Janssen Research & Development

Statistical Analysis Plan

A Phase II, Randomized, Double-Blind, Placebo-Controlled Study of Bermekimab in Patients with Moderate to Severe Hidradenitis Suppurativa

Protocol 77474462HDS2002; Phase 2

JNJ77474462 (Bermekimab)

Status: Approved
Date: 5 June 2020
Prepared by: Janssen Research & Development, LLC
Document No.: EDMS-ERI-206032357

Compliance: The study described in this report was performed according to the principles of Good Clinical Practice (GCP).

TABLE OF CONTENTS

ABBREVIATIONS

ADA	anti-drug antibody
AE	adverse event
ALT/SGPT	alanine aminotransferase
ANCOVA	analysis of covariance
AST/SGOT	aspartate aminotransferase
ATC	anatomic and therapeutic Class
AUC	area under the curve
BMI	body mass index
BSA	body surface area
CI	confidence interval
CL	total systemic clearance
Cmax	maximum concentration
CMH	Cochran-Mantel-Haenszel
CRF	case report form
CRP	C-reactive protein
CSR	Clinical Study Report
CV	coefficient of variation
DBL	Database Lock
DLQI	Dermatology Life Quality Index
ECG	electrocardiogram
eCRF	electronic case report form
EDC	Electronic Data Capture
ESR	Erythrocyte Sedimentation Rate
FAS	full analysis set
FCS	Full Conditional Specification
FDA	Food and Drug Administration
HADS	Hospital Anxiety and Depression Scale
HiSCR	Hidradenitis Suppurativa Clinical Response
HiSQOL	Hidradenitis Suppurativa Quality of Life Measure
HS	Hidradenitis Suppurativa
HS-PGA	Hidradenitis Suppurativa Physician's Global Assessment
HSSD	Hidradenitis Suppurativa Symptom Diary
ICH	International Conference on Harmonization
IHS4	International Hidradenitis Suppurativa Severity Score System
IQ	interquartile
ITT	intent-to-treat
IWRS	interactive web response system
LLOQ	lower limit of quantification
LOCF	last observation carried forward
LS means	Least square means
MAR	missing at random
MedDRA	Medical Dictionary for Regulatory Activities
mHSS	Modified Hidradenitis Suppurativa Score
MMRM	Mixed-Effect Model Repeated Measure
NAb	neutralizing antibodies
NCI-CTCAE	National Cancer Institute - Common Terminology Criteria for Adverse Events
NONMEM	nonlinear mixed-effects modeling
NRS	Numeric Rating Scale
PD	pharmacodynamic
PGA-ISD	Physician's Global Assessment of Inflammatory Skin Disease
PGI-c	Patient's Global Impression of Change
PGI-s	Patient Global Impression of Severity
PK	pharmacokinetic(s)
PP	per protocol

q2w	every other week
qw	every week
REML	restricted maximum likelihood
SAE	serious adverse event
SAP	Statistical Analysis Plan
SC	subcutaneous
SD	standard deviation
SUSAR	suspected, unexpected, serious adverse reaction
TEAE	treatment-emergent adverse event
WHO	World Health Organization
WHO-DD	World Health Organization Drug Dictionary

1. INTRODUCTION

This statistical analysis plan (SAP) contains definitions of analysis sets, derived variables, and statistical methods for the analysis of efficacy, safety, pharmacokinetics (PK), pharmacodynamics (PD), Immunogenicity and health-related quality of life in the JNJ-77474462PT047 study.

1.1. Trial Objectives

Primary Objective

The primary objective of this study is to evaluate initial efficacy, safety of different dose regimens of bermekimab in adult participants with moderate to severe hidradenitis suppurativa (HS).

Secondary Objectives

The secondary objectives of this study are:

- To evaluate the effect of bermekimab on the dermatologic health-related quality of life in adult participants with moderate to severe HS.
- To evaluate the pharmacokinetics (PK), immunogenicity, and pharmacodynamics (PD) of bermekimab therapy in adult participants with moderate to severe HS.

1.2. Trial Design

This is a randomized, double-blind, placebo-controlled study to further establish the efficacy and safety and explore dose effects for bermekimab monotherapy in adults with moderate-to-severe HS. It is anticipated that approximately 50% of patients enrolled will have failed prior targeted biologic therapy(ies) and approximately 50% will be naïve to these forms of biologic therapy in each arm. Patients will be assessed for study eligibility at the initial screening visit after providing informed consent. Thirty days are allowed in the screening window to complete all screening procedures and randomize the participant. During the screening period, certain treatments will be washed out, as applicable, according to eligibility requirements. Patients who fail initially may be re-screened once based on the investigator discretion but only if inclusion criteria are met at time of re-screening.

The study is approximately 40 weeks (30-day screening period, 32-week treatment period and a 4-week follow-up period) for all participants.

A diagram of the study flow (Figure 1) is provided below. Approximately 150 patients who meet eligibility criteria at screening will be randomized in a 1:1:1 fashion to one of three treatment arms:

Bermekimab 400 mg SC (qw) (Treatment Arm 1): 800 mg loading dose of bermekimab (2 x 400 mg SC injections) at weeks 0 and 1 followed by weekly (qw) subcutaneous (SC) injections of 400 mg bermekimab through week 31

Bermekimab 400 mg SC (q2w) (Treatment Arm 2): 800 mg loading dose of bermekimab (2 x 400 mg SC injections) at weeks 0 and 1 followed by subcutaneous (SC) injections of 400 mg bermekimab every other week (q2w) alternating with matching placebo q2w through week 31.

Placebo (Treatment Arm 3): 4mL loading dose of placebo (2 x 2mL SC injections) at weeks 0 and 1 followed by weekly placebo injections from week 2 through week 15. Patients in Arm 3 will begin to receive weekly (qw) blinded subcutaneous (SC) injections of 400 mg bermekimab beginning at week 16 through week 31.

Efficacy assessments (Hidradenitis Suppurativa Clinical Response (HiSCR), Hidradenitis Suppurativa Physician's Global Assessment (HS-PGA), Modified Hidradenitis Suppurativa Score (mHSS), lesion counts, International Hidradenitis Suppurativa Severity Score System (IHS4), Hurley staging, Hidradenitis Suppurativa Symptom Diary [HSSD], Numeric Rating Scale (NRS) for Pain and Itch, Dermatology Life Quality Index [DLQI], Hospital Anxiety and Depression Scale [HADS], Patient's Global Impression of Change [PGI-c], Physician's Global Assessment of Inflammatory Skin Disease (PGA-ISD), Patient Global Impression of Severity (PGI-s), Hidradenitis Suppurativa Quality of Life Measure (HiSQoL), and Health Status Questionnaire (EQ-5D-3L) will be performed according to the Schedules of Activities. Serum samples for PK, immunogenicity, and biomarker analyses will be collected at the timepoints shown in the Study Calendars (Section 5.2 of the protocol).

Blood samples will be collected for laboratory testing including chemistry, hematology, inflammation markers, bermekimab concentration (PK), PD marker(s), and anti-drug antibody.

Figure 1: Study Flow Diagram

	Screening	Treatment Period			Follow-up
		Baseline	Weekly Dosing	Endpoint Assessment	
Visit (V)	D -30 to -1	V1	V2 to V32	V33	V34
Week (W)		W0	W1 to W31	W32	W36
Day (D)		D0	D7 to D217	D224	D252

After completion of the 36-week study period, data will be locked and analyzed for safety and efficacy. In addition to the analyses that will be performed following the formal database lock at the end of the study at Week 36, an additional database lock may be performed at Week 16 to help plan potential future development activities.

1.3. Statistical Hypotheses for Trial Objectives

The primary hypothesis is that bermekimab treatment is superior to placebo as assessed by the proportion of participants achieving Hidradenitis Suppurativa Clinical Response (HiSCR) at Week 12.

- HiSCR is defined as at least 50% reduction in total abscess and inflammatory nodule counts (AN count) with no increase in abscess count and no increase in draining fistula count relative to baseline.

1.4. Sample Size Justification

The primary efficacy endpoint is percent of patients who achieve HiSCR at week 12. The sample size is calculated for the bermekimab weekly injection group (Arm 1) versus placebo group (Arm 3) comparison.

Assuming HiSCR is achieved in 60% of intent-to-treat (ITT) patients treated in the bermekimab weekly injection group (Arm 1) versus 30% in placebo treated patients (Arm 3), then, for a randomization ratio of 1:1:1, a sample size of 126 patients (42 in each treatment group) is required for approximately 80% power at a 2-sided 0.05 significance level using Pearson chi-square test. In order to account for possible dropouts and to provide adequate safety data, approximately 150 patients will be randomized and treated.

1.5. Randomization and Blinding

1.5.1. Procedures for Randomization and Stratification

A centralized block randomization schedule with a 1:1:1 allocation ratio will be utilized using an electronic randomization system. Patients will be randomly assigned to Arms 1, 2, or 3. The randomization schedule will be stratified by two factors: prior biologic use (yes/no) and Hurley Stage (II/III).

