Protocol Number: 22107

Official Title: A Randomized, Double-blind, Placebo-controlled, Multicenter, Phase 3 Study to Evaluate the Efficacy and Safety of Izokibep in Subjects with Moderate to Severe Hidradenitis Suppurativa

NCT Number: NCT05905783

Document Date: 10 December 2024

CONFIDENTIAL

ACELYRIN, INC.

Protocol Number: 22107

A Randomized, Double-blind, Placebo-controlled, Multicenter, Phase 3 Study to Evaluate the Efficacy and Safety of Izokibep in Subjects with Moderate to Severe Hidradenitis Suppurativa

Statistical Analysis Plan (SAP) for Final Analysis

Final Version 1.0

10 Dec 2024

SAP APPROVAL SIGNATURE PAGE

Protocol Number:	22107
Protocol Version and Date:	Version 2.0 – 10 May 2024
Protocol Title:	A Randomized, Double-blind, Placebo-controlled, Multicenter, Phase 3 Study to Evaluate the Efficacy and Safety of Izokibep in Subjects with Moderate to Severe Hidradenitis Suppurativa
SAP Version and Date:	Version 1.0 – 10 Dec 2024

☒ Original Statistical Analysis Plan ☐ Amended Statistical Analysis Plan

Signature below indicates the SAP has been authored by the following:

[redacted] Date

[redacted]
ACELYRIN, INC.

Signatures below indicates the SAP has been reviewed and approved by the following:

[redacted] Date

[redacted] ACELYRIN, INC.

VERSION HISTORY OF FULLY EXECUTED (FINAL) PLANS

Version Number	Date	Version Author	Comments
Final Analysis SAP 1.0	10DEC2024	[redacted]	Initial version of Final Analysis SAP. This document is based on the Statistical Analysis Plan for Primary Analysis Version 2.0 (16JUL2024). Analyses of Period 2 (and combined Period 1 + Period 2) data have been added to the descriptions of the completed primary analyses. Some of the originally planned analyses for Period 1 have been removed. This document incorporates the decisions to terminate this study early and write an abbreviated CSR.

TABLE OF CONTENTS

TABLES

1. LIST OF ABBREVIATIONS

Abbreviation or Term	Definition
ADA	anti-drug antibody
AESI	adverse event of special interest
ALP	Alkaline Phosphatase
ALT	Alanine Transaminase
AN	abscess and inflammatory nodule
AR	autoregressive
AST	Aspartate Transferase
ATC	Anatomical Therapeutic Chemical
BLQ	below limit of quantification
BMI	body mass index
bpm	beats per minute
C	Celsius
CI	confidence interval
cm	centimeters
CRO	clinical research organization
CSR	clinical study report
CS	compound-symmetric
C-SSRS	Columbia-suicide severity rating scale
CV	coefficient of variation
DBP	diastolic blood pressure
DILI	drug-induced liver injury
DLQI	Dermatology Life Quality Index
ECG	electrocardiogram
EQ-5D	European Quality of Life-5 Dimensions
FAS	full analysis set
FASP2	full analysis set period 2
HADS	Hospital Anxiety and Depression Scale
HiSCR	Hidradenitis suppurative clinical response
HiSCR100	Hidradenitis Suppurativa Clinical Response 100
HiSCR50	Hidradenitis Suppurativa Clinical Response 50
HiSCR75	Hidradenitis Suppurativa Clinical Response 75
HiSCR90	Hidradenitis Suppurativa Clinical Response 90
HiSQOL	Hidradenitis Suppurativa Quality of Life
HS-PGA	Hidradenitis Suppurativa-Physician's Global Assessment
ICF	informed consent form
IDMC	independent data monitoring committee
IHS4	International Hidradenitis Suppurativa Severity Score System
IL	interleukin
INR	international normalized ratio
ISR	injection site reaction
IXRS	Interactive Voice/Web Response System
JAK	Janus Kinase
kg	kilograms
LFT	liver function test
LS	least square
MACE	Major adverse cardiovascular and cerebrovascular event
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities

Abbreviation or Term	Definition
mmHg	millimeters of mercury
MMRM	mixed model repeated measures
NRI	non-response imputation
NRS	numeric rating scale
PGIC	Patient Global Impression of Change
PK	pharmacokinetic
PRO	patient reported outcome
PT	preferred term
QW	every week
SAE	serious adverse event
SAF	safety analysis set
SAFP2	safety analysis set period 2
SAP	statistical analysis plan
SBP	systolic blood pressure
SD	standard deviation
SF-12	Short Form-12v2™ Health Survey
SoA	schedule of activities
SOC	system organ class
TEAE	treatment-emergent adverse event
TNFi	tumor necrosis factor-alpha inhibitor
ULN	upper limit of normal
UN	unstructured
WHODD	World Health Organization Drug Dictionary

2. INTRODUCTION

This study will be conducted to evaluate the efficacy and safety of izokibep in the treatment of subjects with moderate to severe Hidradenitis Suppurativa (HS). Izokibep is a small protein molecule that acts as a selective, potent inhibitor of Interleukin (IL) 17A, to which it binds with high affinity.

This study investigates izokibep in subjects with active HS, including tumor necrosis factor-alpha inhibitor (TNFi) naïve subjects, and those who had an inadequate response or intolerance to TNFi, or for whom TNFi is contraindicated. The study is intended to be 1 of 2 adequate and well-controlled studies to support a claim of efficacy of izokibep in subjects with HS. Hidradenitis suppurativa clinical response 75 (HiSCR75) at Week 12 will be utilized to compare izokibep to placebo, for the primary endpoint.

This statistical analysis plan (SAP) will define the final analyses planned for the study.

The study will be terminated early after all subjects completed (or discontinued) treatment at Week 32. A decision was made to write an abbreviated clinical study report (CSR). The analyses described thereafter reflect this decision.

