will be estimated and presented using the Kaplan-Meier product-limit method for each treatment (excluding subjects who were treated with placebo in both Initial Treatment Period and Randomized-Withdrawal Period as they cannot relapse by definition). Subjects who discontinue IMP prior to a recorded relapse will be considered as having relapsed on the date of

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

early discontinuation. Subjects who reach the Week 56 Visit without relapsing will be censored at the date of the Week 56 Visit.

The median time to relapse, including the two-sided 95% CI, will be calculated for each treatment. Between-group differences will be analyzed with the log-rank test stratified by region and prior biologic exposure. These between-group differences will compare each bimekizumab dose (separately for each regimen as well as pooled regimen) to placebo based on the re-randomized treatment groups (i.e., independent of the initial randomization).

8.3.1.4 Percentage of subjects who relapse

Percentage of subjects who relapse during the Randomized-Withdrawal Period is defined as: (Number of subjects who relapse / Number of subjects in WK16ResS) × 100

The percentage of subjects who relapse will be presented by treatment group over time (excluding subjects who were treated with placebo in both Initial Treatment Period and Randomized-Withdrawal Period as they cannot relapse by definition)..

8.3.1.5 Percentage of subjects who rebound

Rebound is defined as when a subject experiences a >=25% increase from Initial Treatment Period Baseline in PASI score occurring within 2 months (60 days) of stopping therapy. The percentage of subjects who rebound during the Randomized-Withdrawal Period is defined as:

(Number of subjects who rebound / Number of subjects in WK16ResS) × 100

The percentage of subjects who rebound will be presented for subjects re-randomized to placebo over time.

8.3.1.6 PASI score

Absolute and percentage 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 using NRI.

8.3.1.7 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 (i.e., 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.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

An mNAPSI 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.

8.3.2 IGA response

- IGA response (Clear with at least 2-category improvement from Baseline) is defined as an IGA score of zero (Table 8-1) with at least 2-category improvement from Baseline. Shift from Baseline in IGA score is defined at each Post-Baseline visit timepoint relative to Baseline.
- IGA response (Clear or Almost Clear with at least 2-category improvement from Baseline) is defined as an IGA score of zero (Clear) or an IGA score of 1 (Almost Clear) with at least a 2-category improvement from Baseline.
- Scalp-specific IGA response (Clear or Almost Clear with at least 2-category improvement from Baseline) definition and derivation are outlined in Section 8.2.1.4.
- Palmoplantar Investigator's Global Assessment (pp-IGA) response (Clear or Almost Clear with at least 2-category improvement from Baseline)

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

Only subjects with scalp 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 the table below.

Table 8–4: 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,

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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 scalp IGA analysis.

A line plot of the IGA (both clear or almost clear and clear) and scalp 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 patients 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 8–3: 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

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

The DLQI 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).

A DLQI absolute score of 0 or 1 indicates no or small impact of the disease on health-related quality of life.

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.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- Percent of subjects achieving a minimal clinically important difference (MCID) 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 that have a Baseline DLQI of at least 4.

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 the PGADA for the arthritis VAS in subjects with PsA at Baseline is defined as Post-Baseline PGADA minus Baseline PGADA.

8.3.7 Patient Global Assessment of PSO score

The Patient Global Assessment of PSO is a PSO-specific item in which the patient responds to the multiple-choice question, "How severe are your psoriasis-related symptoms right now?" Possible response 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)

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

For PSD response score of redness, burning, cracking, dryness, irritation, sensitivity, lesions, thickening, fatigue, embarrassment, choice of clothing, the responder criterions in terms of change from Baseline are set to 3.05, 1.99, 2.01, 3.04, 2.82, 2.60, 2.69, 2.68, 1.51, 2.43, and 2.14 respectively. Number and percentage of subjects who are responders will be summarized based on each of these items.

In addition, cumulative distribution plots will be provided for absolute and percent change from Baseline PSD at Week 16 for each item.

8.3.9 Psoriatic Arthritis Screening and Evaluation (PASE)

The PASE questionnaire is a self-administered tool to screen for active PsA in patients 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.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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.

8.3.10 Short Form 36-item 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 updated 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.

8.3.11 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 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.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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.

8.4 Additional statistical analysis of other efficacy variables

For selected of the other 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 Placebo (Weeks 1, 2, 4, 8, and 12)
- IGA Clear or Almost Clear
  - Bimekizumab vs Placebo (Weeks 1, 2, 4, 8, and 12)
- PASI100
  - Bimekizumab vs Placebo (Weeks 4, 8, and 12)
- IGA Clear
  - Bimekizumab vs Placebo (Weeks 4, 8, and 12)
- PASI75
  - Bimekizumab vs Placebo (Weeks 1, 2, and 16)
- PSD Response – Pain
  - Bimekizumab vs Placebo (Weeks 1, 2, and 4)
- PSD Response – Itch
  - Bimekizumab vs Placebo (Weeks 1, 2, and 4)
- PSD Response – Scaling
  - Bimekizumab vs Placebo (Weeks 1, 2, and 4)
- Scalp IGA Clear or Almost Clear (subjects with Baseline Scalp IGA ≥2)
  - Bimekizumab vs Placebo (Weeks 1, 2, and 4)

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- pp-IGA Clear or Almost Clear (subjects with Baseine pp-IGA ≥2)
  - Bimekizumab vs Placebo (Week 16)
- mNAPSI change from Baseline (subjects with Baseline mNAPSI>0)
  - Bimekizumab vs Placebo (Week 16)
- mNAPSI75 response (subjects with Baseline mNAPSI>0)
  - Bimekizumab vs Placebo (Week 16)
- mNAPSI90 response (subjects with Baseline mNAPSI>0)
  - Bimekizumab vs Placebo (Week 16)
- mNAPSI100 response (subjects with Baseline mNAPSI>0)
  - Bimekizumab vs Placebo (Week 16)
- DLQI change from Baseline
  - Bimekizumab vs Placebo (Week 16)
- DLQI 0/1 response
  - Bimekizumab vs Placebo (Week 16)
- SF-36 PCS change from Baseline
  - Bimekizumab vs Placebo (Week 16)
- SF-36 MCS change from Baseline
  - Bimekizumab vs Placebo (Week 16)
- PASI percentage change from Baseline
  - Bimekizumab vs Placebo (Weeks 1, 2, 4, 8, 12, and 16)
- PASI ≤2 response
  - Bimekizumab vs Placebo (Weeks 1, 2, 4, 8, 12, and 16)
- PASI ≤5 response
  - Bimekizumab vs Placebo (Weeks 1, 2, 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

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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 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. In addition, if the PK sampling date is >1 day after the dosing date, then the plasma concentration from that visit will be excluded from the PK summary.

However, all PK concentrations will also 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 Immunogenicity

The analysis of immunogenicity will be performed on SS.

9.2.1 Auto-antibodies

Not applicable.

9.2.2 Anti-bimekizumab antibodies

Anti-bimekizumab antibodies (ADAb) will be measured using a 3-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 (IL-17A or IL-17F, 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 open label extension study (PS0014) a sample at SFU (20 weeks after the last dose) . For the interim analysis, the SFU visit (after Week 56) will be excluded in the analyses.

In subjects randomized to placebo and have a PASI90 response and remain on placebo for 56 weeks, no samples will be analyzed or reported.

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 ADAb in test sample as above the cut point (ACP) or below the cut point (BCP). Samples presenting ADAb 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

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- 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 (IL-17A or IL-17F 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 initial 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 initial 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-initial 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 initial 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 initial treatment (including SFU) with titer values of the same magnitude as baseline (i.e. ≤ then 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 initial 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- initial treatment samples are missing, while other post-initial treatment samples are ADAb negative, **the subject will be classed as inconclusive.**

Derivation for above classification will be different for the interim analysis and the final analysis. For the interim analysis no SFU data will be considered. Only for the final analysis, when all SFU data will be available, data from the SFU visits will be considered.

Analysis

Immunogenicity will be assessed through summary tables and figures, and listing of individual results by subject. All analyses will be run on the SS, WK16ResS, ESS and AMS.

- 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.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- 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 in categories 2 and 5, summarized as total treatment emergent. In addition, the counts and percentage of subjects who are pre-anti-bimekizumab positive will be calculated (this is the sum of categories 3,4, and 5).
- 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 responders, 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 (excluding SFU for interim analyses and including SFU for final analyses), where a subject has not progressed into the open label extension study (PS0014). 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

10 SAFETY ANALYSES

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

The AMS will be used for summaries of safety that include data from the Initial Treatment Period and/or Randomized-Withdrawal Period.

Additional summaries focusing on safety during the Randomized-Withdrawal Period will also be prepared and will be based on the WK16ResS. Selected summaries of safety during the Escape Period will also be presented and will be based on the ESS.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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., the Initial Treatment Period, the Randomized-Withdrawal Period, and Initial and Randomized-Withdrawal Period). The cumulative study medication duration will be summarized for subjects exposed for given durations of time, the following categories for duration will be used:

- >0 weeks; ≥ 4 weeks; ≥8 weeks; ≥ 12 weeks; ≥ 16 weeks for Initial Treatment Period

>0 weeks; ≥ 4 weeks; ≥8 weeks; ≥ 12 weeks; ≥ 16 weeks; ≥ 20 weeks; ≥ 40 weeks for Randomized-Withdrawal Period >0 weeks; ≥ 4 weeks; ≥8 weeks; ≥ 12 weeks for Escape Period 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 Period are provided as follows:

Study medication duration (days)

- Date of last dose in the Initial Treatment Period – date of first dose in the Initial Period + 28 days.