1.5.2. Maintenance of Blind

The study blind will be maintained for the duration of the study, until after the Final (Week 36) DBL.

In order to maintain blinding, all patients will receive two injections at week 0 and 1, followed by weekly injections from weeks 2 to 31. Additionally, to ensure that no unintentional unblinding occurs during study drug administration, at least one qualified individual will be designated at each study site to perform all preparation and injection of the investigational product. The designated injector(s) should have no other contact with the participant during the study, should not discuss the participant's treatment with the participant, participant's caregiver, or other site personnel at any time, and will not be permitted to review the participant's medical records, questionnaires, or the electronic CRF prior to each injection. The designated injector(s) should be documented in the source documents at each visit.

The investigator will not be provided with randomization codes. The codes will be maintained within the Electronic Data Capture (EDC) system, which has the functionality to allow the investigator to break the blind for an individual participant.

Data that may potentially unblind the intervention assignment (ie, study intervention serum concentrations, anti-bermekimab antibodies, and intervention allocation) will be handled with special care to ensure that the integrity of the blind is maintained and the potential for bias is minimized. This can include making special provisions, such as segregating the data in question from view by the investigators, clinical team, or others as appropriate until the time of DBL and unblinding.

Under normal circumstances, the blind should not be broken until all participants have completed the study and the database is finalized. However, a number of prospectively identified Sponsor individuals will be unblinded at the Week 16 DBL for the purposes of performing data analysis and review. Identification of sponsor personnel who will have access to the unblinded data at participant-level and who will have access to the unblinded data at group-level will be documented prior to unblinding at the Week 16 DBL. All site personnel and participants will remain blinded to the treatment assignments until the last participant completes Week 36 evaluations and the database has been locked.

In the event of an emergency that would require the investigator to be aware of the treatment allocation prior to the end of the trial, the investigator can obtain this information, on a per participant basis, after consultation with the Sponsor's Medical Monitor, if possible. Every effort should be made to consult with the Medical Monitor prior to emergency unblinding. Events that qualify as emergencies are as follows:

- A grade 3 or greater AE which are "probably or definitely" related to study drug, and only if treatment assignment information is essential for the management of the event. This type of reaction would require that the patient receives no further doses, and is followed until the resolution of the toxicity.
- Any suspected, unexpected, serious adverse reaction (SUSAR)
- Pregnancy

The investigator may, in an emergency, determine the identity of the intervention through EDC system. While the responsibility to break the intervention code in emergency situations resides solely with the investigator, it is recommended that the investigator contact the sponsor or its designee if possible to discuss the particular situation, before breaking the blind. Telephone contact with the sponsor or its designee will be available 24 hours per day, 7 days per week. In the event the blind is broken, the sponsor must be informed as soon as possible. The date and reason for unblinding must be documented in the appropriate section of the eCRF, and in the source document. The documentation received from the EDC indicating the code break must be retained with the participant's source documents in a secure manner.

Participants who have had their intervention assignment unblinded by the investigator will not be eligible to receive further study intervention, but should continue to complete early termination visit specified in the Study Calendars section (section 5.2 of the protocol).

2. GENERAL ANALYSIS DEFINITIONS

This analysis plan provides the general analysis definitions and describes the planned subject information, efficacy, safety, pharmacokinetics, and antibody analyses for the two planned DBLs.

2.1. Imputation Rules for Partial or Missing AE Dates

Partial AE onset dates will be imputed as follows:

- If the onset date of an adverse event is missing day only, it will be set to:
  - First day of the month that the AE occurred, if month/year of the onset of AE is different than the month/year of the study agent start
  - The day of study agent start, if the month/year of the onset of AE is the same as month/year of the study agent start date and month/year of the AE resolution date is different
  - The day of study agent start or day of AE resolution date, whichever is the earliest, if month/year of the onset of AE and month/year of the study agent start date and month/year of the AE resolution date are same
- If the onset date of an adverse event is missing both day and month, it will be set to the earliest of:
  - January 1 of the year of onset, as long as this date is on or after the study agent start date
  - Month and day of the study agent start date, if this date is in the same year that the AE occurred
  - Last day of the year if the year of the AE onset is prior to the year of the study agent start date,
  - The AE resolution date.
- Completely missing onset dates will not be imputed.

Partial AE resolution dates not marked as ongoing will be imputed as follows:

- If the resolution date of an adverse event is missing day only, it will be set to the earliest of the last day of the month of occurrence of resolution or the date of death, if the death occurred in that month.
- If the resolution date of an adverse event is missing both day and month, it will be set to the earliest of December 31 of the year or the day and month of the date of death, if the death occurred in that year.
- Completely missing resolution dates will not be imputed.

2.2. Visit Windows

Nominal visits will be used for all by-visit analyses in the study. The study visits scheduled post randomization should occur at the times delineated in the Schedule of Activities. The study visits through Week 36 should occur within ± 3 days of the scheduled visit. The sponsor must be contacted for any significant deviation in the scheduling of a visit outside the appropriate window and determine how the participant should resume his/her normal dosing schedule relative to the baseline visit (Week 0).

2.3. Pooling Algorithm for Analysis Centers

Unless otherwise specified, data from all investigational centers/sites will be pooled for analyses.

2.4. Analysis Sets

There will be two database locks at Weeks 16 and 36 respectively.

Week 16 database lock

Week 16 database lock will include selected efficacy and safety data through Week 16 and PK data through Week 16 for all randomized and treated participants. The data will be primarily used in planning for future bermekimab clinical development in HS. The selected efficacy and safety data include but not limited to following:

1. Lesion count [includes Hidradenitis Suppurativa Clinical Response (HiSCR)]
2. Numeric Rating Scale for Pain & Itch (NRS Pain & Itch) score
3. HS Symptom Diary (HSSD) HS related Pain Symptom Score
4. Baseline Disease Characteristics for 1-3 above
5. Demography
6. AE

Week 36 database lock

Week 36 DBL would include all data through Week 36 for all randomized and treated participants.

2.4.1. Efficacy Analysis Set(s)

2.4.1.1. Full Analysis Set

The full analysis set (FAS) includes all randomized participants who received at least 1 dose of study agent. This analysis set will be used for the efficacy analyses of the endpoints through Week 36, ***unless otherwise specified***.

In the efficacy analyses, participants will be analyzed according to the randomized treatment groups they were assigned to, regardless of the treatments they actually received.

2.4.1.2. Per Protocol Analysis Set

The primary endpoint will also be analyzed using the per-protocol population, which includes participants who are generally compliant with the protocol.

Specifically, the per protocol population includes participants in FAS except those

- who did not meet the inclusion criteria 7 in the protocol as listed below:
  - A total body count of abscesses and inflammatory nodules (A+N) of at least 3.
- who met the exclusion criteria 10 in the protocol as listed below:
  - Has a draining fistula count of >20 at the baseline visit.
- who did not receive all scheduled study agent administrations prior to Week 12.

However, for those who discontinued the study agent prior to Week 12 will be included in the per-protocol analysis and the data handling rule specified in Section 5.2.2 will apply. Participants who were excluded from the per protocol analyses will also be summarized.

2.4.2. Safety Analysis Set

The safety analysis set includes all randomized participants who received at least 1 dose of study agent, i.e., the treated population.

2.4.3. Pharmacokinetics Analysis Set

The pharmacokinetic (PK) analysis set includes all participants who received at least 1 complete dose of bermekimab and had at least 1 valid blood sample drawn for PK analysis after their first dose of bermekimab.

2.4.4. Pharmacodynamics Analysis Set

The Pharmacodynamics (PD) analysis set includes all participants who received at least 1 dose of bermekimab and had at least 1 valid blood sample drawn for PD(including IL-6) analysis after their first dose of bermekimab.

2.4.5. Immunogenicity Analysis Set

The immunogenicity analysis set includes all participants who received at least 1 dose of bermekimab and who have at least 1 sample obtained after their first dose of bermekimab for the detection of antibodies to bermekimab.

2.5. Definition of Subgroups

To evaluate the consistency in the primary efficacy endpoint (proportion of participants who achieve HiSCR at Week 12) over demographics, baseline disease characteristics, and prior HS medication use, subgroup analyses will be performed when sample sizes are permitted. The subgroups include, but are not limited to, the following:

Baseline demographics:

- Sex (male, female)
- Race
- Baseline Age (<40 years, 40 to <65 years, ≥65 years)
- Baseline weight (< median, ≥ median)
- BMI (Normal [<25], Overweight [25 -<30], Obese [≥ 30])

Baseline disease characteristics:

- Baseline Hurley stage status (II, and III)
- Age at diagnosis (years) (< median, ≥ median)
- HS disease duration (years) (< median, ≥ median)
- Baseline AN count category (3 to 5, 6 to 10, ≥11)
- Baseline AN count (< median, ≥ median)
- Baseline HS-PGA (<3, 3, 4, 5)
- Baseline C-reactive protein (CRP) level (< median, ≥ median)
- Baseline Erythrocyte Sedimentation Rate (ESR) (< median, ≥ median)

HS medication:

- Prior biologic therapy used
  - Never used
  - Ever used

2.6. Study Day and Relative Day

Study day is defined as the number of days from the study reference date to the event/visit date. It will be calculated as follows:

- If the event/assessment occurs on or after the reference date, then study day = event/assessment date – reference date + 1.
- If the event/assessment occurs before the reference date, then study day = event/assessment date – reference date.