3. STUDY OBJECTIVES

3.1 Primary Objective

The primary objective of this study is:

- To demonstrate the efficacy of izokibep compared with placebo, as measured by percentage of subjects achieving HiSCR75 at Week 12

3.2 Secondary Objectives

Secondary objectives of this study are:

- To demonstrate that izokibep is efficacious compared with placebo as measured by:
  - Percentage of subjects achieving HiSCR90 at Week 12
  - Percentage of subjects achieving HiSCR100 at Week 12
  - Percentage of subjects achieving HiSCR50 at Week 12
  - Percentage of subjects who experience ≥ 1 disease flare through 12 weeks of treatment
  - Change in Dermatology Life Quality Index (DLQI)
  - Percentage of subjects with baseline Hurley Stage II who achieve abscess and inflammatory nodule (AN) count of 0, 1, or 2 at Week 12
  - Percentage of subjects achieving at least 3-point reduction from baseline in Numeric Rating Scale (NRS) Patient Global Assessment of Skin Pain at its worst among subjects with baseline NRS ≥ 4
- To assess the safety and tolerability of izokibep as measured by the incidence of treatment-emergent adverse events (TEAEs), events of interest, serious adverse events (SAEs), and clinically significant laboratory values and vital signs

3.3 Exploratory Objectives

Exploratory objectives of this study are:

- To demonstrate that izokibep is efficacious compared with placebo at timepoints through Week 16 and to estimate izokibep efficacy at timepoints after Week 16, as measured by:
  - Percentage of subjects who achieve HiSCR75 at Weeks 2, 4, 8, 16, 32, and 52
  - Percentage of subjects who achieve HiSCR90 at Weeks 2, 4, 8, 16, 32, and 52
  - Percentage of subjects who achieve HiSCR100 at Weeks 2, 4, 8, 16, 32, and 52
  - Percentage of subjects who achieve HiSCR50 at Weeks 2, 4, 8, 16, 32, and 52
  - Percentage of subjects with baseline Hurley Stage II who achieve AN count of 0, 1, or 2 at Weeks 4, 8, 16, 24, 32, and 52
  - Percentage of subjects achieving at least 3-point reduction from baseline in NRS Patient Global Assessment of Skin Pain at its worst among subjects with baseline NRS ≥ 4
  - Mean change from baseline in NRS Patient Global Assessment of Skin Pain at its worst (all subjects)
  - Mean change from baseline in NRS Patient Global Assessment of Skin Pain on average (all subjects)
  - Modified Sartorius Score after 4, 8, 12, 16, 24, 32, and 52 weeks of treatment
  - HS flare, rates through 4 and 52 weeks of treatment
  - International Hidradenitis Suppurativa Severity Score System (IHS4) after 4, 8, 12, 16, 24, 32, and 52 weeks of treatment
  - Hidradenitis Suppurativa-Physician's Global Assessment (HS-PGA)
  - Mean change from baseline in draining tunnel count
- To explore the effect of izokibep on Patient Reported Outcomes (PROs) as measured by change from baseline over time in:
  - European Quality of Life-5 Dimensions (EQ-5D)
  - Hidradenitis Suppurativa Quality of Life (HiSQOL)

  - Hospital Anxiety and Depression Scale (HADS)
  - Patient Global Impression of Change (PGIC)
  - DLQI
  - Short Form-12v2™ Health Survey (SF 12v2)
- To estimate the mean trough plasma level izokibep in subjects with HS
- To assess the immunogenicity of izokibep as measured by the presence of anti-drug-antibodies (ADAs)
- To assess biomarkers of anti-IL-17A treatment and biomarkers associated with HS efficacy (United States Only)

4. STUDY DESIGN

4.1 General Design

This is a Phase 3, pivotal, confirmatory, 1:1 randomized, double-blind, placebo-controlled, parallel group study to evaluate the efficacy and safety of izokibep in subjects with moderate to severe HS. Subjects must have HS lesions present in ≥ 2 distinct anatomic areas, 1 of which must be Hurley Stage II or Hurley Stage III, and a total AN count (total sum of abscesses and inflammatory nodules) of ≥ 5. Subjects with draining fistula count of > 20 at screening or Day 1 will be excluded.

Overall, 250 subjects (125 subjects per treatment group) with moderate to severe HS are planned to be enrolled at sites globally in North America, Europe, and Asia Pacific. Additional sites and regions may be added.

Subjects meeting eligibility criteria will be randomized into 1 of 2 treatment groups in a 1:1 ratio as follows:

- Group 1 (n = 125): placebo every week (QW) from Day 1 through Week 15, then izokibep 160 mg QW from Weeks 16 to 51.
- Group 2 (n = 125): izokibep 160 mg QW from Day 1 through Week 51.

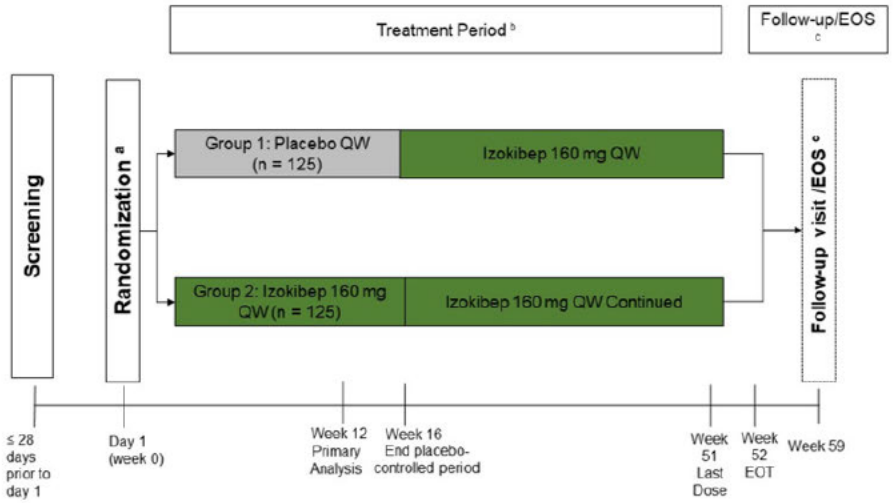
The first dose of study drug (i.e., izokibep or placebo) will be administered on Day 1.

Randomization will be stratified by prior TNFi use for HS (Yes/No) and Hurley Stage (II or III). Subjects enrolled with stable concomitant antibiotic use (dosing regimen has been stable for ≥ 4 weeks prior to first dose of study drug) will be capped at approximately 30%.

The maximum planned length of participation in the study for an individual subject is up to 63 weeks, which includes the following:

- Screening Period of up to 28 days (4 weeks)
- Treatment Period of 51 weeks
- Follow-up visit at 8 weeks after the last dose of study drug.

Figure 1: Study Design Schema



EOS = end of study; EOT = end of treatment; QW = every week

[a] Randomization will be stratified by any prior TNFi use for hidradenitis suppurativa (Yes/No) and Hurley Stage (II or III).

[b] Subjects in Group 1 (placebo QW) will receive placebo QW until Week 15, then izokibep QW beginning at Week 16 until Week 51. Subjects in Group 2 (izokibep QW) will receive izokibep QW.

[c] An 'end of treatment visit' will be conducted at Week 52 (± 5 days). A 'safety follow-up visit' will be conducted at Week 59 (± 5 days). For subjects who early terminate, the end of treatment visit should be completed 1 week after the last dose of study drug and safety follow-up visit completed 8 weeks after the last dose of study drug (± 5 days), where possible.