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

  - Date of first dose in the Randomized-Withdrawal Period – date of first dose in the Initial Period + 1.

  Note: For subjects who die during the Initial Period, if date of last dose in the Initial 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 Period + 1.

Time at risk (days)

- For subjects who complete the final visit of the Initial Period and continue to the Randomized-Withdrawal Period: Date of first dose in the Randomized-Withdrawal Period – date of first dose in the Initial Period + 1.
- For subjects who discontinue on or prior to the final visit of the Initial Period, use the minimum of the following:
  - Date of last dose in the Initial Treatment Period – date of first dose in the Initial Period + 140 days.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- The total number of days in the Initial Treatment Period (112 days). For AEs that emerged after 112 days but still within the 140 days window, those AEs would be classified as TEAE, but will be excluded from the output based on the Initial Treatment Period. However these AEs will be included in the AE summaries for Randomized-Withdrawal Period or Escape Period for subjects in the WK16ResS and ESS respectively.
- 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 the Initial Period +1.

10.1.2 Exposure during the Randomized-Withdrawal Period for WK16ResS

Definitions for study medication duration (days) and time at risk (days) during the Initial Period are provided as follows:

Study medication duration (days)

- Date of last dose in the Randomized-Withdrawal Period – date of first dose in the Randomized-Withdrawal Period + 28 days.
- Note: The use of 28 days assumes a Q4W dosing interval. This should be adjusted based on the dosing interval (i.e., use 56 days for a Q8W dosing interval).

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

  - Final visit date of the Randomized-Withdrawal Period (not including SFU) – date of first dose in the Randomized-Withdrawal Period + 1.

    Note: If date of last dose in the Randomized-Withdrawal Period + 28 days (or 56 days in the case of Q8W dosing) extends to a date beyond the date of the first dose in the 12-Week Escape Period, then this calculation reverts to:

    Date of first dose in the 12-Week Escape Period – date of first dose in the Randomized-Withdrawal Period + 1.

  Note: For subjects who die during the Randomized-Withdrawal Period, then this calculation reverts to:

  - Date of death – date of first dose in the Randomized-Withdrawal Period + 1.

Time at risk (days)

- For subjects who complete the Randomized-Withdrawal Period as planned and continue into an open-label study (and, therefore, do not have the SFU visit in the feeder study): Final visit date of the Randomized-Withdrawal Period – date of first dose in the Randomized-Withdrawal Period + 1.
- For subjects who die prior to the final visit of the Randomized-Withdrawal Period: Date of death – date of first dose in the Randomized-Withdrawal Period + 1.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- For all other subjects, use the minimum of the following:
    - Date of last dose in the Randomized-Withdrawal Period – date of first dose in the Randomized-Withdrawal Period + 140 days.
    - Date of first dose in the 12-Week Escape Period – date of first dose in the Randomized-Withdrawal Period + 1.
    - Date of last clinical contact – date of first dose in the Randomized-Withdrawal Period + 1.

  Note: This group could include subjects who discontinue the Randomized-Withdrawal Period early, subjects who complete the Randomized-Withdrawal 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.

10.1.3 Exposure during the Escape Period for ESS

There are some summaries for exposure and TEAEs that are based only on the Escape Period. Definitions for study medication duration (days) and time at risk (days) during the Escape Period are provided as follows:

Study medication duration (days)

- Date of last dose in the Escape Period – date of first dose in the Escape Period + 28 days.

  Note: If date of last dose in the Escape Period + 28 days extends to a date beyond the date of last visit in the Escape Period, then this calculation reverts to:
    - Date of last visit in the Escape Period – date of first dose in the Escape Period + 1.

  Note: For subjects who die during the Escape Period, if date of last dose in the Escape 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 Escape Period + 1.

Time at risk (days)

- For subjects who complete the final visit of the Escape Period: Date of last visit in the Escape Period – date of first dose in the Escape Period + 1.
- For subjects who discontinue on or prior to the final visit of the Escape Period, use the minimum of the following:
    - Date of last dose in the Escape Period – date of first dose in the Escape Period + 140 days.
    - The total number of days in the Escape Period (84 days).
    - Date of last clinical contact – date of first dose in the Escape Period + 1.
- For subjects who die prior to the final visit of the Escape Period: Date of death – date of first dose in the Escape Period +1.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

10.1.4 Exposure during the Initial and Randomized-Withdrawal for AMS

Definitions for study medication duration (days) and time at risk (days) entire study treatment period (the Initial and the Randomized-Withdrawal Period) are provided as follows:

Study medication duration (days)

For subjects who do not switch study treatments and do not enter the 12-Week Escape Period:

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

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

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

  Note: For subjects who die, 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 + 1.

For subjects who switch study treatments (between Initial and Randomized-Withdrawal Periods) and do not enter the 12-Week Escape Period:

- Initial Period (attributed to initially randomized treatment):
  - Date of last dose in the Initial Period – date of first dose in the Initial Period + 28 days.

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

    - Date of first dose in the Randomized-Withdrawal Period – date of first dose in the Initial Period + 1.
- Randomized-Withdrawal Period (attributed to the treatment initiated in the Randomized-Withdrawal Period):
  - Date of last dose in the Randomized-Withdrawal Period – date of first dose in the Randomized-Withdrawal Period + 28 days.

    Note: The use of 28 days assumes a Q4W dosing interval. This should be adjusted based on the dosing interval (i.e., use 56 days for a Q8W dosing interval).

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

    - Final visit date of the Randomized-Withdrawal Period (not including SFU) – date of first dose in the Randomized-Withdrawal Period + 1.

    Note: For subjects who die during the Randomized-Withdrawal Period, then this calculation reverts to:

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- Date of death – date of first dose in the Randomized-Withdrawal Period + 1.

For subjects who enter the 12-Week Escape Period due to not achieving a PASI90 response at Week 16:

- Initial Period (attributed to initially randomized treatment):
  - Date of last dose in the Initial Period – date of first dose in the Initial Period + 28 days.

    Note: If date of last dose in the Initial Treatment Period + 28 days extends to a date beyond the date of first dose in the 12-Week Escape Period, then this calculation reverts to:
    - Date of first dose in the 12-Week Escape Period – date of first dose in the Initial Period + 1.

For subjects who enter the 12-Week Escape Period due to relapsing during the Randomized Withdrawal Period:

- Initial Period (attributed to initially randomized treatment):
  - Date of last dose in the Initial Period – date of first dose in the Initial Period + 28 days.

    Note: If date of last dose in the Initial Treatment Period + 28 days extends to a date beyond the date of first dose in the Randomized Withdrawal Period, then this calculation reverts to:
    - Date of first dose in the Randomized Withdrawal Period – date of first dose in the Initial Period + 1.
- Randomized-Withdrawal Period (attributed to the treatment initiated in the Randomized-Withdrawal Period):
  - Date of last dose in the Randomized-Withdrawal Period – date of first dose in the Randomized-Withdrawal Period + 28 days.

    Note: The use of 28 days assumes a Q4W dosing interval. This should be adjusted based on the dosing interval (i.e., use 56 days for a Q8W dosing interval).

    Note: If date of last dose in the Randomized-Withdrawal Period + 28 days (or 56 days in the case of Q8W dosing) extends to a date beyond the date of the first dose in the 12-Week Escape Period, then this calculation reverts to:
    - Date of first dose in the 12-Week Escape Period – date of first dose in the Randomized-Withdrawal Period + 1.

<u>Time at risk (days)</u>:

For subjects who do not switch study treatments and do not enter the 12-Week Escape Period:

- For subjects who complete the Randomized-Withdrawal Period as planned and 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.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- For subjects who die prior to the final visit: Date of death – date of first dose in the + 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 Randomized-Withdrawal 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 (in the case of the interim analysis).

For subjects who switch study treatments (between Initial and Randomized-Withdrawal Periods) and do not enter the 12-Week Escape Period:

- Initial Treatment Period (attributed to initially randomized treatment):
    - Date of first dose in the Randomized-Withdrawal Period – date of first dose in the Initial Period + 1. (Note: This assumes that anyone in this category has completed the Initial Treatment Period and doses [with a new study treatment] in the Randomized-Withdrawal Period.)
- Randomized-Withdrawal Period (attributed to the treatment initiated in the Randomized-Withdrawal Period):
    - For subjects who complete the Randomized-Withdrawal Period as planned and continue into an open-label study (and, therefore, do not have the SFU visit in the feeder study): Final visit date of the Randomized-Withdrawal Period – date of first dose in the Randomized-Withdrawal Period + 1.
    - For subjects who die prior to the final visit of the Randomized-Withdrawal Period: Date of death – date of first dose in the Randomized-Withdrawal Period + 1.
    - For all other subjects, use the minimum of the following:
        - Date of last dose in the Randomized-Withdrawal Period – date of first dose in the Randomized-Withdrawal Period + 140 days.
        - Date of last clinical contact – date of first dose in the Randomized-Withdrawal Period + 1.