Hence, the day of reference date is Study Day 1; the previous day is Study Day -1.

2.7. Baseline and Endpoint

In general, the baseline measurement is defined as the closest measurement taken prior to or at the time of the first study agent administration date unless otherwise specified.

3. INTERIM ANALYSIS AND DATA MONITORING COMMITTEE REVIEW

No interim analysis and data monitoring committee review is planned for this study.

4. SUBJECT INFORMATION

Unless specified otherwise, FAS will be used for the subject information analyses and participants will be analyzed according to the treatment group to which they were randomized at Week 0, regardless of the treatment they actually received. The number of participants in FAS will be summarized by treatment group and overall.

Simple descriptive statistics, such as mean, median, standard deviation, interquartile range, maximum, and minimum for continuous variables, and counts and percentages for discrete variables will be used to summarize most data and no formal hypothesis testing for treatment comparisons will be performed. In addition, participant listings will also be used to present the data.

4.1. Demographics and Baseline Characteristics

The FAS will be used for all tabulations. Participants' demographic data and baseline disease characteristics will be summarized by treatment group. If imbalances are found at baseline, then additional analyses may be performed adjusting for baseline differences.

4.1.1. Demographic

Table 1 presents a list of the demographic variables that will be summarized by treatment group, and overall for the FAS.

Table 1: Demographic Variables

Continuous Variables:	**Summary Type**
Age (years)	Descriptive statistics (N, mean, standard deviation [SD], median and range [minimum and maximum], and IQ range).
Weight (kg)	
Height (cm)	
Categorical Variables:	
Age (<40 years, 40 to <65 years, and ≥65 years)	Frequency distribution with the number and percentage of participants in each category.
Sex (male, female)	
Weight (kg) (≤ 95 kg, > 95 kg)	
Race (American Indian or Alaska Native, Asian, Black or African American, White, Other)	
Ethnicity (Hispanic or Latino, Not Hispanic or Latino)	
BMI (Normal [<25], Overweight [25 -<30], Obese [≥30])	

4.1.2. Baseline Disease Characteristics

Baseline disease severity measurements (e.g., duration of HS disease, baseline Hurley stage status, baseline AN count, inflammatory nodule count, abscess count, and draining and non-draining fistula count, HS-PGA, Modified Hidradenitis Suppurativa Score) and baseline patient reported outcomes (e.g, DLQI, HSSD, NRS Pain, NRS Itch, EQ-5D-3L, HADS) will be summarized by treatment group.

4.1.3. Medical History

Summaries of participants' surgical history, alcohol intake, and smoking status will be provided by treatment group. Medical history will be listed including body systems and condition/non-drug allergy as captured on the eCRF.

4.2. Disposition Information

The number of participants in the following disposition categories will be summarized by treatment group and overall:

- Participants screened: met the inclusion criteria regarding the target indication and signed the informed consent
- Participants randomized
- Participants randomized and received at least 1 study agent
- Participants who completed the study through Week 16 and Week 36
- Participants who discontinued study and the reasons for discontinuation through Week 16 and Week 36

Listings of participants will be provided for the following categories:

- Participants who discontinued study Participants who were unblinded during the study period

4.3. Treatment Compliance

Study agent compliance will be summarized descriptively through Week 16 and through the end of study (Week 36) for the FAS. Number of the participants receiving each scheduled treatment will be summarized. In addition, treatment compliance will also be assessed by protocol deviations related to study agent administration (i.e., incorrect study agent or dose received).

4.4. Extent of Exposure

The exposure data through Week 16 and through Week 36 will be summarized. The number and percentage of participants who receive study agent will be summarized by treatment group for the safety analysis set. Descriptive statistics will be presented for the following parameters:

- Number of administrations
- Cumulative total dose

In addition, the average exposure (number of administrations) and average duration of follow-up (weeks) will also be summarized by treatment group in the safety tables through different study time periods.

4.5. Protocol Deviations

In general, the following list of major protocol deviations may have the potential to impact participants' rights, safety or well-being, or the integrity and/or result of the clinical trial. Participants with major protocol deviations will be identified prior to database lock and will be summarized by category by treatment group through Week 36 for the FAS.

- Entered study and did not satisfy entry criteria
- Received wrong treatment or incorrect dose
- Received excluded concomitant or non-drug therapy
- Out of window visit
- Out of window procedure
- Other

A listing of participants with major protocol deviations and a listing of patients who missed scheduled study agent administration will also be provided by randomized treatment group. In addition, analyses of COVID-19 related protocol deviations will be provided.

4.6. Prior and Concomitant Medications/Surgeries

Prior and concomitant medications will be summarized by treatment group for the FAS. Prior medications are defined as any therapy used before the day of first dose (partial or complete) of study agent. Concomitant medications are defined as any therapy used on or after the same day as the first dose of study agent, including those that started before and continue on after the first dose of study agent.

Prior biologic used that was captured in eCRF, summary of concomitant medications with potential to improve HS and summary of concomitant analgesic therapy will be summarized by randomized treatment group for the FAS. a listing of participants who received a concomitant treatment that could improve HS, a listing of participants who received concomitant analgesic therapy, and a listing of participants who received concomitant corticosteroids will be provided.

5. EFFICACY

5.1. General Method of Analysis

Unless specified otherwise, efficacy data summaries will be provided by treatment group for the FAS. Statistical comparisons will be made between each of the bermekimab treatment groups (400 mg SC qw, 400 mg SC q2w) and the placebo treatment group.

In general, for response efficacy endpoints, treatment comparisons will be performed using a Cochran-Mantel-Haenszel (CMH) test stratified by the prior biologic use (yes/no) and screening Hurley stage (II/III). For continuous efficacy endpoints, treatment comparisons will be performed using either a Mixed-Effect Model Repeated Measure (MMRM) model. All of the models will have treatment group, prior biologic use (yes/no), screening Hurley stage (II/III), and baseline value for the corresponding efficacy endpoint as explanatory factors. The MMRM model will also include visit, treatment group by visit interaction, prior biologic use by visit, screening Hurley stage by visit, and baseline value by visit interaction as additional explanatory factors. In addition, treatment differences and their associated 95% confidence intervals will be presented.

For the MMRM model, an unstructured covariance matrix for repeated measure within a participant will be used. In case that convergence cannot be achieved, autoregressive (1) covariance structure will be used.

Descriptive statistics, such as mean, median, standard deviation, minimum and maximum, interquartile range for continuous variables, and counts and percentages for categorical variables will be used to summarize the data. Graphical data displays and subject listings may also be used to summarize the data.

5.2. Analysis Specifications

5.2.1. Level of Significance

Unless otherwise specified, all statistical testing procedures will be performed at a 2-sided significance level of 0.05. Multiplicity adjustment was planned for the primary endpoint as specified in protocol Section 10.5.3 to maintain an overall Type I error of 0.05 or less for the primary endpoint analysis. Nominal p-values will be reported for other secondary analyses.

However, since this is a Phase 2 trial and stricter multiplicity adjustment for the primary endpoint at an overall alpha level of 0.05 is not required; a less stringent type I error controlled will be implemented instead. The study would be considered as positive if any of the comparisons for primary endpoint analysis (weekly injection group vs placebo, and every other week injection group vs placebo) is less than or equal to 0.05. Therefore, an overall Type I error rate will be maintained at 0.1 or less for the primary endpoint analysis. Nominal p-values will be reported for secondary analyses.

5.2.2. Data Handling Rules

No imputation will be performed for missing baseline values.

For the main analysis of primary endpoint of HiSCR at Week 12, participants who discontinued treatment will be considered as a non-responder from that point onward. The remaining missing data will be imputed as non-responder. Supplementary analyses will be performed and include (1) using observed data regardless of whether the participant discontinued treatment or have missing data, and (2) data after treatment discontinuation will be considered as missing, missing response will be imputed by multiple imputation methods under the assumption of missing at random. Similar data handling rules (i.e., main analysis with non-responder imputation and

2 supplementary analyses) will also be applied for HiSCR response over time through Week 16. The following data handling rules will be applied for other endpoints:

Binary secondary endpoints:

- Analyses will be similar to the main analysis of the primary endpoint. A participant who discontinued treatment will be considered as a non-responder from that point onward. The remaining missing data will be imputed as non-responder.