The primary analysis of primary and secondary endpoints will be conducted after all subjects have the opportunity to complete Week 12 assessments or early terminate from the study (all subjects complete week 12, or complete an assessment after week 12, or permanently discontinue from the study).

A separate, final, analysis of data will occur after all subjects have completed or discontinued from the study. This document defines the final analysis.

The study will be terminated early after all subjects completed (or discontinued) treatment at Week 32. A decision was made to write an abbreviated clinical study report (CSR). The analyses described thereafter reflect this decision.

4.2 Randomization and Treatment Assignments

Subjects will be randomized to the study drug or placebo (in a 1:1 ratio) on Day 1 by the Interactive Voice/Web Response System (IXRS).

Randomization will be stratified by prior TNFi use for HS (Yes/No) and Hurley Stage (II or III). Subjects on a stable dose of antibiotic will be capped at a rate of approximately 30%. Izokibep and matching placebo will be visually indistinguishable to prevent unblinding during preparation or administration of study treatment.

4.3 Blinding/Unblinding

The subject, site personnel, and sponsor/Contract Research Organization (CRO) study personnel and designees will be blinded to the randomization treatment group assignment (ie, izokibep or placebo). A subject's treatment assignment should be unblinded only when knowledge of the treatment is essential for the further management of the subject in this study. Unblinding at the site for any other reason will be considered a protocol deviation. The investigator should contact the medical monitor before unblinding any subject's treatment assignment, whenever possible. If an urgent therapeutic intervention

is necessary that warrants unblinding prior to contacting the medical monitor, the investigator can directly access the IXRS, or call the 24-hour emergency medical coverage line to unblind without medical monitor notification or agreement but must contact the medical monitor within 1 working day after the unblinding event.

If an SAE requires an expedited regulatory report to be sent to 1 or more regulatory agencies, sponsor/designee's safety staff may unblind the treatment assignment for the subject. A copy of the report, identifying the subject's treatment assignment, may be sent to that regulatory agency in accordance with local regulations.

4.4 Sample Size Considerations

In prior studies, HiSCR75 response rates among placebo recipients after 12 to 16 weeks of treatment have been approximately 10% to 20%. At Week 12 in this study, HiSCR75 response rates are expected to be 15% for placebo recipients and 35% for izokibep recipients. With 125 subjects receiving each treatment, this study will have approximately 95% power to demonstrate a difference in HiSCR75 response rates between arms. This calculation uses an unstratified test, with $\alpha = 0.05$, 2-sided.

5. CHANGES IN THE ANALYSES SPECIFIED IN THE PROTOCOL

The secondary endpoint, "Percentage of subjects who experience ≥ 1 disease flare through 12 weeks of treatment" has been moved down the hierarchy of secondary endpoint testing to be the (last) seventh one in the order of testing.

The exploratory endpoint "Change of Modified Sartorius Score from baseline, Weeks 4, 8, 12, and 16" will not be analyzed due to inconsistent and unreliable data collection.

Due to the early termination of the study and the decision to write an abbreviated clinical study report, exploratory efficacy analyses will not be performed as written in the protocol. The efficacy endpoints for which analyses will be done as-written will be limited to the following:

- Primary:
  - HiSCR75 at Week 16
- Secondary/Exploratory:
  - HiSCR90 at Weeks 2, 4, 8, 16, and 32
  - HiSCR100 at Weeks 2, 4, 8, 16, and 32
  - HiSCR50 at Weeks 2, 4, 8, 16, and 32
  - AN count of 0, 1, or 2 at Weeks 4, 8, 16, 24, and 32
  - Change in NRS Patient Global Assessment of Skin Pain at its worst in subjects with baseline NRS ≥ 4 at Weeks 4, 8, 16, 24, and 32
  - HS flares through Weeks 4, 16 and 52
  - Change from baseline to Weeks 4, 8, 16, and 24 in DLQI

All remaining exploratory efficacy endpoint analyses will not be performed.

Limited safety summaries are planned for the purposes of abbreviated CSR. Clinical laboratory data, ECGs, vital signs, and physical exam results will not be summarized. These data will be provided as datasets.

6. STUDY ENDPOINTS

6.1 Schedule of Assessments

The assessments planned for each visit are presented in the Protocol in Table 1, Schedule of Activities (SoA).

6.2 Timepoint Definitions

6.2.1 Relative Day

All assessment days will be relative to the date of randomization. The date on which the subject is randomized will be considered as relative Day 1, while the day before that will be relative Day -1. There is no relative Day 0. Relative days will be calculated using the following formula only if the date of the assessment is fully known (i.e., the relative day will be missing if the assessment date is a partial date):

Relative Day =

- Date of Assessment – Date of Randomization+1 when Date of Assessment ≥ Date of Randomization.
- Date of Assessment – Date of Randomization when Date of Assessment < Date of Randomization.

The date of randomization for each subject will be taken from the Disposition eCRF page. The date of the last dose of study treatment for each subject will be taken from the Study Drug Administration form, with the last study drug administration date as the latest date of study drug regardless of the visit.

6.2.2 Visit and Timepoint Windows

In general, data will be analyzed using the nominal study visits as defined in the SoA in the protocol and reported on the eCRF.

If a subject discontinues the study early, the subject will have an early termination visit which includes efficacy and safety assessments. If the early termination visit occurs in a timeframe that corresponds to a scheduled efficacy or safety assessment (week 2, 4, 8, 12, 16, 20, 24, 32, or 52), and if the subject does not have an actual visit at or after the corresponding scheduled efficacy or safety assessment, the day ranges below will be used to map the early termination visit to a scheduled efficacy assessment for use in all efficacy summaries at that timepoint.

In addition, unscheduled visit results will not be included in analyses, unless the unscheduled visit corresponds to a nominal visit for which the result is missing or if the unscheduled visit is selected as baseline. If the nominal visit result is missing, then the unscheduled visit result will also be mapped according to the ranges in Table 1and Table 2 below for efficacy and safety assessments, respectively. If multiple unscheduled visits fall into 1 analysis window, then the one closest to the scheduled day will be used for analysis. If the 2 unscheduled visits are equidistant from the scheduled day, the later assessment will be used.