      Note: This group could include subjects who discontinue the Randomized-Withdrawal Period early, subjects who complete the Randomized-Withdrawal 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 who enter the 12-Week Escape Period due to not achieving a PASI90 response at Week 16:

- Initial Treatment Period (attributed to initially randomized treatment):
    - Date of first dose in the 12-Week Escape Period – date of first dose in the Initial Period + 1. (Note: This assumes that anyone in this category has completed the

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Initial Treatment Period and doses [with BKZ 320mg Q4W] in the 12-Week Escape Period.)

For subjects who enter the 12-Week Escape Period due to relapsing during the Randomized Withdrawal Period:

- Initial Period (attributed to initially randomized treatment):
    - Date of first dose in the Randomized-Withdrawal Period – date of first dose in the Initial Period + 1. (Note: This assumes that anyone in this category has completed the Initial Treatment Period and doses [with a new study treatment] in the Randomized-Withdrawal Period.)
- Randomized-Withdrawal Period (attributed to the treatment initiated in the Randomized-Withdrawal Period):
    - Date of first dose in the 12-Week Escape Period – date of first dose in the Randomized-Withdrawal Period + 1.

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 from placebo to bimekizumab at Week 16 or who relapse on placebo during the Randomized-Withdrawal Period and enter the Escape Period, AEs will be allocated to the treatment the subject was on the day the event occurred. If the event occurs on the day of the switch, it will be attributed to placebo. 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 12.1)
- Events with a reported term containing the term “hypersensitivity”
- 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 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 the Section 4.2.2.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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:

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

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:

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

where N_{AE} is the total number of AEs.

No confidence interval will be computed for EAER.

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

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

where IP_{BKZ} is the incidence proportion for the bimekizumab-treated group and IP_{PBO} is the incidence proportion for the placebo group. Note that incidence proportion simply refers to the

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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:

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

where n_{BKZ} is the number of subjects in the bimekizumab-treated group and n_{PBO} is the number of subjects in the placebo 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). The risk difference and corresponding CI will be displayed as percentage.

10.2.2 AE Summaries

The following summaries will be provided by treatment group for the Initial Treatment Period, Randomized-Withdrawal Period, and the Initial and Randomized-Withdrawal Period combined based on the SS, the WK16ResS, and the AMS, respectively. In addition, all summaries of TEAEs based on "100 subject years" will include EAIR (with 95% confidence interval) and EAER. Additional summaries for the ESS will include only the TEAE overview, TEAEs and SAEs per 100 subject years, and TEAEs tables for safety topics of interest. For AEs that emerged after 112 days but still within the 140 days window, those AEs would be classified as TEAE. These AEs will be excluded from the outputs based on the Initial Period but included in the AE summaries for Initial and Randomized-Withdrawal Period.

- 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 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

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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

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 placebo).

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 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 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 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").

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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 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)".

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

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

(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 event

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

Adjudicated 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 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

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

10.2.3.5 Suicidal Ideation and Behavior

- **Incidence of Suicidal Ideation or Behavior 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 a 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.

10.2.3.6 Inflammatory bowel disease

- **Incidence 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". The table will be 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 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 (refer to Appendix A in Section 12) for acute anaphylactic events (reported on either the same day as when an injection was administered or one day after).

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".

10.2.3.8 Hepatic events and DILI

- **Incidence 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.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

A by-subject listing of all AEs of safety topics of interest will be presented by type of safety topics of interest.

10.3 Clinical laboratory evaluations

Laboratory values, including markedly abnormal laboratory values will be presented descriptively by treatment group for the SS, WK16ResS and AMS. In addition, the ESS will be summarized for only markedly abnormal laboratory values and Hy's Law outputs.

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 minimum of period of interest (Week 16) or date of last dose + 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
- A by-subject listing of all laboratory data (including urinanalysis) 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 2017). 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 liver function test values, Table 10-2 for markedly abnormal biochemistry values and Table 10-3 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

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

of study medication are not considered. The laboratory results classified as Grade 3 or Grade 4 will be summarized and listed separately.

Table 10–1: Definitions of Markedly Abnormal Liver Function Values

Parameter name	Conventional		Standard		Abnormal Designation
	Unit	Criteria	Unit	Criteria	
Alkaline Phosphatase	U/L	>5.0 x ULN	U/L	>5.0 x ULN	AH
ALT	U/L	>5.0 x ULN	U/L	>5.0 x ULN	AH
AST	U/L	>5.0 x ULN	U/L	>5.0 x ULN	AH
Total Bilirubin	mg/dL	>3.0 x ULN	umol/L	>3.0 x ULN	AH
GGT	U/L	>5.0 x ULN	U/L	>5.0 x ULN	AH

Table 10–2: 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 10–3: 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

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Parameter name	Conventional		Standard		Abnormal Designation
	Unit	Criteria	Unit	Criteria	
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 table for markedly abnormal liver function tests (LFTs) will contain data beyond the CTCAE Grade 3 thresholds outlined in Table 10-1 above in order to allow for a more thorough review of elevated LFTs. There will be one table which will list the count and percentage of subjects meeting the below criteria at any time during the study:

- 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 two definitions 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, if applicable) 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 summaries will be provided for the SS, WK16ResS, and AMS:

- A summary of the absolute and change from Baseline value for each vital sign variable by treatment group and visit
- 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 10-4, by treatment group and visit.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Table 10–4: 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 will 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 for SS.

10.4.3 Electrocardiograms

Electrocardiogram data will be analyzed by treatment group and visit for the SS, WK16ResS and AMS.

A summary of the number and percentage of subjects with normal, 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 listing 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 listing of all 12-lead ECG data will be provided based on interpretation from central reader and from site, respectively.

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 Initial and Randomized-Withdrawal Period.

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

REDACTED COPY
This document can only be used to support any marketing authorization application and any extensions or variations thereof.

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

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 subjects and percentage with (i) suicidal ideation, (iii) suicidal behavior, (iii) suicidal ideation or behavior, and (iv) self-injurious behavior without suicidal intent.

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

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

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

[redacted]

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.

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.

10.4.4.5 Patient Health Questionnaire (PHQ)-9 scores

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

REDACTED COPY
This document cannot be used to support any marketing authorization application or its extensions or variations thereof.

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.

Change from Baseline in PHQ-9 is derived as post-Baseline score minus Baseline score.

Different to other safety variables, PHQ-9 will be summarized using the MCMC/monotone regression approach described for continuous variable.

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

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

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

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.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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

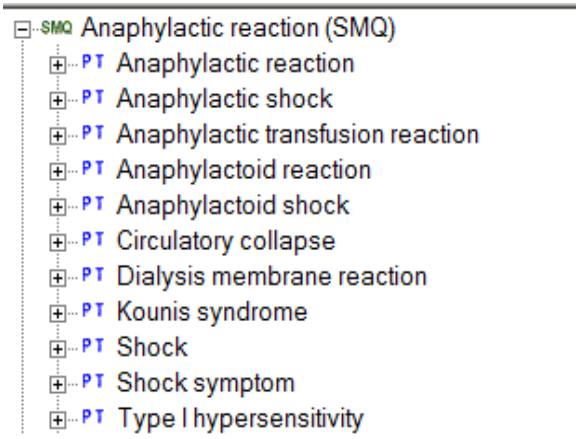
U.S. Department of Health and Human Services, Common Terminology Criteria for Adverse Events (CTCAE), Version 5.0, June 2017.

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



- 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

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

– 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 either **(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.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

12.2 Appendix B: 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

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

Note that subjects who meet the decreased potassium criterion of 3.0< LLN, which is specified as the decreased potassium lab criterion for both CTCAE Grade 1 and Grade 2, will be counted as Grade 2.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

13 AMENDMENT(S) TO THE STATISTICAL ANALYSIS PLAN

13.1 AMENDMENT 1

Rationale for the amendment

The main purpose of this amendment is to align the SAP with protocol amendment 2 and to achieve consistency with other SAP of the program.

Global Changes

Typos and formatting were updated throughout the document.

Specific changes

Change #1
Title page

- Draft SAP 30Aug2018

Has been changed to:

- Final SAP 30Aug2018
- Amendment 1 16 Sep 2019

Change #2
List of abbreviation

The following abbreviations have been added:

- LFT = liver function test
- SIB = suicidal ideation and behavior

The following abbreviations have been deleted:

- ADR = adverse drug reaction
- DAP = data analysis plan
- HLGT= Higher level group term
- LTB = latent tuberculosis
- SOP = standard operation procedure

Change#3
Section 1 Introduction

- This SAP is based on the following study document: Protocol Amendment 2, 06 Apr 2018.

Has been changed to:

- This SAP is based on the following study document: Protocol Amendment 3, 21 May 2018.