Continuous secondary endpoints:

- If a participant discontinued treatment, zero will be assigned to improvement or percent improvement. Remaining missing data will not be imputed. A MMRM model will adjust for missing post baseline data and all available data from the 3 treatment groups through Week 16 or from Week 16 through Week 36 will be included. Under the assumption of missing at random (MAR), the missing data will be accounted for through correlation of repeated measures in the model.
- In addition, supplementary analyses will be performed for the analyses related to (a) change from baseline in AN count (Section 5.4.3.2.2) and (b) change from baseline in HSSD past 24 hours pain (Section 5.4.3.5.3.2) endpoints using
  - Observed data regardless of whether the subject discontinued treatment or have missing data.
    - Data after treatment discontinuation will be considered as missing, missing data will be addressed by MMRM model.

5.2.3. Treatment Groups

In the efficacy analyses, full analysis set will be used and the participants will be analyzed according to the randomized treatment groups they were assigned to, regardless of the treatments they actually received. Unless otherwise specified, efficacy analyses during each of study periods below are in general displayed as follows:

- Analysis through Week 16
  - Efficacy analyses for placebo comparison at Week 16 or through Week 16 will be summarized by randomized treatment group at Week 0:
    - **Placebo**: Participants randomized to placebo group at Week 0.
    - **Bermekimab 400 mg q2w**: Participants randomized to bermekimab 400 mg SC q2w group at Week 0.
    - **Bermekimab 400 mg qw**: Participants randomized to bermekimab 400 mg qw group at Week 0.
    - **Combined Column**: Combined column which combines the data from the randomized and treated participants who are assigned to bermekimab treatment groups at baseline.

- Analysis from Week 16 through Week 36
  - Efficacy data from Week 16 through Week 36 will be summarized by the following treatment groups:
    - **Placebo → Bermekimab 400 mg qw**: efficacy in participants randomized to placebo group at week 0 who crossed over to receive Bermekimab 400 mg qw at Week 16.
      - Only placebo participants crossover to bermekimab 400 mg SC qw and received bermekimab at Week 16 will be included in the summary for the visits after Week 16.
    - **Bermekimab 400 mg q2w**: Participants randomized to bermekimab 400 mg SC q2w group at Week 0.
    - **Bermekimab400 mg qw**: Participants randomized to bermekimab 400 mg SC qw group at Week 0.
    - **Combined Column**: Combined column which combines the data from the randomized and treated participants who are assigned to bermekimab treatment groups at baseline.

These presentations allow assessment of efficacy over time between treatment groups.

5.3. Primary Efficacy Endpoint(s)

5.3.1. Definition

The primary efficacy endpoint is the proportion of participants achieve a HiSCR at Week 12. HiSCR is defined as a ≥ 50% reduction in inflammatory lesion AN count (sum of abscesses and inflammatory nodules), and no increase in abscesses or draining fistulas in HS compared with the lesions counted at baseline.

5.3.2. Estimands

Primary Estimand:

The **Composite Estimand** will be targeted for the primary endpoint. The Composite Estimand for the primary endpoint is defined by the following 5 attributes:

- Treatment:
  - Bermekimab 400mg q2w
  - Bermekimab 400mg qw
  - Placebo

  Note: treatment regimen include protocol specified antibiotic rescue therapy/lesion intervention as needed.
- Population: Adult participants with moderate to severe Hidradenitis Suppurativa for at least 1 year.
- Variable: HiSCR binary responder variable at Week 12, where a participant who discontinues treatment is considered as non-responder after discontinuation.

- Intercurrent event: The intercurrent event of treatment discontinuation is addressed by the composite strategy, as reflected in the variable definition.
- Population-level summary: Difference in the proportion of participants achieving HiSCR at Week 12 between the bermekimab and placebo treatment groups.

Supplementary Estimands for primary endpoint:

Supplementary Estimand 1

Same components as the primary estimand, except for the strategy used for treatment discontinuation:

Hypothetical strategy: assess the treatment effect as if the participants would not have discontinued treatment

Note: Under this hypothetical strategy, the data of HiSCR status after treatment discontinuation will not be used and will be set to missing in analysis.

Supplementary Estimand 2

Same components as the primary estimand, except for the strategy used for treatment discontinuation:

Treatment policy strategy: assess the treatment effect regardless of treatment discontinuation.

Note: Under the treatment policy strategy, the data of HiSCR status after treatment discontinuation will also be used in analysis.

5.3.3. Data Handling

Participants who discontinued treatment prior to Week 12 will be considered HiSCR non-responder at Week 12. In addition, participants with missing lesion count evaluations at Week 12 will be considered as HiSCR non-responders at Week 12.

5.3.4. Analysis Methods

The primary endpoint will be compared between the each bermekimab group and the placebo group. The primary endpoint will be analyzed at Week 12 based on the composite strategy (Section 5.3.2) and the data from all participants in FAS (Section 2.4.1.1) will be analyzed according to randomized treatment group regardless of the treatment actually received.

In this primary analysis, the number and proportion of participants who achieve a HiSCR at Week 16 will be summarized for each treatment group. To address the primary objective, a CMH chi-square statistic stratified by stratified by prior biologic use (yes/no) and screening Hurley stage (II/III) at an alpha level of 0.05 will be used.

The proportion of participants who achieve HiSCR at Week 12 will be compared between each of the 2 bermekimab treatment groups and the placebo group. The proportion difference between each bermekimab group and placebo group and its 2-sided 95% CI will be provided based on normal approximation with Mantel-Haenszel weights adjusting for prior biologic use (yes/no) and screening Hurley stage (II/III).

The study would be claimed as positive if any of the comparisons for primary endpoint analysis (weekly injection group vs placebo, and every other week injection group vs placebo) is less than or equal to 0.05. Therefore, an overall Type I error rate will be maintained at 0.1 or less for the primary endpoint analysis.

5.3.5. Analysis Corresponding to Supplement Estimand

5.3.5.1. Supplementary Analysis 1 (Hypothetical Estimand)

The first supplementary analysis will be performed utilizing the Hypothetical Estimand by using multiple imputations (MI)[1] by treatment group based on HiSCR responses respectively. For participants who do not return for evaluation at Week 12 after discontinuation of study treatment for any reason, the data will be considered missing. The fully conditional specification (FCS) method will be used to impute the missing data for the HiSCR response at Week 12.

More specifically, the missing HiSCR response status will be imputed using FCS method with 500 imputed data sets with seed = 1231 to fill in the missing HiSCR responses through Week 12 in each of the 500 copies of datasets. The proportion difference of HiSCR at Week 12 adjusted for prior biologic use (yes/no) and screening Hurley stage (II/III) using Mantel-Haenszel weight between the bermekimab groups and the placebo group and its 95% CI combining multiple datasets will be provided. In addition, the p-value for testing the treatment difference will be obtained from SAS PROC MIANALYZE based on these combined Mantel-Haenszel estimates from the multiple imputation datasets.

5.3.5.2. Supplementary Analysis 2 (Treatment Policy Estimand)

The primary endpoint will be analyzed utilizing the Treatment Policy Estimand. For participants who do not return for evaluation at Week 12, the analysis will be performed using observed data regardless of study treatment discontinuation.

5.3.6. Subgroup Analysis

For each of the subgroups defined in Section 2.5, the difference between each of the bermekimab treatment group and placebo group in the proportion of participants achieving HiSCR at Week 12 and its 95% continuity adjusted confidence interval (when the number of participants permits) will be calculated. Subgroup analyses will not be stratified by prior biologic use (yes/no) and screening Hurley stage (II/III).

In addition, the proportion of participants achieving HiSCR at Week 12 by investigator site/region will be summarized. The same data handling rules specified in Section 5.2.2 will apply for subgroup analysis.

5.3.7. Per-protocol Analysis

The primary analyses for primary endpoint will also be performed on per-protocol analysis set (Section 2.4.1.2). Similar data handling rules specified in Section 5.2.2 will apply.

5.4. Secondary Efficacy Variable(s)

In addition to the primary, the analyses for secondary efficacy endpoints will be performed and nominal p-values will be provided. Other efficacy endpoints include the endpoints related to

- HiSCR
- Lesion Count
- Modified Hidradenitis Suppurativa Score (mHSS)
- HS-PGA
- Patient-reported Outcomes
  - Numerical Rating Scale (NRS) for Pain & Itch [take home patient diary]
  - DLQI
  - HSSD
  - HADS
  - Health Status Questionnaire (EQ-5D-3L)
  - Patient's Global Impression of Severity (PGI-s)
  - Patient's Global Impression of Change (PGI-c)
- Exploratory endpoints
  - International Hidradenitis Suppurativa Severity Score System (IHS4)
  - Hidradenitis Suppurativa Quality of Life Measure (HiSQOL)
  - Physician's Global Assessment of Inflammatory Skin Disease (PGA-ISD)

5.4.1. Analysis Methods

The secondary efficacy analyses outlined in the following sections in general will be carried out at key endpoints (Week 12 and Week 16) and over time.