Table 1: Visit Windowing for Efficacy Assessments

Nominal Visit	Scheduled Day	Analysis Window (Relative Days)
Day 1/Week 0 (Baseline)	1	[1, 1]
Week 2	14	[2, 21]
Week 4	28	[22, 42]
Week 8	56	[43, 70]
Week 12	84	[71, 98]
Week 16	112	[99, 140]
Week 24	168	[141, 196]
Week 32	224	[197, 294]
Week 52	364	[295, later]

Table 2: Visit Windowing for Safety Assessments:

Nominal Visit	Scheduled Day	Analysis Window (Relative Days)
Day 1/Week 0 (Baseline)	1	[1, 1]
Week 2	14	[2, 21]
Week 4	28	[22, 42]
Week 8	56	[43, 70]
Week 12	84	[71, 98]
Week 16	112	[99, 126]
Week 20	140	[127, 154]
Week 24	168	[155, 196]
Week 32	224	[197, 294]
Week 52	364	[295, later]

6.2.3 Screening Period

The Screening Period starts when the subject signs and dates the informed consent form (ICF) and ends when the subject is enrolled/randomized, or screen failed. The Screening Period is up to 28 days. Certain initial Screening Period procedures may be repeated during the original initial Screening Period. Repeating procedures during the original initial Screening Period is a part of screening and is not considered "rescreening." These procedures include laboratory assessments due to value(s) out of range due to a potential sampling error or that could be within range with repeat sampling.

6.2.4 Baseline

Baseline will be identified as the last non-missing assessment result taken prior to the date/time of the randomization/enrollment.

6.2.5 Treatment Period

Data collected on Day 1 will be assigned to the Treatment Period unless the time (HH:MM) of data collection and time (HH:MM) of randomization are both recorded, and the data collection time is before the time of randomization. In this case, the assessment will be assigned to the Screening Period.

If the time (HH:MM) of data collection is not recorded but the protocol and/or CRF includes an instruction to the effect that all Day 1 assessments are to be performed prior to randomization, the data collected at Day 1 will be assigned to the Screening Period. However, adverse events and concomitant

medications starting on Day 1, will be assigned to the Treatment Period, unless the time (HH:MM) demonstrates otherwise. Adverse events occurring within 4 weeks of the last dose of study drug will also be assigned to the Treatment Period. Adverse events occurring > 4 weeks after the last dose of study drug will be assigned to the "Safety Follow-up Period".

The Treatment Period will be defined as the time from the date/time of randomization to 1 week after the last dose or when the subject completes their End of Treatment visit. Adverse events that occur within 4 weeks of the last dose of study drug will be attributed to the Treatment Period as well.

Treatment Period will also be divided into 2 periods, Period 1 and Period 2:

- Period 1 is the placebo-controlled period, and will start at the date of randomization and end at Week 16 when Week 16 assessments are completed and before administration of study drug in Period 2 (at or after Week 16 dose date/time). If a subject discontinues study drug prior to Week 16 and continues to attend visits without treatment through Week 16 (e.g. eCRF visit or analysis visit Week 8, Week 12, Week 16), then these visits' results will be considered in Period 1, and should be included in efficacy/safety summaries.
- Period 2 will start at the time of first administration of study drug at or after Week 16 and end at Week 56 or early termination/end of treatment visit. If a subject discontinued study drug prior to Week 16 and continues to attend visits without treatment after Week 16 (e.g. eCRF visit or analysis visit Week 20, 24, etc.), then these visits' results will be considered in Period 2 and should be included in efficacy/safety summaries.

The Safety Follow-up Period will start after completion of the End of Treatment visit (either Week 52 or EOT for subjects who discontinue treatment) or 8 days after the last dose of study drug (for subjects who don't complete the End of Treatment visit). Adverse events that occur within 4 weeks of the last dose of treatment will not be assigned to the Safety Follow-up Period (rather, they will be assigned to the Treatment Period, either Period 1 or Period 2), but any adverse events after that time will be assigned to the Safety Follow-up Period. All other data collected in the defined Safety Follow-up Period will be assigned to it as stated.

6.3 Disposition of Subjects

Disposition categories are defined in Table 3, below:

Table 3: Disposition Categories

Category	Definition
Study Disposition	
Completed Week 12 Assessment	Subject has a non-missing efficacy result at their Week 12 visit.
Completed Study at Week 12	Subject has a non-missing Week 12 analysis visit (includes windowed early termination/unscheduled visit results) or has any visit after Week 12.
Discontinued Study Before Week 12	Subject has Completed = "No" or status = "Discontinued" on the End of Study Form and missing Week 12 analysis visit (includes windowed early termination/unscheduled visit

	results) and has no additional results at visits after Week 12
Completed Week 16 Assessment	Subject has a non-missing efficacy result at their Week 16 visit.
Completed Study at Week 16 (Period 1)	Subject has a non-missing Week 16 analysis visit (includes windowed early termination/unscheduled visit results) or has any visit after Week 16.
Discontinued Study before Week 16 (within Period 1)	Subject has Completed = "No" on the End of Study Form and missing Week 16 analysis visit (includes windowed early termination/unscheduled visit results) and has no additional results at visits after Week 16.
Discontinued Study before Week 52 (within Period 2)	Subject has Completed = "No" on the End of Study Form and has at least 1 visit in Period 2 completed (includes windowed early termination/unscheduled visit results)
Completed Study at Week 52	Subject has a non-missing Week 52 analysis visit (includes eCRF visits and windowed early termination/unscheduled visit results) or has any visit after Week 52.
Discontinued Study any time before Week 52	Subject has Completed = "No" on the End of Study Form and missing Week 52 analysis visit (includes windowed early termination/unscheduled visit results) and has no additional results at visits after Week 16.
Completed Study	Subject has Completed = "Yes" on the End of Study Form
Discontinued Study	Subject has Completed = "No" on the End of Study Form
Treatment Disposition	
Completed Treatment Through Week 11	At least one dose of study drug has been administered on or after Week 11
Discontinued Treatment Prior to Week 11	Completed = "No" on the End of Treatment form and no dose administered on or after Week 11.
Completed Treatment Through Week 15 (Period 1)	A dose of study drug has been administered on or after Week 15
Discontinued Treatment before Week 15 (within Period 1)	Completed = "No" on the End of Treatment form and no dose administered on or after Week 15.
Discontinued Treatment before Week 51 (within Period 2)	Completed = "No" on the End of Treatment form and at least one dose of study drug administered in Period 2
Completed Treatment through Week 51	Completed = "Yes" on the End of Treatment Form
Discontinued Treatment any time before Week 51	Completed = "No" on the End of Treatment form

6.4 Protocol Deviations

Protocol deviations will be determined outside of the eCRF. The protocol deviations review will include a determination of which deviations are identified as important. Representatives from ACELYRIN, INC. will participate in the review of the deviations.