Change #4
Section 2.2.1.2 Secondary efficacy variables

- Change from Baseline in the Patient Symptom Diary responses for itch, pain, and scaling at Week 16

Has been changed to:

- Patient Symptom Diary responses for itch, pain, and scaling at Week 16

Change #5

Section 2.2.1.2 Secondary efficacy variables

The following variable has been added:

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- IGA response (defined as Clear with at least a 2-category improvement relative to Baseline) at Week 16

Change #6
Section 2.2.1.3 Other efficacy variables

- Percentage of subjects who rebound (defined **as a >125% increase** from Baseline in PASI score occurring within 2 months of stopping therapy) during the Randomized-Withdrawal Period

Has been changed to:

- Percentage of subjects who rebound (defined **as a ≥25% increase from** Baseline in PASI score occurring within 2 months of stopping therapy) during the Randomized-Withdrawal Period

Change #7
Section 2.2.1.3 Other efficacy variables

The following has been moved from Other efficacy variable to Other safety variable:

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

Change #8
Section 2.2.1.3 Other efficacy variables

- Change from Baseline in the Patient Symptom Diary

Has been changed to:

- Patient Symptom Diary responses

Change #9
Section 2.2.1.3 Other efficacy variables

- IGA response (Clear)

Has been changed to:

- IGA response (Clear with at least 2 category improvement from Baseline)

Change #10
Section 2.2.3.2 Other safety variables

- Change from Baseline in clinical laboratory values (chemistry, hematology, and urinalysis)

Has been changed to:

- Change from Baseline in clinical laboratory values (chemistry and hematology)

Change #11
Section 3.1 General presentation of summaries and analyses

The following has been added:

- Per protocol, visit windows of ±3 days from the first dose to Week 24 and ±7 days from Week 28 to Week 52 are permissible. For the Week 56 visit, the visit window is±3 days. For the SFU Visit, visit window is ±7 days from final dose. All by-visit summaries will contain

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

nominal (i.e. scheduled) visits only. Unscheduled visits will not be mapped to scheduled visits except for assessments that occur within 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.

Change #12
Section 3.2 Definition of Baseline Values

The following clarification has been added:

- 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.

Change #13
Section 3.2 Definition of Baseline Values

- For subjects who are re-randomized at the Week 16 Visit, change from Baseline variables during the Randomized-Withdrawal Period may be evaluated relative to both the Baseline (first dose) Visit and the Week 16 Visit. For the latter, the Baseline is defined as the measurement at Week 16.

Has been edited to

- For subjects who are re-randomized at the Week 16 Visit, change from Baseline variables during the Randomized-Withdrawal Period may be evaluated relative to both the Baseline (first dose) Visit and the Week 16 Visit. For the latter, the Baseline is defined as **the latest measurement on or prior to the first dose in the Randomized-Withdrawal Period (intended to be Week 16**).

Change #14
Section 3.2 Definition of Baseline Values

- For the latter, the Baseline is defined as the measurement at Week 16 or the visit at which escape treatment is initiated. If the Week 16 measurement or measurement at the visit at which escape treatment is initiated is missing, the subject will be excluded from the respective change from Baseline summaries for that measurement.

Has been edited to

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- For the latter, the Baseline is defined as the **latest measurement on or prior to the first dose in the Escape Period (for those who escape due to lack of response at Week 16) or the latest measurement on or prior to the visit at which escape treatment is initiated**.

Change #15
Section 3.2 Definition of Baseline Values

- For the variables listed below, additional baselines for Week 16 and the visit at which escape treatment is initiated are defined as well.

Has been edited to

- An additional Baseline for Week 16 and the visit at which escape treatment was initiated **(as described above)** are defined for the following variables:

Change #16
Section 3.2 Definition of Baseline Values

The following sentences have been deleted:

- Unless otherwise specified the last valid measurement before study medication administration in the initial period will be used as the baseline value
- If a scheduled baseline assessment is taken on the same day as first administration of study medication and no time is collected, it will be assumed to have been taken prior to study medication.

Change #17
Section 3.2 Definition of Baseline Values

'Laboratory data' has been added in the list of variables and additional baseline value for Week 16 and the visit at which escape treatment is initiated are defined:

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 #18
A new section 'Section 3.6.1 Subjects incorrectly entered Randomized-Withdrawal or Escape period' has been added with clarification on handling of such subjects

Subjects who enter the Escape Period without meeting the escape criteria, or subjects who meet the criteria but don't enter the Escape Period, will be summarized as follows for efficacy analyses:

4. If a subject meets escape criterion at Week 16 (<PASI90) but does not enter into the Escape Period at Week 16, then their data during the Randomized-Withdrawal Period will not be summarized in any table as they did not fulfill the definition for inclusion in the WK16ResS. Selected data from the Randomized-Withdrawal Period will be listed only.
5. If a subject meets escape criterion during the Randomized-Withdrawal Period (<PASI75) but does not enter into the Escape Period at the corresponding visit, then their data following the

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

visit at which they meet the escape criterion will be treated as missing in summaries based on the WK16ResS and will be subject to imputation as appropriate.

6. If a subject enters the Escape Period without meeting escape criterion at the corresponding visit, they will not be in the ESS. These subjects also do not qualify for inclusion in the WK16ResS. Selected data corresponding to the Randomized Withdrawal Period and Escape Period for these subjects will be listed only.

For subjects incorrectly entered into the Randomized-Withdrawal or Escape Period, a listing for some efficacy variables (e.g. PASI and IGA) and listing for all TEAE will be provided.

Change #19
Section 3.7 Center pooling strategy

Taiwan was removed from table 3-1.

Change #20
Section 3.8 Coding dictionaries

- SOP

Has been changed to:

- Standard operation procedure

Change #21
Section 3.9 Relative Day

The following has been removed:

- 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 '+'

Change #22
Section 3.10 Changes to Protocol-defined analysis

The following text has been added:

The following variables were added as a secondary efficacy variable and added to the sequential testing procedure:

- IGA response (defined as Clear with at least a 2-category improvement relative to Baseline) at Week 16

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

- Palmoplantar Investigator's Global Assessment (pp-IGA) response (clear or almost clear with at least a two-category improvement from Baseline) for subjects with palmoplantar PSO at Baseline
- Change from Baseline in Patient Global Assessment (PGA) of PSO score

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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

- Percentage of subjects who rebound (defined as a ≥25% increase from Baseline in PASI score occurring within 2 months of stopping therapy) during the Randomized-Withdrawal Period
- 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%
- Change from Baseline in clinical laboratory values (chemistry and **hematology) will be summarized, urinalysis results will be listed**

The following additional changes were made:

- The calculation of nominal p-values has been added for selected of the other efficacy variables.
- Modifications have been made to the text regarding the multiple imputation algorithm for clarity.
- Subgroup analyses will be performed for PASI75/90/100 and IGA variables only and not all co-primary and secondary efficacy variables for the Initial Treatment Period that are part of the fixed sequence testing procedure.
- Prior primary failure to biologic (yes/no) was removed as a subgroup.
- Selected TEAE tables were added to include calculation of risk differences.
- Selected summaries of safety during the Escape Period will also be presented and will be based on the ESS.

Change #23
Section 4.2.1 Handling of missing data for efficacy variables

The following text was added:

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

Change #24
Section 4.2.1.3 Handling missing data for the other efficacy variables

- For other continuous efficacy variables, the MI-MCMC/monotone regression method will be used to impute missing data as the primary method.

Has been edited to

- For other continuous efficacy variables, the MI-MCMC/monotone regression method will be used to impute missing data as the primary method, except for analysis based on the ES

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

where the OC method will be used due to the short duration of the Escape Period and the impact of missing data in this period.

Change #25
Section 4.2.1.5 Missing Data Overview and Summary

The sensitivity OC analysis for other continuous variables was removed from Table 4-1. Superscripts have been updated to reflect the changes in footnotes.

Table 4-1 in Footnote

PSD of PSO for pain, itch, and scaling has been moved to

- [c] Includes IGA response, Scalp IGA, PASI75, PASI90, PASI100 and PSD of PSO for pain, itch and scaling.
- [f] For PASE, OC is the primary analysis method.

Has been modified to:

- [f] For PASE and analysis of the Escape Period, OC is the primary analysis method.

The following text has been added:

[d] Includes Patient Global Assessment of PSO, IGA score and EQ-5D-3L **responses.**

Change #26
Section 4.2.1.6 Missing Data algorithms

- Create a data set, one for each treatment of subjects with observed values and those needing estimation by multiple imputation. The intermittent missing PASI/IGA values in each data set (i.e., 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 analyses will use this same seed as well). For monotone missing data (i.e., 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.

Has been reworded to:

- 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 (i.e., 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 (i.e., where

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

all subject data is missing after a given time point), monotone regression will then be used to impute missing data. This will be based on the100 sets of imputations already created using the MCMC method such that there will be 100 imputation 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.

Change #27
Section 4.2.1.6 Missing Data algorithms

- Three more indicator variables will be defined similarly replacing North America with Central/Eastern Europe, and Western Europe, and Asia/Australia.

Has been reworded to:

- 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.

Change #28
Section 4.2.1.6 Missing Data algorithms

- 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.

Has been modified to:

- 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.

Change #29
Section 4.2.1.6 Missing Data algorithms

The following has been added:

Additional ranges for values for **secondary and** other variables are defined in Table 4-2.

The following table has been added:

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Table 13–5: 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
PHQ-9	0	27	Yes

Change #30
Section 4.2.1.6 Missing data algorithms

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.

Has been changed to

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.