Most of the secondary efficacy analyses described in sections below will be based on FAS. For participants randomized to placebo, only participants who crossed over to receive bermekimab at Week 16 will be included in the efficacy summaries for the visits after Week 16. All statistical testing will be performed at the 2-sided 0.05 significance level. Nominal p-values will be presented.

5.4.1.1. At Week 12 and Week 16

For continuous endpoints where longitudinal data were collected prior to the analysis visits; treatment comparisons will be performed using Mixed-Effect Model Repeated Measure (MMRM) model, with fixed effects for treatment, week, stratification factors (prior biologic use (yes/no), screening Hurley stage (II/III)) when sample size in each category of the stratification variables is permitted, and baseline value in the model. The MMRM model will also include prior biologic use by visit, screening Hurley stage by visit, and baseline value by visit interaction and the treatment-by-week interaction as additional explanatory factors.

The model will be performed based on all available data from the 3 treatment groups through Week 16. The treatment difference between a guselkumab group and the placebo group will be estimated by the difference in the least squares means (LSmeans). The 95% CIs for the differences in LSmeans and p-values will be calculated.

For binary response endpoints treatment comparisons will be performed using a Cochran-Mantel-Haenszel (CMH) test stratified by the prior biologic use (yes/no) and screening Hurley stage (II/III) with missing data imputed with Non-Responder Imputation (NRI) defined in Section 5.2.2.

For the PRO endpoints with ordinal scale, the Cochran-Mantel-Haenszel (CMH) mean scores test stratified by prior biologic use (yes/no) and screening Hurley stage (II/III) will be used to compare bermekimab and placebo. The distribution of ordinal scale category will also be provided.

5.4.1.2. Over Time Summaries

In general, all endpoints with over time analyses will be descriptively summarized by treatment groups using descriptive statistics, such as mean, median, standard deviation, interquartile range, minimum and maximum for continuous variables, and counts and percentages for categorical variables.

For continuous secondary endpoints, the change from baseline at each week will be analyzed using a restricted maximum likelihood (REML)-based mixed model for repeated measures (MMRM) with fixed effects for treatment, visit, stratification factors (prior biologic use (yes/no) and screening Hurley stage (II/III)), baseline value, prior biologic use by visit, screening Hurley stage by visit, baseline value by week interaction, and the treatment-by-week interaction. Least Square means (LSmeans), LS means differences and their corresponding 95% confidence interval will be provided over time through Week 16 by visit and treatment group in addition to the descriptive summary statistics. For over time summaries from Week 16 through Week 36. Least Square means (LSmeans) and their corresponding 95% confidence interval will be provided.

Additionally, graphical data displays may also be used to summarize the over time data if applicable.

5.4.2. Data Handling Rule

Unless otherwise specified, data handling rules specified in Section 5.2.2 will be applied to these secondary analyses. For binary endpoints, participants with missing response status will be considered non-responders for the over time analyses.

5.4.3. Secondary Efficacy Endpoints

5.4.3.1. Hidradenitis Suppurativa Clinical Response (HiSCR)

5.4.3.1.1. Definition

- HiSCR75 is defined as a ≥ 75% reduction in inflammatory lesion AN count (sum of abscesses and inflammatory nodules) from baseline and no increase in abscesses or draining fistulas in HS compared with the lesions counted at baseline.
- HiSCR90 is defined as a ≥ 90% reduction in inflammatory lesion AN count (sum of abscesses and inflammatory nodules) from baseline and no increase in abscesses or draining fistulas in HS compared with the lesions counted at baseline.

Please also refer to Section 5.3.1 for the definition of the HiSCR responder status.

5.4.3.1.2. Analysis Related to HiSCR

- The proportion of participants who achieved HiSCR at Week 16 will be compared between each of the bermekimab groups and the placebo group.
- The proportion of participants who achieved HiSCR75 and HiSCR90 at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group.
- The proportion of participants who achieved HiSCR, the proportion of participants who achieved HiSCR by baseline weight category (≤ 95 kg, > 95 kg), and the proportion of participants who achieved HiSCR by baseline Hurley stage (II/III) will be summarized by treatment group over time.
- The proportion of participants who achieved HiSCR75 and HiSCR90 will be summarized by treatment group over time.

Subgroup analyses specified in Section 5.3.6 will also be performed at Week 16. In addition, for participants who discontinued study treatment or have missing data, the proportion of participants who achieve HiSCR through Week 16 will be analyzed utilizing the hypothetical estimand (i.e. MI method) and treatment policy estimand (i.e. observed data).

5.4.3.2. Lesion Count

5.4.3.2.1. Definition

Lesion counts are defined as the number of abscesses, inflammatory nodules, non-inflammatory nodules, draining fistulas, and non-draining fistulas. Lesions will be counted during each visit.

- AN count is defined as total abscess and inflammatory nodule count.
- AN50 is defined as at least a 50% reduction in AN count relative to baseline.

- AN75 is defined as at least a 75% reduction in AN count relative to baseline.
- AN90 is defined as at least a 90% reduction in AN count relative to baseline.
- AN100 is defined as a 100% reduction in AN count relative to baseline (ie, AN count = 0).

5.4.3.2.2. Analysis Related to Lesion Count

AN Count

- The change from baseline in total AN count at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group respectively.
- The change from baseline in total AN count will be summarized by treatment groups over time.
- The proportion of participants achieving at least 50%, 75%, 90%, and 100% reduction in total AN count at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group.
- The proportion of participants achieving at least 50%, 75%, 90%, and 100% reduction in total AN count will be summarized by treatment group over time. In addition, the proportion of participants achieving at least 50%, 75%, 90%, and 100% reduction in total AN count will be summarized in subgroups of baseline weight category (≤ 95 kg, > 95 kg) and baseline Hurley stage (II/III) by treatment group over time.
- The proportion of participants achieving AN count of 0/1 and 0/1/2 respectively at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group.
- The proportion of participants who experienced at least one flare (defined as at least a 25% increase in AN count with a minimum increase of 2 AN count relative to baseline) over 16 weeks will be compared between each of the bermekimab groups and the placebo group at Week 12 and Week 16. Furthermore, the proportion of participants who experience at least one flare over 36 weeks will be summarized at Week 36 by treatment groups.

In addition, two supplementary analyses will be performed for the analyses (a) change from baseline in AN count at Week 12 and week 16 and (b) change from baseline in AN count overtime (Section 5.2.2) by using

- Observed data regardless of whether the subject discontinued treatment or have missing data.
- Data after treatment discontinuation will be considered as missing, missing data will be addressed by MMRM model.

Abscess

- The proportion of participants who achieved complete elimination of abscesses at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group among participants who had any abscess at baseline.

- The change from baseline in number of abscesses at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group respectively.
- The change from baseline in number of abscesses will be summarized by treatment group over time.

Fistulas

- The proportion of participants who achieved complete elimination of draining fistulas at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group among those with draining fistulas at baseline.
- The change from baseline in number of draining fistulas at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group respectively.
- The change from baseline in number of draining fistulas and non-draining fistula will be summarized by treatment group over time.

Inflammatory Nodules non-inflammatory Nodules

- The proportion of participants who achieved complete elimination of inflammatory nodules at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group among those with inflammatory nodules at baseline.
- The change from baseline in number of inflammatory nodules at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group respectively.
- The change from baseline in number of inflammatory nodules will be summarized by treatment group over time.
- The change from baseline in number of non-inflammatory nodules will be summarized by treatment group over time.

5.4.3.3. Modified Hidradenitis Suppurativa Score (mHSS) (Modified Sartorius Score)

5.4.3.3.1. Definition

The Sartorius Scale is used to quantify the severity of HS. Points are awarded for 12 body areas (left and right axillae, left and right sub/inframammary areas, intermammary area, left and right buttocks, left and right inguino-crural folds, perianal area, perineal area, and other): points were awarded for nodules (2 points for each); abscesses (4 points); fistulas (4 points); scars (1 point); other findings (1 point); and longest distance between two lesions (2-6 points, 0 if no lesions); and if lesions are separated by normal skin (yes-0 points; no-6 points). The total Sartorius score is the sum of the 12 regional scores.

5.4.3.3.2. Analysis Related to Modified Hidradenitis Suppurativa Score (mHSS)

- The change from baseline in modified Hidradenitis Suppurativa score at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group respectively.
- The change from baseline in modified Hidradenitis Suppurativa score will be summarized by treatment group over time.