6.5 Demographics and Baseline Characteristics

The following assessments will be collected at baseline:

- Demographic parameters
  - Age (years) at Screening
  - Sex
  - Race
  - Ethnicity
- Vital Signs
  - Height (centimeters [cm])
  - Weight (kilograms [kg])
  - Body Mass Index (BMI; kg/m^2) – calculated using the following formula: *Weight (kg) ÷ (Height (cm)/100)^2*
  - Body Temperature (Celsius [C]) – temperatures collected in Fahrenheit will be converted to C using the following formula: *(temperature in Fahrenheit – 32) x 5/9*
  - Respiratory Rate (breaths per minute)
  - Systolic Blood Pressure (SBP [mmHg])
  - Diastolic Blood Pressure (DBP [mmHg])
  - Heart Rate (beats per minute [bpm])
- Disease Characteristics
  - Time since diagnosis with HS (years) – calculated using the following formula: *(date ICF signed – date of diagnosis + 1)/365.25*
    - *If the date of diagnosis is incomplete, a date based on the same logic as used for concomitant medications and adverse events (described in Section 8.3) will be imputed*
  - AN Count
  - Abscess Count
  - Draining Fistula Count
  - Non-Draining Fistula Count
  - Total Fistula Count – calculated using the following formula:
    - *draining count + non-draining count*
  - Inflammatory Nodule Count
  - Non-inflammatory Nodule Count
  - Hurley Stage (II, III)
  - Location of lesions
- Other Baseline Characteristics
  - HADS Total Score
  - HADS-Anxiety
  - HADS-Depression
  - PGA Assessment

  - DLQI Score
  - EQ-5D (Domain Scores [Mobility, Self-Care, Pain/Discomfort, Usual Activities, and Anxiety/Depression], Index Score, Visual Analog Scale (VAS) Score)
  - SF -12 (Mental Component Score and Physical Component Score)
  - Columbia-Suicide Severity Rating Scale (C-SSRS)
  - NRS Pain Score (At its Worst and On Average)
  - IHS4
  - HiSQOL
  - Childbearing potential (Yes, No)
  - Alcohol Use (Never, Current, Former)
  - Nicotine Use (Never, Current, Former)
- Medical History
  - Medical history is defined as any condition, apart from the study indication, that the subject may have prior to enrollment in the study. Medical occurrences that begin before the start of study drug administration, but after obtaining ICF will be recorded as medical history/current medical conditions, not as AEs. The only exception to this rule is for serious adverse events – these are collected as such starting from the date of obtaining ICF.
  - Medical history will be coded using the latest Medical Dictionary for Regulatory Activities (MedDRA) version around the time of Database Lock.

6.6 Prior and Concomitant Medications

The latest World Health Organization Drug Dictionary (WHODD) version around the time of Database Lock will be used to code the recorded medications to Anatomical Therapeutic Chemical (ATC) level and standardized (generic) medication name.

Prior medications are those taken with a stop date prior to the start of the Treatment Period. Concomitant medications are those with a start date and time on or before the end of the Treatment Period and a stop date on or after the start of the Treatment Period, including prior medications that are ongoing at the start of the Treatment Period. Analogous logic is used to assign concomitant medication to Period 1 and Period 2. Post-treatment medications are those with a start date after the end of the Treatment Period.

See Section 8.3 below for details on how to identify prior vs. concomitant when the start/end dates for the medication usage are partial dates or missing.

6.7 Primary Efficacy Endpoint

The primary endpoint of the study is HiSCR75, the proportion of subjects achieving at least a 75% reduction from baseline in the total AN count, with no increase from baseline in abscess or draining fistula count, at Week 12.

The proportion of subjects achieving HiSCR75 will be defined as meeting all 3 criteria below at a specific assessment:

- [(AN count at baseline – AN count at current visit) / AN count at baseline] × 100% ≥ 75%
- Abscess count at baseline ≥ abscess count at the current visit

- Draining fistula count at baseline ≥ draining fistula count at the current visit

where AN count is abscess count + inflammatory nodule count, and abscess count, inflammatory nodule count and fistula count are summed across all regions and are inclusive of all abscesses, inflammatory nodules and fistula that have previously undergone localized surgical or medical intervention and are no longer present.

6.7.1 Estimand for Primary Endpoint

The primary estimand is defined through the 5 attributes described in the following sections.

6.7.1.1 Treatment Condition of Interest

The primary treatment condition of interest is izokibep versus placebo.

6.7.1.2 Population of Interest

The population of subjects is those who are 18 years and older with moderate to severe HS (HS lesions present in ≥ 2 distinct anatomic areas, 1 of which must be Hurley Stage II or Hurley Stage III, and a total AN count of ≥ 5) who meet all inclusion criteria and exclusion criteria.

6.7.1.3 Variable of Interest

The endpoint for each subject to address the clinical question is achievement of HiSCR75 after 12 weeks of treatment.

6.7.1.4 Intercurrent Event (ICE)

The treatment policy strategy approach will be used in general. Thus, all subjects will be included using observed data at Week 12 regardless of treatment discontinuation, concomitant medication/rescue medication, protocol deviations, or other actions.

Subjects who receive antibiotic therapy that could affect HS will be included in the primary analysis by assigning non-response at subsequent timepoints. A list of all oral antibiotic therapy that results in such imputation will be finalized after review of all oral antibiotic therapy received by any subject and before unblinding of the study, and will include tetracycline, clindamycin, and possibly other products. Similarly, subjects who discontinue treatment due to an AE or lack of efficacy, and subsequently have missing efficacy responses, will be included by imputing non-response.

Multiple Imputation will be used to handle other types of missing data as described in Section 9.5.

6.7.1.5 Population Level Summary

The population level summary is the response rate, or proportion of subjects who meet HiSCR75 after 12 weeks of treatment.

6.8 Secondary Efficacy Endpoints

The secondary efficacy endpoints are defined in Table 4:

Table 4: Secondary Efficacy Endpoints

Endpoint	Derivation
HiSCR90 at Week 12	At least 90% reduction from baseline in the total AN count, with no increase from baseline in abscess or draining fistula count at Week 12. Must meet all 3

	criteria as described for the primary endpoint replacing 75% with 90%
HiSCR100 at Week 12	At least 100% reduction from baseline in the total AN count, with no increase from baseline in abscess or draining fistula count at Week 12. Must meet all 3 criteria as described for the primary endpoint replacing 75% with 100%
HiSCR50 at Week 12	At least 50% reduction from baseline in the total AN count, with no increase from baseline in abscess or draining fistula count at Week 12. Must meet all 3 criteria as described for the primary endpoint replacing 75% with 50%
HS Flares Through Week 12	≥ 25% increase in AN count with a minimum increase of 2 AN relative to baseline, i.e. subject must meet all the following criteria at Week 12: • (AN count at current visit- AN count at baseline) / AN count at baseline ×100% ≥ 25% • AN count at current visit- AN count at baseline ≥ 2 A subject will be counted as having HS flare if they have a HS flare at any assessment after baseline, up to and including the assessment at Week 12 (including scheduled and unscheduled visits and early termination visits)
Change in DLQI from baseline to Week 12	(DLQI score at current visit – DLQI score at baseline) at each post baseline visit until Week 12 DLQI comprises 10 items arranged in 6 categories: symptoms and feelings (questions 1 and 2), daily activity (3 and 4), leisure (5 and 6), work or study (7), interpersonal relationships (8 and 9), and treatment (10). The total score can vary from 0 (no impact) to 30 (maximum impact).
AN count of 0, 1, or 2 at Week 12	Observed values of 0,1, or 2 for AN count (abscess count + inflammatory nodule count)
Change in pain score from baseline in NRS Patient Global Assessment of Skin Pain at its worst at Week 12	NRS in Patient Global Assessment of Skin Pain at its worst is the answer from Question on the eCRF: In the last 24 hours, which number best describes your skin pain at its WORST due to your HS? The answer is in the range of 0 (No skin pain) to 10 (Skin pain bad as you can imagine). The skin pain score at each visit will be calculated using average of daily scores among the 7 days up to and including the day of visit, with a minimum of 4 days (consecutive or non-consecutive) with scores required. The skin pain at baseline will be calculated using average of daily scores among the 7 days up to and including the first study drug administration day, with a minimum of 4 days with

	scores required. For any visit, if Patient Global Assessment of Skin Pain at its worst is available for 3 or few of the 7 days up to and including the day of visit, the pain score will be missing for analysis.

6.8.1 Estimand for Secondary Efficacy Endpoints

Estimands for secondary efficacy endpoints are defined in Table 5, below:

Table 5: Estimands for Secondary Efficacy Endpoints

Endpoint	Estimand Attribute	Explanation
HiSCR90 at Week 12	Treatment Condition of Interest	izokibep versus placebo
	Population of Interest	Subjects with HS who meet all inclusion/exclusion criteria
	Variable of Interest	Percentage of subjects achieving HiSCR90 at Week 12
	Intercurrent Event	Treatment policy strategy, in general, with all subjects included in the analysis regardless of treatment discontinuation, concomitant medication/rescue medication, protocol deviations, or other actions. Subjects who receive antibiotic therapy that could affect HS will be included in the primary analysis by assigning non-response at subsequent timepoints. Multiple imputation and non-response imputation will be used to handle certain other types of missing data, analogously to the primary endpoint.
	Population Level Summary	Response Rate of HiSCR90 at Week 12 (Yes/No)
HiSCR100 at Week 12	Treatment Condition of Interest	izokibep versus placebo
	Population of Interest	Subjects with HS who meet all inclusion/exclusion criteria
	Variable of Interest	Percentage of subjects achieving HiSCR100 at Week 12
	Intercurrent Event	Treatment policy strategy, in general, with all subjects included in the analysis regardless of treatment discontinuation, concomitant medication/rescue medication, protocol deviations, or other actions. Subjects who receive antibiotic therapy that could affect

		HS will be included in the primary analysis by assigning non-response at subsequent timepoints. Multiple imputation and non-response imputation will be used to handle certain other types of missing data, analogously to the primary endpoint.
	Population Level Summary	Response Rate of HiSCR100 at Week 12 (Yes/No)
HiSCR50 at Week 12	Treatment Condition of Interest	izokibep versus placebo
	Population of Interest	Subjects with HS who meet all inclusion/exclusion criteria
	Variable of Interest	Percentage of subjects achieving HiSCR50 at Week 12
	Intercurrent Event	Treatment policy strategy, in general, with all subjects included in the analysis regardless of treatment discontinuation, concomitant medication/rescue medication, protocol deviations, or other actions. Subjects who receive antibiotic therapy that could affect HS will be included in the primary analysis by assigning non-response at subsequent timepoints. Multiple imputation and non-response imputation will be used to handle certain other types of missing data, analogously to the primary endpoint.
	Population Level Summary	Response Rate of HiSCR50 at Week 12 (Yes/No)
HS Flares Through Week 12	Treatment Condition of Interest	izokibep versus placebo
	Population of Interest	Subjects with HS who meet all inclusion/exclusion criteria
	Variable of Interest	Change in HS flares after 12 weeks of treatment
	Intercurrent Event	Treatment policy strategy, with all subjects included in the analysis regardless of treatment discontinuation, concomitant medication/rescue medication, protocol deviations, or other actions. All subjects with missing abscess, inflammatory nodule, and draining tunnel counts will be

		included using multiple imputation, analogously to the analysis of the primary endpoint. Multiple Imputation described in Section 9.6.1. Subjects who receive antibiotic therapy that could affect HS will be included in the primary analysis by assigning non-response at subsequent timepoints. Multiple imputation and non-response imputation will be used to handle certain other types of missing data, analogously to the primary endpoint.
	Population Level Summary	Improvement in HS flare occurrence (Yes/No)
Change in DLQI from baseline to Week 12	Treatment Condition of Interest	izokibep versus placebo
	Population of Interest	Subjects with HS who meet all inclusion/exclusion criteria
	Variable of Interest	DLQI after 12 weeks of treatment
	Intercurrent Event	Treatment policy strategy, in general, with all subjects included in the analysis regardless of treatment discontinuation, concomitant medication/rescue medication, protocol deviations, or other actions
	Population Level Summary	Mean change from baseline in DLQI
AN count of 0, 1, or 2 at Week 12	Treatment Condition of Interest	izokibep versus placebo
	Population of Interest	Subjects with HS who meet all inclusion/exclusion criteria and with baseline Hurley Stage II
	Variable of Interest	Percentage of subjects achieving AN count of 0, 1, or 2 at Week 12
	Intercurrent Event	Treatment policy strategy, in general, with all subjects included in the analysis regardless of treatment discontinuation, concomitant medication/rescue medication, protocol deviations, or other actions. Subjects who receive antibiotic therapy that could affect HS will be included in the primary analysis by assigning non-response at subsequent timepoints. Multiple imputation and non-response imputation will be used to handle

		certain other types of missing data, analogously to the primary endpoint.
	Population Level Summary	Response Rate of AN count of 0, 1, or 2 at Week 12 (Yes/No)
Change in pain score from baseline in NRS Patient Global Assessment of Skin Pain at its worst at Week 12	Treatment Condition of Interest	izokibep versus placebo
	Population of Interest	Subjects with HS who meet all inclusion/exclusion criteria and have a baseline numeric rating scale (NRS) ≥ 4
	Variable of Interest	Change in NRS score after 12 weeks of treatment
	Intercurrent Event	Hybrid estimand, with treatment policy strategy used for most ICEs and hypothetical strategy used to account for use of prohibited pain medications due to non-HS-related pain, when taken near a visit. All pain scores after such use of prohibited pain medication (within 28 days of start date of pain medication) will be omitted from the dataset and replaced via multiple imputation in a model analogous to that used for multiple imputation of the primary endpoint. Additional details of multiple imputation described in Section 9.6.4. If prohibited pain medication is used for HS purposes, then any visit within 28 days of the start date of the pain medication will be imputed as NRI. Multiple imputation and NRI will be used to handle other types of missing data, analogously to the primary endpoint.
	Population Level Summary	Response rate, reduction in NRS scores, specifically, at least a 3 point reduction from baseline in NRS in Patient Global Assessment of Skin Pain at its worst, at Week 12 (Yes/No)