Change #31
Section 4.2.1.6 Missing Data algorithms

The following text has been added for CMH test:

- Appropriate transformations to the standard errors should also be made in order to get the correct confidence intervals.

Has been updated to

- 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).

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Change #32
Section 4.2.1.6 Missing Data algorithms

The followings have been added:

7. 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.
8. The change from Baseline values will be computed based on the complete data sets.
9. The analysis model will be based on ANCOVA (see above) as opposed to the CMH test.

Change #33
Section 4.2.1.6 Missing Data algorithms

The following sentence has been deleted:

1) the imputation model will use the change from Baseline (instead of actual) values by visit and no dichotomization will be necessary; and 2) the analysis model will be based on an analysis of covariance (ANCOVA) model.

Change #34
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 and no dichotomization will be necessary;

Has been updated to

The MI procedure will also be similar to that described above with the following differences: 1) no dichotomization will be necessary;

Change #35
Section 4.2.1.6 Missing Data algorithms

The following has been added:

Further details about imputation for different analysis sets are 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 2 randomized treatment groups (placebo, Bimekizumab 320mg Q4W) including all scheduled assessment visits from Baseline to Week 16.
- WK16ResS: When programming multiple imputation based on the WK16ResS, PROC MI will be used with a separate data set for each of the 3 re-randomized treatment groups (Bimekizumab 320mg Q4W/Placebo, Bimekizumab 320mg Q4W/Q8W, Bimekizumab 320mg Q4W/Q4W) including all scheduled assessment visits from Week 16 to Week 56. Note: The Placebo/Placebo responders will not be included in these tables or in the MI procedure.

Change #36
Section 4.3 Interim analyses and data monitoring

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

After the final Week 56 visit, an interim analysis will be performed and a corresponding interim clinical study report (CSR) may be written. Note that some subjects may still be ongoing in the Escape Period when the interim data cut if taken, after the final subject completes the Week 56 visit. 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 relevant safety data from this study and advise UCB. Details will be provided in the DMC Charter and in the Adjudication Committee Charters.

Futher details related to the DMC will be outlines in a separate analysis plan.

Has been changed to:

After the final Week 56 visit, an interim analysis will be performed on all available data at that time point (including all subjects who are in the Escape Period when the interim data cut is taken) and a corresponding interim clinical study report (CSR) will 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. Further details related to the DMC will be outlined in a separate analysis plan.

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

Change #37
Section 4.4 Multicenter studies

However, if the model is unable to converge due to a low number of subjects at a given center, a pooling by region will be applied in order to allow the model to converge.

Has been changed to

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 is still not achieved a pooling by region will be applied.

Change #38
Section 4.5 Multiple comparison/multiplicity

H_{10} Has been added in the following sentence:

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

Figure 4-1 was updated to add H_{10}.

Change # 39
Section 4.8 Examination of subgroups

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

The following characteristic has been removed:

- Prior primary failure to biologic (yes, no)

Change #40
Section 4.8 Examination of subgroups

In addition, a subgroup analysis will be performed on PASI90/100 and IGA to assess predictability of future response at Week 4 using the following subgroups:

- PASI75 responders at Week 4 (yes, no) to predict PASI90/100 and IGA through Week 56
- PASI90 responders at Week 16 (yes, no) to predict PASI100 through Week 56

For responder analysis, the definition of subgroups of the PASI75/90/100 should be based on observed values. All summaries will be based on imputed data as appropriate and will include descriptive statistics only.

Has been updated to

In addition, a subgroup analysis will be performed on PASI90/100 and IGA to assess predictability of future response at Week 4 and Week 16 using the following subgroups:

- PASI75 responders (OC) at Week 4 (yes, no) to predict PASI90/100 and IGA (NRI) through Week 56
- PASI90 responders (OC) at Week 16 (yes, no) to predict PASI100 (NRI) through Week 56

For responder analysis, the definition of subgroups of the PASI75/90/100 will be based on observed values. All summaries will be based on imputed data as appropriate and will include descriptive statistics only.

Change #41
Section 4.8 Examination of subgroups

The following text has been added:

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

Change #42

Section 5.1 Subject disposition

ESS was added for an additional summary of disposition of analysis sets:

Summaries of reasons for screen failures (for all subjects screened), disposition of subjects (for all subjects screened), disposition of analysis sets (for RS **and ESS**), disposition and discontinuation reasons in the Initial Treatment period (for RS) and in the Randomized-Withdrawal period (for WK16ResS) and Escape Treatment period (for ESS), as well as the

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

subjects who discontinued due to AEs in the Initial Treatment period (for RS) and in the Randomized-Withdrawal period (for WK16ResS) and Escape Treatment period (for ESS) will be produced.

Change # 43
Section 6.2 Other Baseline characteristics

The following characteristics have been added:

- Nail involvement (yes, no)
- Scalp involvement (yes, no)
- Palmoplantar involvement (yes, no)
- Prior primary failure to biologic (yes, no)
- Baseline nail, scalp, and palmoplantar involvement are based on the number of subjects achieving mNAPSI>0, Scalp IGA>0, and pp-IGA>0, respectively.
- 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.

Change # 44
Section 7 Measurement of treatment compliance

The following sentence has been deleted:

- All summaries detailed in this section will be performed on the SS

Change # 45
Section 7 Measurement of treatment compliance

- A summary of percent treatment compliance categorized as ≤80% and >80% will be provided by treatment group and study periods (Initial Treatment Period, Randomized-Withdrawal Period and the Initial and Randomized-Withdrawal Period).

Has been changed to:

- A summary of percent treatment compliance categorized as <75% and ≥75% will be provided by treatment group and study periods (Initial Treatment Period for the RS, Randomized-Withdrawal Period for the WK16ResS and ESS, and the Initial and Randomized-Withdrawal Period for the AMS).

Change # 46
Section 8 Efficacy analysis

All efficacy analyses will be performed on the RS unless otherwise specified. All efficacy summary tables will be displayed by treatment group unless otherwise specified.

Has been updated to

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

All efficacy analyses of primary and secondary variables will be performed on the RS unless otherwise specified. All efficacy analyses of other efficacy variables will be performed on the RS, Wk16ReS and ESS unless otherwise specified. All efficacy summary tables will be displayed by treatment group unless otherwise specified.

Change # 47
Section 8.1.2 Primary analysis of the co-primary efficacy variables; section 8.1.3.1 Sensitivity analysis #1 and section 8.1.3.2 Sensitivity analysis #2

The following sentence has been added:

- If one of the treatment groups has zero or very low response where the CMH odds ratio can no longer be calculated, the logit method will be applied instead.

Change #38
Section 8.1.2 Primary analysis of the co-primary efficacy variables

The following text was deleted:

The number and percentage of subjects who are PASI90 responders and ~~those who are non-responders~~ at Week 16 will be summarized.

Change # 48

Section 8.1.3.7 Sensitivity Analysis #7
The following has been added:

- In order to achieve model convergence, prior biologic exposure may be dropped from the model. If convergence still cannot be achieved, this analysis will not be performed.

Change # 49
A new section 8.2.1.2 IGA response at Week 16 has been added with the following sentence:

- A categorical response variable, IGA response (Clear with at least 2-category improvement from Baseline) is defined as an IGA score of zero (Table 8-1) with at least 2-category improvement from Baseline.

Change #50
Section 8.2.1.3 Patient Symptom Diary (PSD)

- A weekly average is the sum of the scored item over the course of the study week 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

Has been changed to:

- A weekly average is the sum of the scored item over the course of the study week divided by the number of days on which the item was completed. The weekly averages should be relative to the respective visit date except for baseline which should follow the usual convention and should be based on at least 4 non-missing values (do not need to be consecutive). Otherwise, the PSD will be set to missing. The Baseline value will be computed in the same manner.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Change #51

Section 8.2.1.3 Patient Symptom Diary (PSD)

The followings sentences have been added:

- The PSD will be computed based on the responder. Each of the 3 PSD response scores included in the statistical testing procedure (itch, pain, and scaling) will be characterized in terms of the cumulative percent of subjects demonstrating a pre-specified point improvement at Week 16. The threshold for the PSD response score of itch, pain, and scaling is 2.39, 1.98, and 2.86, respectively. The responder analysis will be limited to the subjects with a Baseline PSD response score at or above the applicable threshold score (i.e., at least 2.39, 1.98, and 2.86 for itch, pain, and scaling, respectively).
- The derivations for absolute and percent change from Baseline are as follows:

Change #52
Section 8.2.1.3 Patient Symptom Diary (PSD)

The following paragraph has been deleted:

- In addition, each of three PSD response scores -itch, pain and scaling, will be characterized in terms of the cumulative percent of subjects demonstrating a pre-specified point improvement at Week 16. The threshold for the PSD response score of itch, pain, and scaling is 2.39, 1.98, and 2.86, respectively. For this responder analysis, it will be limited to the subjects with a Baseline PSD response score at or above the applicable threshold score (i.e., at least 2.39, 1.98, and 2.86 for itch, pain and scaling respectively).