5.4.3.4. Hidradenitis Suppurativa Physician's Global Assessment (HS-PGA)

5.4.3.4.1. Definition

HS-PGA is defined as: a) clear when the total number of abscesses is 0, the total number of draining fistulas is 0, the total number of inflammatory nodules is 0 and the total number of non-inflammatory nodules is 0; b) minimal when the total number of abscesses is 0, the total number of draining fistulas is 0, the total number of inflammatory nodules is 0 and there is presence of non-inflammatory nodules; c) mild when the total number of abscesses is 0, the total number of draining fistulas is 0, and the total number of inflammatory nodules is 1-4, or when there is presence of one abscess or draining fistula and absence of any inflammatory nodules; d) moderate when the total number of abscesses is 0, the total number of draining fistulas is 0 and the total number of inflammatory nodules is at least 5; or when there is presence of one abscess or draining fistula and at least one inflammatory nodule; or when there are 2-5 abscesses or draining fistulas and fewer than 10 inflammatory nodules; e) severe when the total number of abscesses or draining fistulas is 2-5 and the total number of inflammatory nodules is at least 10; and f) very severe when there are more than 5 abscesses or draining fistulas[1].

5.4.3.4.2. Analysis Related to HS-PGA

- The proportion of participants who achieved HS-PGA of clear (0), clear (0) or minimal (1) or HS-PGA score of mild or better (≤ 2) with at least a 2-grade improvement relative to baseline at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group respectively.
- The proportion of participants with 1-grade and 2-grade improvement in HS-PGA scores from baseline will be summarized over time respectively.
- The proportion of participants who achieved HS-PGA of clear (0), clear (0) or minimal (1) or HS-PGA score of mild or better (≤ 2) with at least a 2-grade improvement relative to baseline will be summarized over time by treatment group.

5.4.3.5. Patient Reported Outcome

5.4.3.5.1. Numerical Rating Scale (NRS) for Pain and Itch [take home diary]

5.4.3.5.1.1. Definition

Patients will be given a diary to complete each night before bed. Patients will be asked to report "average pain", and "worst moment pain" as well as "average itch" and "worst moment itch" on a 0-10 numeric rating scale. Patient diaries will be collected weekly. A higher score indicates more severe disease (0 indicate no itch or no pain). Both pain and itch rating scale will be averaged into

a weekly score respectively (ie 7 days [from day -7 to -1] prior to a visit). Four days out of 7 days (either consecutive or nonconsecutive) are necessary to derive a weekly score; otherwise data are considered missing for that week. The baseline is defined as the average score of last 7 days prior to Week 0 study agent administration. Similarly, If there are more than 3 days missing data, then baseline is set as missing. Missing baseline will not be imputed.

- NRS30 for numerical rating pain score (weekly average): at least 30% reduction from baseline in numerical rating scale for pain.

5.4.3.5.1.2. Analysis Related to NRS (Pain & Itch) Score

- Proportion of participants who achieved NRS30 for pain score – at worst (weekly average) at Week 12 and Week 16 among participants with baseline NRS pain score ≥ 3 will be compared between each of the bermekimab groups and the placebo group respectively.
- Proportion of participants who achieved NRS30 for pain score – on average (weekly average) at Week 12 and Week 16 among participants with baseline NRS pain score ≥ 3 will be compared between each of the bermekimab groups and the placebo group respectively.
- Proportion of participants who achieved NRS30 – at worst (weekly average) and on average (weekly average) will be summarized over time by treatment group among participants with baseline NRS score ≥ 3 respectively.
- Proportion of participants who achieve ≥3-point and ≥4-point improvement (reduction) in NRS itch score – at worst (weekly average) from baseline at Week 12 and Week 16 will be compared between each bermekimab group and the placebo group among participants with baseline NRS itch score ≥ 3 and ≥4 respectively.
- Proportion of participants who achieve ≥3-point and ≥4-point improvement (reduction) in NRS itch score – on average (weekly average) from baseline at Week 12 and Week 16 will be compared between each bermekimab group and the placebo group among participants with baseline NRS itch score ≥ 3 and ≥4 respectively.
- Proportion of participants who achieve ≥3-point and ≥4-point improvement (reduction) in NRS itch score – at worst (weekly average) and on average (weekly average) from baseline will be summarized over time by treatment group among participants with baseline NRS itch score ≥ 3 and ≥ 4 respectively.
- Change from baseline in NRS pain score and itch score – at worst (weekly average) at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group respectively.
- Change from baseline in NRS pain score and itch score – on average (weekly average) at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group respectively.
- Change from baseline in NRS pain and itch score -at worst (weekly average) and on average (weekly average) will be summarized over time by treatment group.

5.4.3.5.2. DLQI

5.4.3.5.2.1. Definition

The DLQI is a dermatology-specific quality of life instrument designed to assess the impact of the disease on a participant's quality of life. It is a 10-item questionnaire that, in addition to evaluating overall quality of life, can be used to assess 6 different aspects that may affect quality of life: symptoms and feelings, daily activities, leisure, work or school performance, personal relationships, and treatment. The DLQI is calculated by summing the score of each question resulting in a maximum of 30 and a minimum of 0. A higher score indicates more severe disease. A score of ≤1 indicates no effect at all of disease on participant's health related quality of life, and a reduction of 5 points or more in total DLQI score is considered clinically meaningful improvement.

For a partially answered questionnaire (eg, not all 10 answers in the DLQI questionnaire were available):

- If one question's answer is not available, this question will be scored 0. The total score will then be calculated.
- If two or more questions' answers are unavailable, the questionnaire is not scored. Hence, the total score and each of the 6 component scores will be set to missing.
- If one question from one of the 6 component scores is missing, the affected component score will be set to be missing.

5.4.3.5.2.2. Analysis related to DLQI

- The change from baseline in DLQI scores at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group respectively. In addition, the change from baseline in DLQI scores will also be summarized over time.
- The proportions of participants with DLQI score of 0 and 1 at Week 12 and Week 16 for the participants with baseline DLQI score >1 will be compared between each of the bermekimab groups and the placebo group respectively and will also be summarized by treatment groups over time.
- The proportion of participants with a reduction of 5 or more points in DLQI score for the participants with baseline DLQI score ≥ 5 at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group and will also be summarized by treatment groups over time.
- The change from baseline in DLQI component scores at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group.

5.4.3.5.3. Hidradenitis Suppurativa Symptoms Diary (HSSD)

5.4.3.5.3.1. Definition

The HSSD is a 7-item patient self-reported questionnaire that assesses 5 HS-related symptoms including pain, tenderness, hot skin feeling, odor, and itchiness. The participants are asked to rate the severity of each symptom on a 0 to 10 numerical rating scale, with 0 indicating no symptom

experience and 10 indicating the worst possible symptom experience. All 5 symptoms have a recall period of the past 7 days, except for 2 additional questions on pain which evaluate current pain and pain in the past 24 hours. Each individual symptom scale score, ranging from 0-10, will be summarized. A total symptom score, which will also range from 0-10, will be derived by averaging the 5 individual scale scores that utilize the past 7-day recall period.

5.4.3.5.3.2. Analysis Related to Hidradenitis Suppurativa Symptoms Diary (HSSD)

- The change from baseline in HS-related skin pain in the past 24 hours based on HSSD at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group respectively.
- The change from baseline in each HSSD symptom scale score (7 days recall) and the change from baseline in current pain at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group.
- The change from baseline in HSSD total symptom score at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group.
- Number of participants with a reduction of ½ SD or more from baseline in HSSD total symptom score at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group.
- The change from baseline in each HSSD symptom scale score (7 days recall), the change from baseline for the past 24 hours pain and current pain, and the change from baseline in HSSD total symptom score will be summarized over time.

In addition, two supplementary analyses will be performed for the analyses (a) change from baseline in HS-related skin pain in the past 24 hours based on HSSD at Week 12 and week 16 and (b) change from baseline in HS-related skin pain in the past 24 hours based on HSSD overtime (Section 5.2.2) by using

- Observed data regardless of whether the subject discontinued treatment or have missing data.
- Data after treatment discontinuation will be considered as missing, missing data will be addressed by MMRM model.

5.4.3.5.4. Hospital Anxiety and Depression (HADS)

5.4.3.5.4.1. Definition

The HADS is an instrument for screening anxiety and depression in non-psychiatric populations; repeated administration also provides information about changes to a patient's emotional state. The HADS consists of 14 items, 7 each for anxiety and depression symptoms; Each item on the questionnaire is scored from 0-3 resulting in a score between 0 and 21 for either anxiety or depression. The following cut-off scores are recommended for both subscales: 7 to 8 for possible presence, 10 to 11 for probable presence, and 14 to 15 for severe anxiety or depression. The HADS is provided in the study reference manual.

If one or more of the items within each domain are left unanswered, that HADS component score will be considered missing.

5.4.3.5.4.2. Analysis Related to Hospital Anxiety and Depression

- The change from baseline in Hospital Anxiety and Depression scale (HADs) at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group.
- The change from baseline in Hospital Anxiety and Depression scale will be summarized over time by treatment groups.
- The proportion of participants with hospital anxiety scale <8 and depression scale <8 at Week 12 and Week 16 will be compared between each of the bermekimab groups and placebo groups, among participants with a baseline hospital anxiety scale and depression scale >=8.
- Hospital anxiety scale and depression scale shift from baseline with respect to the hospital anxiety and depression scale category (<8, >=8) will be summarized over time.