6.9 Exploratory Endpoints

The exploratory endpoints are listed below. Due to the decision to write an abbreviated CSR, many exploratory endpoints defined in the protocol have been removed (See Section 5 Changes in the Analysis Specified in the Protocol):

- HiSCR75 at Weeks 2, 4, 8, 16, 32
- HiSCR90 at Weeks 2, 4, 8, 16, 32
- HiSCR100 at Weeks 2, 4, 8, 16, 32
- HiSCR50 at Weeks 2, 4, 8, 16, 32
- AN count of 0, 1, or 2 at Weeks 4, 8, 16, 24, 32
- Change in NRS Patient Global Assessment of Skin Pain at its worst in subjects with baseline NRS ≥ 4 at Weeks 4, 8, 16, 24, 32
- HS flares through Weeks 4, 16 and 52
- Change from baseline to Weeks 4, 8, 16, and 24 in DLQI
- Trough plasma concentrations of izokibep at collected timepoints
- ADAs

6.10 Safety Endpoints

6.10.1 Study Drug Exposure and Compliance

The subjects' exposure to study drug will be calculated separately for Week 16, and Week 52.

The number of doses administered per subject will count the number of administrations of each study drug until Week 16 and until Week 52. A partial dose in CRF will be treated as a half dose.

The duration of exposure will summarize the total time in days while taking study drug. The duration of exposure will be calculated, separately, for Period 1 and the full study period: date of last dose of study drug (prior to Week 16/52) – date of first dose of study drug (Day 1/Week 0) +1 .

For subjects who continue into Period 2, the Period 1 duration calculation will be derived with a slightly different formula: date of Week 16 dose – date of first dose of study drug. If the subject is missing the Week 16 dose but subsequent dosing was done, then 112 days will be imputed for the duration in Period 1.

An additional summary will be provided for Period 2 duration, which will summarize the total time in days taking the study drug in Period 2. The duration of exposure will be calculated as the date of last dose of study drug – date of first dose of study drug in Period 2 (Week 16 for most subjects) + 1.

If the date of first dose(Day 1/Week 0) is missing, then the date of dispense will be used. If the date of dispense is not available, but there is evidence of the subject taking at least 1 dose of the drug, then the date of randomization will be used.

Study drug compliance is defined as the number of doses that were taken relative to the number of doses that should have been taken as per the protocol for the planned duration of treatment exposure. Compliance to study treatment (%) will be calculated as follows for each subject at Week 12 and Week 16:

At Week 16:

$$\frac{\text{number of doses taken from Day 1 until Week 16}}{\text{number of doses expected to be taken from Day 1 until Week 16}} * 100\%$$

At Week 52 (Period 1 + Period 2, Total Study):

$$\frac{\text{number of doses taken from Day 1 until Week 52}}{\text{number of doses expected to be taken from Day 1 until Week 52}} * 100\%$$

At Week 52 (Period 2):

$$\frac{\text{number of doses taken from Week 16 until Week 52}}{\text{number of doses expected to be taken from Week 16 until Week 52}} * 100\%$$

The number of doses expected to be taken by Week 16 is 16, by Week 52 is 52, and by Week 52 (Period 2 only) is 36.

The calculated percentage compliance will be categorized at Week 16, Week 52 (Period 1 + Period 2, Total Study), and Week 52 (Period 2) as:

- < 80% compliance
- ≥ 80% compliance

6.10.2 Adverse Events

All AEs recorded on the CRF will be coded using the latest MedDRA dictionary around the time of Database Lock and classified as either AEs or TEAEs as follows:

- TEAEs in Period 1 are defined as:
  - Any adverse event with a start date/time after the first administration of study drug in the Treatment Period until the date/time of the Week 16 dose administration.
    - If the subject discontinued study treatment prior to taking the Week 16 dose, then 28 days from last dose of study drug administration will be used as the upper limit of the time-period, or
    - If a subject does not have treatment discontinuation prior to Week 16 dose and is missing the Week 16 dose and all subsequent doses, then the study day 112 will be used as the upper limit of the time-period, or
    - In the case of the subject missing Week 16 dose and continuing treatment in the study, the date/time of the first dose of drug in Period 2 will be used as the upper limit of the time-period
    -
- TEAEs in Period 2
  - Events with start date and time after the start of administration of Period 2 treatment and up to 4 weeks after the end of the Treatment Period, or events with start date and time prior to the start of administration of Period 2 treatment whose severity worsens on or after the start of administration of Period 2 treatment and up to 4 weeks after the last dose of study treatment.
    - For subjects who receive treatment in period 2, TEAEs with start date after the start of period 2 will be assigned to period 2 only, not to period 1
- Treatment-emergent SAE are TEAEs that are indicated as Serious ="Yes".

- The relationship between a TEAE and treatment is assessed as related or not related. Treatment-related TEAEs are TEAEs that are indicated as "related to treatment". If relationship to treatment is missing or unknown, then the TEAE is considered "related to treatment".
- TEAEs leading to discontinuation of treatment are TEAEs where the "Action Taken with study treatment" = "Drug Withdrawn"
- Assessment of AE severity will be based on CRF AEs form. TEAEs with missing severity will be included in the counts as severe (Grade 3) unless the TEAE has an outcome of fatal; in that case, the TEAE with missing severity will be included in the counts as death (Grade 5).

6.10.2.1 Injection Site Reaction

Injection Site Reaction (ISR) is defined as

- Any adverse event where "Please check if this event is an injection site reaction" is checked "Yes" or MedDRA High Level Term = "injection site reactions".
    - ISRs are mapped to a particular injection/visit based on the following logic:
        - If the ISR occurs on the day of an injection after administration of treatment or prior to the date/time of the next injection, it should be attributed to the visit where the latest injection occurred.
        - If the ISR occurs after administration of treatment at a particular visit and additional injections do not occur, then the ISR should also be attributed to the visit where the latest injection occurred.