Change #53

Section 8.2.1.3 Patient Symptom Diary (PSD)

The following bullet was deleted:

- Percent change from Baseline in PSD responses for itch, pain, and scaling at Week 16 is defined as

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

Change #53
Section 8.2.1.3 Patient Symptom Diary (PSD)

The following paragraph has been deleted:

- Weekly average score of PSD response for itch, pain and scaling will be included in the analysis for secondary efficacy variables. Baseline in PSD response is defined as the average score of the last 7 days before study medication administration at Baseline visit as long as at least 1 day is available. In the event there are no values for days prior to Baseline, the diary score from the day of Baseline will be used as Baseline

Change #54

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extension or variations thereof.

Section 8.2.2 Primary analysis of the secondary efficacy variables

The following text was deleted:
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.

Change #55
Section 8.2.2.1 PASI responses

The following text was added:
(based on subjects who have switched to placebo from bimekizumab at Week 16, to evaluate the efficacy of continuous treatment with bimekizumab versus treatment withdrawal (placebo))

Change #56
A new section 8.2.2.2 IGA response has been added.

- For IGA response (clear) at Week 16, the stratified CMH test similar to the primary analysis will be applied. A line plot of IGA responder rate (clear) over time, by treatment group will be produced.

Change #57
Section 8.2.2.3 PSD

The following has been added:

- The PSD response variables will be characterized in terms of the cumulative percent of subjects demonstrating a pre-specified point improvement at Week 16. The threshold for the PSD response score of itch, pain, and scaling is 2.39, 1.98, and 2.86, respectively. This responder analysis will be limited to the subjects with a Baseline PSD response score at or above the applicable threshold score (i.e., at least 2.39, 1.98, and 2.86 for itch, pain and scaling respectively). The stratified CMH test, similar to the primary efficacy analysis, will be applied to each of these responder analyses. These are the inferential analyses that will be used in the multiplicity-controlled testing procedure for itch, pain, and scaling.

Change #58
Section 8.2.2.3 PSD

The following has been deleted:

- ~~The percentage change as well as the absolute change from Baseline to the Week 16 for the weekly average will be analyzed in separate analyses of covariance (ANCOVA) with fixed effects of treatment, region and prior biologic exposure, and Baseline value as a covariate. The analysis based on the absolute change from Baseline will be the primary analysis that will be considered in the fixed sequence testing procedure. Difference between treatment groups will be determined using least square means and t-tests using the pooled error term from the linear model. 95% confidence intervals for the treatment difference will be calculated based on differences in least squares means estimates. All continuous variables in the model will be centered with respect to their mean values.~~

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- 1) The imputation model will use change from Baseline (instead of actual) values by visit and no dichotomization will be necessary; and
- 2) The analyses model will be based on analysis of covariance (ANCOVA) as opposed to the Cochran-Mantel-Haenszel (CMH) test.

Change #59
Section 8.2.2.3 PSD

The followings have been deleted:

- Absolute and percent changes from Baseline PSD for itch, pain and scaling will be summarized by visit for each treatment group. Also, the number and percentage of subjects who have worsened or no improvement, at least 1 point, at least 2 points, at least 3 points, at least 4 points, and at least 5 points improvement at Week 16 visit will be summarized for each of the itch, pain and scaling scores.
- In addition, each of three PSD response scores - itch, pain, and scaling, will be characterized in terms of the cumulative percent of subjects demonstrating a pre-specified point improvement at week16. The threshold for the PSD response score of itch, pain, and scaling is 2.39, 1.98, and 2.86, respectively. For this responder analysis, it will be limited to the subjects with a Baseline PSD response score at or above the applicable threshold score (i.e., at least 2.39, 1.98, and 2.86 for itch, pain and scaling respectively). Stratified CMH test, similar to the primary efficacy analysis, will be applied within each of these responder analyses.

Change #60
Section 8.2.3 Sensitivity analyses of the secondary efficacy variables

The following text was updated:

For binary response variables (PASI100 at Week 16, PASI75 at Week 4, ~~and~~ Scalp IGA at Week 16, **and PSD response for itch, pain and scaling**), sensitivity analyses #1 and #3 (Section 8.1.3.1, Section 8.1.3.3) will be performed.

Change #61
Section 8.2.3 Sensitivity analyses of the secondary efficacy variables

The following text was added:

Missing data for these PSD variables will use the MI procedure similar to the sensitivity analysis #1 (Section 8.1.3.1) for the co-primary efficacy variables with the following differences:

10. The data included in the imputation model will be limited only to those subjects with sufficiently high Baseline values to achieve the response for a given item (based on observed cases).
11. The PSD item value will be imputed. Once imputation has been performed across the 100 iterations specified, any values below 0 and above 10 will be truncated to 0 and 10, respectively.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

12. The change from Baseline values will be computed based on the complete data sets.
13. The analysis model will be based on the CMH test.There may be cases where the multiple imputation model fails to converge. In such situations, the LOCF approach will instead be used to impute the missing data. If LOCF is used instead of multiple imputation for this reason, this will be clearly specified in the corresponding table summary. Note that for LOCF imputation, any missing data or data collected following discontinuation of IMP will be imputed with the most recent previous value.

Change #62

Section 8.2.3 Sensitivity analyses of the secondary efficacy variables

The following text was added:

Missing data for continuous secondary efficacy variables will be imputed using MI. The MI procedure for continuous variables will be similar to sensitivity analysis #1 described in Section 8.1.3.1 for the co-primary efficacy endpoints with the following difference – no dichotomization will be necessary. Also, no requirement of meeting minimum threshold at baseline is needed to be included in imputation for a continuous variable.

Change #63

Section 8.3 Statistical analysis of the secondary efficacy variables

- Note that in addition, for subjects who change treatment at Week 16, change from two Baselines (i.e Baseline at Week 0 visit and Baseline at Week 16) will be summarized in the WK16ResS.

  For subjects that enter the escape arm at Week 16 or later, change from two Baselines (i.e. Baseline at Week 0 visit and Baseline at which escape treatment was initiated) will be summarized in the ESS.

Have been edited to:

- Note that in addition, for subjects who change treatment at Week 16, change from two Baselines (i.e Initial Period Baseline and Randomized Withdrawal Period Baseline) will be summarized in the WK16ResS for PASI and DLQI.

  For subjects that enter the escape arm at Week 16 or later, change from two Baselines (i.e. Initial Period Baseline and Baseline at which escape treatment was initiated) will be summarized in the ESS for PASI and DLQI.

Change #64

Section 8.3 Statistical analysis of the secondary efficacy variables

The following has been deleted:

- Missing data for continuous secondary efficacy variables will be imputed using MI. The MI procedure for continuous variables will be similar to sensitivity analysis #1 described in Section 8.1.3.1 for the co-primary efficacy endpoints with the following differences: 1) the imputation model will use the change from Baseline (instead of actual) values by visit and no

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

dichotomization will be necessary; and 2) the analysis model will be based on ANCOVA as opposed to the stratified CMH test

Change #65

Section 8.3.1.2 Time to PASI50, PASI75, PASI90, and PASI100 response

The following has been added:

- All visits including unscheduled visits are considered.

Change #53
Section 8.3.1.2 Time to PASI50, PASI75, PASI90, and PASI100 response

The following text has been added:

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 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 the** Week 16 Visit. **Subjects will be censored at Baseline if there is no Baseline PASI assessment or no Post Baseline PASI assessment.**

Change #66

Section 8.3.1.3 Time to relapse and section 8.3.1.4 Percentage of subjects who relapse

The following has been added:

- (excluding subjects who were treated with placebo in both Initial Treatment Period and Randomized-Withdrawal Period as they cannot relapse by definition).

Change #67
Section 8.3.1.5 Percentage of subjects who rebound

Rebound is defined as when a subject experience a >=125% increase from Baseline in PASI score occurring within 2 months (60 days) of stopping therapy

Has been changed to

Rebound is defined as when a subject experiences a >=25% increase from Initial Treatment Period Baseline in PASI score occurring within 2 months (60 days) of stopping therapy

Change #68
Section 8.3.1.5 Percentage of subjects who rebound

The percentage of subjects who rebound will be presented by treatment group over time.

Has been updated to

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

The percentage of subjects who rebound will be presented for subjects re-randomized to placebo over time.

Change #69
Section 8.3.1.6 PASI score

The following text was added:

using NRI

Change # 70
Section 8.3.1.7 Modified Nail Psoriasis Severity Index (mNAPSI) score

The following clarification has been added.

- If any of the 7 response items that contribute to mNAPSI is present, while other items are missing (i.e., 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 #71
Section 8.3.1.7 Modified Nail Psoriasis Severity Index (mNAPSI) score

The following sentences have been deleted:

- 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.

Change #72
Section 8.3.2 IGA response

- IGA response (Clear) is defined as an IGA score of zero (Table 8-1). Shift from Baseline in IGA score is defined at each Post-Baseline visit timepoint relative to Baseline.

Has been updated to

- IGA response (Clear with at least 2-category improvement from Baseline) is defined as an IGA score of zero (Table 8-1) with at least 2-category improvement from Baseline. Shift from Baseline in IGA score is defined at each Post-Baseline visit timepoint relative to Baseline.

Change #73
Section 8.3.2 IGA response

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

Has been updated to

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

A line plot of the IGA (both clear or almost clear and clear) and scalp IGA responder rate over time, by treatment group will be produced.