5.4.3.5.5. EuroQol Health Status Questionnaire (EQ-5D-3L)

5.4.3.5.5.1. Definition

The EQ-5D-3L consists of 2 pages: the EQ-5D descriptive system and the EQ visual analogue scale (EQ VAS).The EQ-5D-3L descriptive system comprises the following five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has 3 levels: no problems, some problems, and extreme problems. The patient is asked to indicate his/her health state by ticking the box next to the most appropriate statement in each of the five dimensions. This decision results into a 1-digit number that expresses the level selected for that dimension. The digits for the five dimensions can be combined into a 5-digit number that describes the patient's health state. The EQ VAS records the patient's self-rated health on a vertical visual analogue scale where the endpoints are labelled 'Best imaginable health state (scale of 100) ' and 'Worst imaginable health state (scale of 0)'. The VAS can be used as a quantitative measure of health outcome that reflects the patient's own judgement.

5.4.3.5.5.2. Analysis Related to Health Status Questionnaire (EQ-5D-3L)

- The change from baseline in EQ VAS scale at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group.
- The change from baseline in EQ VAS scale will be summarized over time by treatment group.
- Number of participants with a reduction of ½ SD or more from baseline in EQ VAS scale at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group.
- The distribution of health state scale for each of the EQ-5D-3L dimension at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group; and will also be summarized over time.

5.4.3.5.6. Patient's Global Impression of Change (PGI-c)

5.4.3.5.6.1. Definitions

The PGI-c is a single item patient reported outcome that assesses change in severity of skin pain due to HS. Participants will rate how his/her HS has changed since the beginning of the study using a 7-point scale ranging from 1 which indicates "very much better" to 7 which indicates "very much worse" with a neutral center point 4 which indicates ("no change").

5.4.3.5.6.2. Analysis Related to PGI-c

- The distribution of the PGI-c scale will be summarized over time by treatment group.

5.4.3.5.7. Patient Global Impression of Severity (PGI-s)

5.4.3.5.7.1. Definition

The PGI-s is a single item patient reported outcome that assesses change in a patient's impression of their disease severity. The PGI-s item asks the respondent to best describe how his/her HS symptoms are now ("check the one number that best describes how your HS symptoms are now") on a 4-point scale scored as: "normal" (1), "mild" (2), "moderate" (3), or "severe" (4)".

5.4.3.5.7.2. Analysis Related to PGI-s

- The distribution of the PGI-s scale at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group.
- The proportion of participants who achieved PGI-s of normal (0), normal (0) or mild (1) at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group respectively.
- Summary of PGI-s symptom status will be summarized over time.

5.4.3.6. Exploratory Endpoints

In addition to the above efficacy analyses, the follow efficacy analyses and summaries will be explored :

5.4.3.6.1. International Hidradenitis Suppurativa Severity Score System (IHS4)

5.4.3.6.1.1. Definition

IHS4 is a dynamic severity assessment of HS. IHS4 score is arrived at by the number of nodules (multiplied by 1) plus the number of abscesses (multiplied by 2) plus the number of draining tunnels (multiplied by 4). A total score of 3 or less signifies mild, 4-10 signifies moderate and 11 or higher signifies severe disease.

5.4.3.6.1.2. Analysis Related to IHS4

- The change from baseline in international hidradenitis suppurativa severity score at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group.
- The change from baseline in international hidradenitis suppurativa severity score will be summarized over time by treatment group.

5.4.3.6.2. Hidradenitis Suppurativa Quality of Life Measure (HiSQOL)

5.4.3.6.2.1. Definition

HiSQOL is a patient self-reported questionnaire that assesses various HS-related quality of life measures. This instrument includes 17 items that are categorized into 3 subscales as 'Symptom', 'Psychosocial', and 'Activities'. A five-point item response scale incorporating 'extremely', 'very much', 'moderately', 'slightly' or 'not at all' was used for all items. The point values assigned to these responses were 4, 3, 2, 1 and 0, respectively. For some items, respondents were given the additional option of 'unable to do, due to my HS' and/or 'I do not normally do this, HS did not influence'. The former option was assigned a score of 4 to indicate the severity of the impact of HS, whereas the latter option was assigned a score of 0 to indicate that HS did not impact it.

5.4.3.6.2.2. Analysis Related to HiSQOL

- The change from baseline in HiSQOL and HiSQOL subscales at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group.
- The change from baseline in HiSQOL and HiSQOL subscales will be summarized over time by treatment group.

5.4.3.6.3. Physician's Global Assessment of Inflammatory Skin Disease (PGA-ISD)

5.4.3.6.3.1. Definition

The PGA-ISD is a tool developed by the International Dermatology Outcome Measures (IDEOM) organization to serve as a potential standardized outcome measure for assessment of inflammatory skin diseases.

- PGA-ISD score of 0: No inflammatory activity
- PGA-ISD score of 1: Barely discemible inflammatory activity
- PGA-ISD score of 2: Mild inflammatory activity
- PGA-ISD score of 3: Moderate inflammatory activity
- PGA-ISD score of 4: Severe inflammatory activity
- PGA-ISD score of 5: Extremely severe inflammatory activity

5.4.3.6.3.2. Analysis Related to PGA-ISD

- The distribution of the PGA-ISD score at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group and will be summarized over time.
- The proportion of participants who achieved PGA-ISD score of no inflammatory activity (0), no inflammatory activity (0) or barely discemible inflammatory activity (1) or PGA-ISD score of mild inflammatory activity or better (≤ 2) at Week 12 and Week 16 will be compared between each of the bermekimab groups and the placebo group respectively.
- The proportion of participants who achieved PGA-ISD score of no inflammatory activity (0), no inflammatory activity (0) or barely discemible inflammatory activity (1) or PGA-ISD score of mild inflammatory activity or better (≤ 2) will be summarized over time by treatment group.

6. SAFETY

Safety will be assessed by summarizing the incidence and type of AEs and examining changes in laboratory parameters (hematology, chemistry, and inflammatory markers), and vital signs.

In all the safety analysis, randomized participants who received at least 1 (partial or complete) dose of study agent administration will be included. For safety analyses, participants will be analyzed according to the randomized treatment groups they were assigned to, regardless of the treatments they actually received. No formal statistical comparison is planned.

Depending on the safety data categories, the cumulative safety data will be analyzed through different study periods through Week 16 and Week 36. Unless otherwise specified, tabular summaries of safety events for key study periods are in general presented as following:

Summaries through Week 16 (placebo controlled):

Safety data through Week 16 will be summarized by treatment groups

1. Placebo
2. Bermekimab 400 mg q2w
3. Bermekimab 400 mg qw
4. Combined Bermekimab

This allows between-group comparisons of safety between the bermekimab regimens and the placebo group based on similar follow-up period in each group.

Summaries through Week 36

Safety data through Week 36 will be summarized by treatment groups including

1. Placebo → bermekimab 400 mg qw
2. Bermekimab 400 mg q2w
3. Bermekimab 400 mg qw
4. Combined Bermekimab

This allows safety comparisons between each bermekimab regimen through Week 36 based on the similar follow-up time in each group.

The list of actual treatment groups for safety analyses and inclusions of participants and safety events/measurements in each group are defined as follows:

1. **Placebo:** all participants who were randomized to placebo at Week 0 and received treatment with placebo only or received treatment with placebo prior to receiving treatment with bermekimab. For participants who started treatment with placebo but later received bermekimab, the safety events/measurements on and after the first dose of bermekimab will be excluded from this group. Only the safety events/measurements that occurred while the participants had been receiving placebo only will be included in this group.

2. **Placebo → bermekimab 400 mg qw:** All participants who were randomized to placebo at Week 0 started treatment with placebo only and later crossed over to receive treatment with bermekimab 400 mg SC (qw). Only the safety events/measurements from these participants that occurred on and after their first administration of bermekimab will be included in this group.

3. **Bermekimab 400 mg q2w:** all participants who were randomized to bermekimab 400 mg SC (q2w) at Week 0 and received bermekimab. All the safety events/measurements from these participants that occurred from Week 0 will be included in this group.

4. **Bermekimab 400 mg qw:** all participants who were randomized to bermekimab 400 mg SC (qw) at Week 0 and received bermekimab. All the safety events/measurements from these participants that occurred from Week 0 will be included in this group.

6.1. Adverse Events

The verbatim terms used in the eCRF by investigators to identify adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). Any AE starting at or after the initial administration of study agent through the end of the trial is considered to be treatment emergent. If the event occurs on the day of the initial administration of study agent, and either event time or time of administration are missing, then the event will be assumed to be treatment emergent. If the event date is recorded as partial or completely missing, then the event will be considered to be treatment emergent unless it is known to be prior to the first administration of study agent based on partial onset date or resolution date. All reported treatment-emergent adverse events will be included in the analysis.

The following analyses will also be used to assess the safety of participants in the study:

- The incidence and type of AEs.
- The incidence and type of SAEs.

- AEs of severe intensity
- AEs that led to permanent discontinuation of study agent administration
- The incidence and type of reasonably related AEs.

In addition, a participant who had injection site reaction(s) in 30 minutes (+/- 10) post injection at Weeks 0, 1, and 2 were recorded in eCRF. The incidence of injection-site reaction during this time frame will be summarized.

These summary tables will provide the count and percentage of participants with 1 or more of the specified AEs by treatment group. In addition to the summary tables, listings will be provided for participants who:

- Had SAEs.
- Had AEs leading to discontinuation of study.
- Had AEs of severe or life threating, or fatal based on CTCAE (grade ≥3) intensity.
- Had anaphylactic or serum sickness-like reactions.

Since safety should be assessed relative to exposure and follow-up, most AE summary tables will include average weeks of follow-up and average number of study agent administrations, infusions and injections for each treatment group. A listing of participants who died will be provided.

6.2. Clinical Laboratory Tests

All clinical laboratory reports will be displayed for the participants included in the safety analysis set. The clinical laboratory parameters to be evaluated by the central laboratory include but are not limited to:

- **Hematology**: Complete Blood Count (CBC) with differential, platelets, and hemoglobin.
- **Chemistry**: albumin, alkaline phosphatase, alanine aminotransferase, aspartate aminotransferase, total bilirubin, blood urea nitrogen/urea, calcium, chloride, creatinine, GGT, glucose, potassium, total protein, sodium, bicarbonate.

Box plots of laboratory measurements and change from baseline will be provided for the selected laboratory measurement.

Applicable laboratory results will be graded according to National Cancer Institute's Common Terminology Criteria for Adverse Events (NCI-CTCAE version 5.0). The worst NCI-CTCAE will be summarized by treatment group.

A listing of participants with 1 or more NCI-CTCAE toxicity grade ≥2 in hematology and clinical chemistry laboratory measurements will be provided

6.3. Vital Signs and Physical Examination Findings

Vital signs including pulse, blood pressure (systolic and diastolic), body temperature, peripheral blood oxygen saturation (SpO2) and respiratory rate will be measured at visits (pre-injection and

post-injection) as per the schedule of activities in the protocol. Descriptive statistics of the observed value and change from baseline of the vital signs will be summarized by treatment group over time.

Physical exam findings will not be analyzed. When physical exam findings are captured as AEs, those will be included in the analyses of AEs.

7. PHARMACOKINETICS/PHARMACODYNAMICS

7.1. Pharmacokinetics

All PK analyses will be based on the PK analysis set (Section 2.4.3). Participants will be analyzed according to the randomized treatment groups they were assigned to, regardless of the treatments they actually received. Unless otherwise specified, tabular summaries of PK for key study periods are in general presented as following:

Summaries through Week 16:

PK data through Week 16 will be summarized by treatment groups

- Bermekimab 400 mg q2w
- Bermekimab 400 mg qw

Summaries through Week 36:

PK data through Week 36 will be summarized by treatment groups including

- Placebo → Bermekimab 400 mg qw
- Bermekimab 400 mg q2w
- Bermekimab 400 mg qw

7.1.1. Bermekimab Concentrations

Blood samples for measuring bermekimab concentrations will be collected from all participants at the specified visits as shown in the schedule of activities in the protocol. Samples will also be collected at the final visit from participants who terminate study participation early. Samples must be collected before study agent administration at visits when a study agent administration is scheduled.

Bermekimab concentrations will be summarized using descriptive statistics, including n, arithmetic mean, SD, coefficient of variation (%CV), median, interquartile range, range (minimum and maximum), by treatment group at each PK sampling time where appropriate. PK data may be displayed graphically. The following analyses will be performed as appropriate:

- Summary of bermekimab concentrations at each visit by treatment group.
- Proportion of participants without detectable bermekimab concentration (below the lower limit of quantification) at each visit by treatment group.

- Summary of bermekimab concentrations at each visit by treatment group and baseline body weight (quartiles). Other covariates may also be applied.
- Summary of bermekimab concentrations at each visit by treatment group and baseline Hurley stage status .
- Summary of bermekimab concentrations at each visit by treatment group and prior biologic use status.
- Plot of mean (SD) bermekimab concentrations over time by treatment group.
- Plot of median (IQ) bermekimab concentrations over time by treatment group.

7.1.1.1. Data Handling Rules

Unless otherwise specified, the following data handling rules will apply to PK analyses:

- Participants will be analyzed according to the treatment groups that they actually received.
- All concentration summaries for a particular timepoint will include data obtained from treated participants at the timepoint of interest without imputing any missing data.
- A concentration not quantifiable (below the lower limit of quantification) will be treated as 0 in the summary statistics and shown as the lower limit of quantification (< LLOQ) in the data listings.
- The data from a participant who meets 1 of the following dosing deviation criteria (if applicable) will be excluded from the by-visit data analyses from that point onwards:
- Discontinue bermekimab administrations.
- Skipped a bermekimab administration.
- Received an incomplete/ incorrect dose (if applicable).
- Received an incorrect study agent (if applicable).
- Received an additional bermekimab dose.

In addition, if a participant has an administration outside of dosing windows (Table 2), the concentration data collected at and after that visit will be excluded from the by-visit data analyses. Additional exclusions for incongruous PK data to be implemented based on Janssen SOP-07948. All participants and samples excluded from the analysis will be clearly documented in the Clinical Study Report.

Table 2: Dosing Window

Visit	Window
Week 0 through Week 32	± 3 days from scheduled visit day
Final Safety and Efficacy Follow-up visit	± 7 days from scheduled visit day

7.1.2. PK vs Efficacy

The relationship between bermekimab concentrations and efficacy endpoints may be explored, e.g.:

- The relationship between bermekimab concentrations (quartiles) and proportion of participants achieving HiSCR and change from baseline in AN count.

7.1.3. Population PK Analysis

When appropriate, population PK analysis may be performed using bermekimab concentration-time data with the nonlinear mixed-effects modeling (NONMEM) approach. Details will be provided in a separate technical report.

7.2. Immunogenicity

Immunogenicity analyses will be based on the Immunogenicity Analysis Set (Section 2.4.5). Participants will be analyzed according to the randomized treatment groups they were assigned to, regardless of the treatments they actually received. No imputation for missing concentration data will be performed.

Immunogenicity data through Week 36 will be summarized by treatment groups including

- Placebo → Bermekimab 400 mg qw
- Bermekimab 400 mg q2w
- Bermekimab 400 mg qw
- Combined Bermekimab

7.2.1. Antibodies to Bermekimab

Blood samples will be collected to examine the formation of antibodies to bermekimab at the specified visits as shown in the schedule of activities in the protocol. Samples will also be collected at the final visit from participants who terminate study participation early.

The antibodies to bermekimab status (positive, negative) and titers will be summarized by treatment group for participants who receive at least one dose of bermekimab and have appropriate samples for detection of antibodies to bermekimab. The maximum titers of antibodies to bermekimab will be provided for participants who are positive for antibodies to bermekimab.

A listing of participants who are positive for antibodies to bermekimab will be provided. This listing will provide information regarding dose administered, bermekimab concentration, and antibody status for all visits.

7.2.2. Antibody vs PK/Efficacy/Safety

To explore the relationship between antibodies to bermekimab status and serum bermekimab concentrations, efficacy and safety, the following analysis may be performed if sufficient numbers of participants are positive for antibodies:

- Summary of bermekimab concentrations by antibodies to bermekimab status.
- Plots of mean (SD) trough bermekimab concentrations over time by antibodies to bermekimab status.
- Summary of clinical response status (e.g., proportion of participants achieving HiSCR and change from baseline in AN count) by antibodies to bermekimab status.
- Summary of injection-site reactions by antibodies to bermekimab status if applicable.
- List of antibodies to bermekimab status in participants who discontinued study early.

7.3. Biomarkers

Biomarker assessments will include the evaluation of relevant markers in serum for all participants. The samples will be used to better understand the biology of HS in some or all of the following ways: to provide mechanistic assessment of the pharmacodynamic response of participants to treatment with bermekimab, to analyze differences between responders and nonresponders, and to determine if the markers might be used to classify participants as potential responders prior to treatment.

Serum samples will be analyzed for circulating factors such as cytokines and other inflammatory markers (eg, IL-17A, IL-17F, IL-23, TNFα, IL-6, IL-22), and other categories of biomarkers potentially associated with the development and progression of HS or related to the bermekimab mechanism of action. These analyses are considered exploratory and will be summarized in a separate biomarker technical report.

REFERENCES

1. Rubin DB, Multiple Imputations for Nonresponse in Surveys. New York: John Wiley & Sons. 1987.