Change #74
Patient Symptom Diary (PSD)

The following text was deleted:

Definitions and derivations of absolute change and percent change from Baseline in the Patient Symptom Diary responses, responder at Week 16 based on each PSD item have been outlined in Section 8.2.1.3.

Change #75
Patient Symptom Diary (PSD)

For PSD response score of redness, burning, cracking, dryness, irritation, sensitivity, lesions, thickening, fatigue, embarrassment, choice of clothing, the responder criterions in terms of change from Baseline are set to 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 PSD response scores and PSD responder data for each item, the analyses specified in Section 8.2.2 will be provided

In addition, cumulative distribution plots will be provided for absolute and percent change from Baseline PSD at Week 16 for each item.

Has been updated to

For PSD response score of redness, burning, cracking, dryness, irritation, sensitivity, lesions, thickening, fatigue, embarrassment, choice of clothing, the responder criterions in terms of change from Baseline are set to 3.05, 1.99, 2.01, 3.04, 2.82, 2.60, 2.69, 2.68, 1.51, 2.43, and 2.14 respectively. Number and percentage of subjects who are responders will be summarized based on each of these items.

In addition, cumulative distribution plots will be provided for absolute and percent change from Baseline PSD at Week 16 for each item.

Change #76
Section 8.3.10 Short Form 36-item Health Survey (SF-36)

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

Has been updated to

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

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Change #77
Section 8.3.11 Euro-Quality of Life 5-Dimensions, 3 levels (EQ-5D-3L)

The following text was added:

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 #78
Section 8.3.12 Patient Health Questionnaire (PHQ)-9 Scores

This section was deleted.

Change #79
New Section 8.4 Additional statistical analysis of other efficacy variables

Change #80
Section 9.1 Pharmacokinetics

The following sentence has been deleted:

- Additional PK analysis will be described in the separate data analysis plan (DAP).

Change #81

Section 9.1 Pharmacokinetics

The following sentence has been added:

- In addition, if the PK sampling date is >1 day after the dosing date, then the plasma concentration from that visit will be excluded from the PK summary.

Change # 82
Section 9.2 Immunogenicity

- The analysis of immunogenicity will be performed on PK-PPS.

Has been changed to:

- The analysis of immunogenicity will be performed on SS.

Change #83
Section 9.2.2 Anti-bimekizumab antibodies

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 (20 weeks after the last dose).

Has been updated to

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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 open label extension study (PS0014) a sample at SFU (20 weeks after the last dose). For the interim analysis, the SFU visit (after Week 56) will be excluded in the analyses.

Change #84
Section 9.2.2. Anti-bimekizumab antibodies

The following text was updated:
Subject Classification:

- For subjects who are negative at baseline, and antibody negative at all sampling points post **initial** 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 **initial** 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-**initial** 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 **initial** 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 **initial** treatment (including SFU) with titer values of the same magnitude as baseline (i.e. ≤ then a ~~predefined~~ **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 **initial** treatment (including SFU) with increased titer values compared to baseline (above a ~~predefined~~ **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- **initial** treatment samples are missing, while other post-**initial** treatment samples are ADAb negative, the subject will be classed as inconclusive.

Derivation for above classification will be different for the interim analysis and the final analysis. For the interim analysis no SFU data will be considered. Only for the final analysis, when all SFU data will be available, data from the SFU visits will be considered.

Change # 85
Section 9.2.2 Anti-bimekizumab antibodies

The following text has been 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:

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- 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.
- A summary table to present the incidence (%) of subjects classified as treatment emergent ADAb positive with incidence of IL-17A and IL-17F neutralizing ADAb, respectively over the 56 weeks.

Change # 86
Section 9.2.2 Anti-bimekizumab antibodies

- 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 positive - treatment emergent reduced 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 in categories 2 and 5, summarized as total treatment emergent. In addition, the counts and percentage of subjects who are pre-anti-bimekizumab positive will be calculated (this is the sum of categories 3,4, and 5).

Change # 87
Section 9.2.2 Anti-bimekizumab antibodies

- 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

Has been changed to:

- Individual plots of Bimekizumab Concentrations/ADAb titer and % Change from baseline (all Y-axes) versus time (x-axis) for the full treatment period (excluding SFU for interim analyses and including SFU for final analyses), where a subject has not progressed into the open label extension study (PS0014). Plots should be labeled and grouped into the 6 sub-categories

Change # 88
Section 9.2.2 Anti-bimekizumab antibodies

The following has been deleted:

A summary table to present the incidence (%) of subjects classified as treatment emergent ADAb positive with incidence of IL-17A and IL-17F neutralizing ADAb, respectively over the 56 weeks.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Change #89

Section 10 Safety Analysis

The following sentence has been added:

- Selected summaries of safety during the Escape Period will also be presented and will be based on the ESS

Change #90

Section 10.1 Extent of exposure

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

Has been changed to:

- >0 weeks; ≥ 4 weeks; ≥8 weeks; ≥ 12 weeks; ≥ 16 weeks for Initial Treatment Period
- >0 weeks; ≥ 4 weeks; ≥8 weeks; ≥ 12 weeks; ≥ 16 weeks; ≥ 20 weeks; ≥ 40 weeks for Randomized-Withdrawal Period
- >0 weeks; ≥ 4 weeks; ≥8 weeks; ≥ 12 weeks for Escape Period

Change #91

Section 10.1.1 Exposure during the Initial Treatment Period for SS

The following text has been added:

- The total number of days in the Initial Treatment Period (112 days**). For AEs that emerged after 112 days but still within the 140 days window, those AEs would be classified as TEAE, but will be excluded from the output based on the Initial Treatment Period. However these AEs will be included in the AE summaries for Randomized-Withdrawal Period or Escape Period for subjects in the WK16ResS and ESS respectively.**

Change #92

Section 10.1.2 Exposure during the Randomized Withdrawal Period for WK16ResS

The following text has been added:

Note: If date of last dose in the Randomized-Withdrawal Period + 28 days (or 56 days in the case of Q8W dosing) extends to a date beyond the date of the first dose in the 12-Week Escape Period, then this calculation reverts to:

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Date of first dose in the 12-Week Escape Period – date of first dose in the Randomized-Withdrawal Period + 1.

Change #93
Section 10.1.2 Exposure during the Randomized Withdrawal Period for WK16ResS

The following has been added:

- For all other subjects, use the minimum of the following:
    - Date of last dose in the Randomized-Withdrawal Period – date of first dose in the Randomized-Withdrawal Period + 140 days.
    - **Date of first dose in the 12-Week Escape Period – date of first dose in the Randomized-Withdrawal Period + 1.**
    - Date of last clinical contact – date of first dose in the Randomized-Withdrawal Period + 1.

Change #94

A new section, Section 10.1.3 Exposure during the Escape Period for ESS, has been added.

Change #95
Section 10.1.4 Exposure during the Initial and Randomized-Withdrawal for AMS

The following text was updated:

Definitions for study medication duration (days) and time at risk (days) entire study treatment period (the Initial and the Randomized-Withdrawal Period~~, and the Escape Period~~) are provided as follows:

Change #96
Section 10.1.14 Exposure during the Initial and Randomized-Withdrawal for AMS

The following text was updated:

*For subjects who do not switch study **treatments and do not enter the 12-Week Escape Period**:*

Change #97
Section 10.1.4 Exposure during the Initial and Randomized-Withdrawal and 12-Week Escape Period for AMS

The following text has been deleted:

- 12-Week Escape Period (attributed to BKZ 320mg Q4W):
    - Date of last dose in the 12-Week Escape Period – date of first dose in the 12-Week Escape Period + 28 days

REDACTED COPY
This document can't be used to support any marketing authorization application and any extensions or variations thereof.

Note: If date of last dose in the 12-Week Escape Period + 28 days extends to a date beyond the final visit date of the 12-Week Escape Period (not including SFU), then this calculation reverts to:

- Final visit date of the 12-Week Escape Period (not including SFU) – date of first dose in the 12-Week Escape Period + 1.

Note: For subjects who die during the 12-Week Escape Period, then this calculation reverts to:

- Date of death – date of first dose in the 12-Week Escape Period + 1.

The followings have been deleted:

- 12-Week Escape Period (attributed to BKZ 320mg Q4W):
  - For subjects who complete the 12-Week Escape Period as planned and continue into an open-label study (and, therefore, do not have the SFU visit in the feeder study): Final visit date of the 12-Week Escape Period – date of first dose in the 12-Week Escape Period + 1.
  - For subjects who die prior to the final visit of the 12-Week Escape Period: Date of death – date of first dose in the 12-Week Escape Period + 1.
  - For all other subjects, use the minimum of the following:
    - Date of last dose in the 12-Week Escape Period – date of first dose in the 12-Week Escape Period + 140 days.
    - Date of last clinical contact – date of first dose in the 12-Week Escape Period + 1.

Change #98
Section 10.2 Adverse events

- Events that fulfill the anaphylaxis criteria for acute events (refer to Appendix 1in "Bimekizumab Safety Topics of Interest.docx")

Has been updated to:

- Events that fulfill the anaphylaxis criteria for acute events (refer to Appendix A in Section 12.1)

Change #99
Section 10.2.1 Data considerations

The following description has been added:

- Selected summaries, as specified in Section 10.2.2, will include the risk difference between bimekizumab and placebo. The risk difference is calculated as: $RD = IP_{BKZ} - IP_{PBO}$ where IP_{BKZ} is the incidence proportion for the bimekizumab-treated group and IP_{PBO} is the incidence proportion for the placebo group. Note that incidence proportion simply refers to the percentage of subjects within the specified treatment group that experienced a given

REDACTED COPY
This document cannot be used to support any marketing authorisation application and any extensions or variations thereof.

adverse event. The standard error for the risk difference is calculated as follows: $SE_{RD} = \sqrt{\left(IP_{BKZ} \times \left(\frac{1-IP_{BKZ}}{n_{BKZ}}\right)\right) + \left(IP_{PBO} \times \left(\frac{1-IP_{PBO}}{n_{PBO}}\right)\right)}$

where n_{BKZ} is the number of subjects in the bimekizumab-treated group and n_{PBO} is the number of subjects in the placebo 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). The risk difference and corresponding CI will be displayed as percentages.

Change #100
Section 10.2.1 Data considerations

The following text was deleted:
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.

Change #101
Section 10.2.2 AE Summaries

The following has been added:
Additional summaries for the ESS will include only the TEAE overview, TEAEs and SAEs per 100 subject years, and TEAEs tables for safety topics of interest. For AEs that emerged after 112 days but still within the 140 days window, those AEs would be classified as TEAE. These AEs will be excluded from the outputs based on the Initial Period but included in the AE summaries for Initial and Randomized-Withdrawal Period.

Change #102
Section 10.2.2 AE Summaries

The following clarifications have been added;

- 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 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 placebo).

Change # 103
Section 10.2.2 AE Summaries

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- Incidence of Serious Adverse Drug Reactions by SOC, HLT, and PT
- Incidence of Non-Serious TEAEs by Relationship SOC, HLT, 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

Have been update to

- Incidence of Related Serious TEAE by SOC, HLT, and PT
- Incidence of Non-Serious TEAEs by Maximum Relationship SOC, HLT, and PT
- Incidence of Related TEAEs by SOC, HLT, and PT
- Incidence of Related TEAEs Above Reporting Threshold of 5% by SOC and PT

Change #105
Section 10.2.3 Other Safety topics of interest

The following sentence has been added:

- 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 #106
Section 10.2.3.2 Infections (serious, opportunistic, fungal, and TB)

The following text was deleted:

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

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:

- 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.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Additionally, a column will be included where the study physician can document their decision on the case.

- 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'.
- 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 final 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.

Change #107
Section 10.2.3.2 Infections (serious, opportunistic, fungal, and TB)

The following text was added:

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

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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 #108
Section 10.2.3.2 Malignancies, including lymphoma has been renamed Malignancies.

Change #109
Section 10.2.3.3 Adjudicated Major adverse cardiac event

The following has been added:

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 #110
Section 10.2.3.5 Neuropsychiatric events (in particular depression and suicide)

The title has been changed to Suicidal Ideation and Behavior.

The following has been deleted:

- This table is based on SMQ of "Depression and suicide/self-injury" (all TEAE which code to a PT included in the Scope-Broad and/or Scope=Narrow).

The following has been added:

- . A **table and a** 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.**

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Change #11
Section 10.2.3.6 Inflammatory bowel disease

The following has been deleted:
The incidence and events rate will be presented by history of IBB at Baseline.

The following has been added:

The table will be 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 #112
Section 10.2.3.7 Hypersensitivity (including anaphylaxis)

The title was updated to Hypersensitivity (including analphylaxis).

- This table will be prepared based on the MedDRA anaphylaxis algorithm (refer to Appendix 1 A in Bimekizumab Safety-Topics of Interest.docx section 12) for acute events (reported on the same day) and for treatment-emergent PTs including the term “hypersensitivity” reported at any time.

Has been changed

- A separate table will be prepared based on the MedDRA anaphylaxis algorithm (refer to Appendix A in Section 12) for acute anaphylactic events (reported on either the same day as when an injection was administered or one day after). 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.

Change #113
Section 10.2.3.8 Hepatic events and DILI

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

Has been changed to:

- Cases of Hy’s Law will be reported separately in a liver function test table.

Change #114
10.3 Clinical laboratory evaluations

Laboratory values, including markedly abnormal laboratory values and urinary values will be presented descriptively by treatment group for the SS, WK16ResS and AMS.

Has been updated to:

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Laboratory values, including markedly abnormal laboratory values will be presented descriptively by treatment group for the SS, WK16ResS and AMS. In addition, the ESS will be summarized for only markedly abnormal laboratory values and Hy's Law outputs.

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

Change #115

10.3 Clinical laboratory evaluations

- A summary of the number and percentage of subjects experiencing markedly abnormal values by laboratory variable, treatment group and visit

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

Change #116

10.3 Clinical laboratory evaluations

The following item has been added:

- 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

Change #117

10.3 Clinical laboratory evaluations

- A shift table of the number and percentage of subjects experiencing low, normal or high values at Baseline to maximum post-Baseline value (i.e., low, normal, high), by laboratory variable and treatment group

Has been changed to:

- 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

Change #118

10.3 Clinical laboratory evaluations

The year in the following has been updated from 2010 to 2017:

U.S. Department of Health and Human Services 2017

Change #119

10.3 Clinical laboratory evaluations

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

The following items have been deleted:

- A Shift table of the number and percentage of subjects experiencing low, normal or high values at Baseline to minimum post-Baseline value (i.e., 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 (i.e., 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.

Change #120

10.3 Clinical laboratory evaluations

The following text has been added:

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.

Change #121

10.3 Clinical laboratory evaluations

- AST: >2xULN, >3xULN, >5xULN, >10xULN, >20xULN
- ALT: >2xULN, >3xULN, >5xULN, >10xULN, >20xULN
- AST or ALT: >2xULN, >3xULN, >5xULN, >10xULN, >20xULN
- Bilirubin: >1xULN, >1.5xULN
- ALP: >2xULN

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

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

In order to meet either of the above criteria, a subject must experience the elevation in bilirubin and ALT or AST (and the absence of ALP elevation, if applicable) at the same visit.

Has been updated to:

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

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- 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 two definitions 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, if applicable) at the same visit.

Change #121

10.3 Clinical laboratory evaluations

The following clarifications have been added:

- In addition, the ESS will be summarized for only markedly abnormal laboratory values and Hy's Law outputs.
- The markedly abnormal tables and those based on CTCAE grade will be produced only for selected laboratory variables.

Change #122

10.3 Clinical laboratory evaluations

The following sentences have been deleted:

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.

Change #123

10.3 Clinical laboratory evaluations

- AST: >2xULN, >5xULN, >10xULN, >20xULN
- ALT: >2xULN, >5xULN, >10xULN, >20xULN
- AST or ALT: >3xULN, >5xULN, >8xULN, >10xULN, >20xULN
- Bilirubin: >1xULN, >1.5xULN

Have been updated to:

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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

Change #124

10.3 Clinical laboratory evaluations

The following item has been deleted.

- [AST ≥3xULN or ALT ≥3xULN] and Total Bilirubin ≥2xULN
- ALP>2xULN

Change #125

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 updated to:

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

Change #126
Section 10.4.3 Electrocardiograms

A by-subject listing of all 12-lead ECG data will be provided.

Has been updated to:

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

Change #127
Section 10.4.4.1 Assessment and management of TB and TB risk factors

By-subject listing of the result of chest x-ray for tuberculosis will be provided by treatment group based on interpretation from central reader and from site.

Has been updated to:

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

Change #128

10.4.4.5 Patient Health Questionnaire (PHQ)-9 Score

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

This section has been moved from section 8.3.12

Change #130
Section 11: References
The following reference has been deleted:
Basra, M. K., Fenech, R., Gatt, R. M., Salek, M. S., & Finlay, A. Y., The Dermatology Life Quality Index 1994-2007: a comprehensive review of validation data and clinical results. Br J Dermatol., 2008; 159, 997-1035.

Change #131
Section 11: References
The following reference has been added:
Ratitch, B., Lipkovich, I., O'Kelly, M. Combining Analysis Results from Mutiply Imputed Categorical Data, PharmaSUG, 2013, SP03.

Change #132
Section 11: References
The version and year in the following have been updated from 4.0 to 5.0 and 2010 to 2017:

- U.S. Department of Health and Human Services, Common Terminology Criteria for Adverse Events (CTCAE), Version 5.0, June 2017.

Change #133
Section 12.1 Appendix A: MedDRA algorithmic approach to anaphyalaxis

Appendix A was updated according to program standards.

Change #134

Section 12.2 Appendix B: MedDRA algorithmic approach to anaphylaxis

Appendix B was added.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

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.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Approval Signatures

Name: PS0013-sap-amendment1

Version: 1.0

Document Number: CLIN-000138962

Title: PS0013-sap-amendment1

Approved Date: 16 Sep 2019

Document Approvals	
Approval Verdict: Approved	Name: [redacted] Capacity: Clinical Date of Signature: 15-Sep-2019 20:35:26 GMT+0000
Approval Verdict: Approved	Name: [redacted] Capacity: Clinical Date of Signature: 16-Sep-2019 10:58:51 GMT+0000

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.