6.10.2.2 Adverse Event of Special Interest (AESI)

The following AESIs will be included:

- Candida infection
- Inflammatory bowel disease
- Suicidal ideation
- Malignancies
- Major adverse cardiovascular and cerebrovascular events (MACE) (cerebrovascular accident and transient ischemic attack, non-fatal myocardial infarction or unstable angina, and cardiovascular death)
- Tuberculosis
- Infections (opportunistic, serious, or fungal only)
- Cytopenias (anemia, neutropenia, lymphopenia, monocytopenia, thrombocytopenia)
- Systemic hypersensitivity reaction

Adverse events of special interest will be identified manually. After all data are entered into the database, a list of all unique preferred terms (PTs) will be generated. ACELYRIN pharmacovigilance and clinical development representatives will review the list and determine which are TEAEs of special interest in each category listed above. This determination will be documented in a memo, signed and dated before database lock for the final analysis and stored in the study trial master file. The memo will specify the subjects and preferred terms to be included as AESIs, including justification for including or excluding events as appropriate. The database will not be locked until the memo is finalized and signed.

6.10.3 Clinical Laboratory Assessments

All laboratory data will be converted to SI units. The change from baseline at each visit until Week 52 for all hematology, clinical chemistry, routine urinalysis, and other tests will be calculated as: *the current visit value - the baseline value*.

All quantitative hematology and clinical chemistry laboratory test values at each assessed visit through Week 52 will be compared with the relevant reference range in SI units and categorized as:

- Low: Below the lower limit of the reference range.
- Normal: Within the reference range (upper and lower limits included).
- High: Above the upper limit of the reference range.

Clinical laboratory results will not be summarized or listed for the purposes of abbreviated CSR, but will be available in the analysis datasets.

6.10.3.1 Liver Function Tests

Subjects who met any (1 or more) of the following liver function test (LFT) criteria at any point after starting study treatment and prior to Week 16 will be included in consideration of potential drug-induced liver injury (DILI) in Period 1:

- Alanine Transaminase (ALT) or Aspartate Transferase (AST) ≥5 x upper limit of normal (ULN) at any time
- ALT or AST ≥3 x ULN at all (2 or more) assessments in a span of ≥4 weeks
- ALT or AST ≥3 x ULN and total bilirubin ≥2 x ULN (>35% direct bilirubin) simultaneously at any time
- ALT or AST ≥3 x ULN and international normalized ratio (INR) ≥1.5 simultaneously at any time
- ALT or AST ≥3 x ULN at any time during the study and weekly assessments after this finding are not available for at least 4 weeks
- ALT or AST ≥3 x ULN associated with symptoms (new or worsening) believed to be related to liver injury or hypersensitivity at any time

To determine subjects who meet the last criterion, a listing of all subjects who have at least 1 ALT or AST ≥3 x ULN at any time after starting study treatment and prior to Week 16, and who do not meet any of the first 5 criteria, will be reviewed by ACELYRIN pharmacovigilance and clinical development representatives along with a listing of all AEs reported by such subjects. In consideration of the investigator's causality assessment, the medical professionals will assess whether any AE is believed to be related to liver injury or hypersensitivity and, if so, whether the elevated ALT or AST is in temporal association with the AE. Any subjects who meet these criteria will be included on a memo that is finalized prior to database lock and that will be filed in the trial master file (TMF).

If there is crossover between Period 1 and Period 2 in any of the instances listed above that require several events over a particular time-span, the LFT flag will be assigned to the period in which the first finding occurs.

Hy's Law is met when all 3 of the following criteria is met simultaneously (at the same draw):

- ALT or AST is ≥3 x ULN
- Total bilirubin ≥2 x ULN

- Alkaline Phosphatase (ALP) is <2 x ULN

At each assessment through Week 16, subjects will be classified as meeting Hy's Law or not meeting Hy's Law.

An analogous approach will be used for evaluating Period 2 for potential DILI. Visits after Week 16 will be considered. Only subjects who took at least one dose of treatment in Period 2 are evaluated for Period 2 potential DILI.

6.10.4 Electrocardiograms (ECGs)

Triplicate 12-lead ECGs will be obtained after subject has been supine for at least 5 minutes using an ECG machine that automatically calculates the heart rate and measures PR, QRS, and QT intervals. The overall interpretation of findings, any abnormal findings, and the clinical significance status are collected on the eCRF.

ECG results will not be listed, but will be available in the analysis datasets.

6.10.5 Vital Signs

The change from baseline at each visit through Week 52 will be calculated as: *the current visit value minus the baseline value*, for the following vital sign parameters:

- SBP (mmHg)
- DBP (mmHg)
- Heart Rate (bpm)
- Respiration Rate (breaths/minute)
- Body Temperature (C) – temperatures collected in Fahrenheit will be converted to C using the following formula: *(temperature in Fahrenheit – 32) x 5/9*

6.10.6 Physical Examination

Clinically significant findings observed prior to the first dose of study drug will be entered on the CRF in medical history in the CRF and reported as medical history. Clinically significant findings observed after the first dose of study drug will be reported as AEs in the CRF and reported as AEs.

Physical examination results will not be listed, but will be available in the analysis datasets..

7. ANALYSIS SETS

The analysis sets are defined in Table 6 below:

Table 6: Analysis Sets

Analysis Set	Definition	Used for Analyses
All Screened Set	All subjects who signed the informed consent form.	Study Disposition
Full Analysis Set (FAS)	All subjects who were randomized. Subjects will be analyzed based on their planned treatment.	Treatment Disposition, Baseline, Protocol Deviations, Efficacy
Full Analysis Set P2 (FASP2)	All subjects who were randomized and took at least one dose of treatment in Period 2.	Efficacy analyses of "Period 2 only".

Analysis Set	Definition	Used for Analyses
	Subjects will be analyzed based on their planned/randomized treatment course.	
Safety Analysis Set (SAF)	All subjects who were randomized and received at least 1 administration of study treatment. Subjects will be analyzed based on the actual treatment received. If a subject receives both treatments, the subject will be grouped with the treatment received most often.	All safety analyses, PK, ADA analyses.
Safety Analysis Set P2 (SAFP2)	All subjects who were randomized and received at least 1 administration of study treatment in Period 2 of the study. Subjects will be analyzed based on the actual treatment course received (Placebo QW – Izokibep 160 mg QW or Izokibep 160 mg QW). If a subject receives both treatments in Period 1, the subject will be grouped with the treatment received most often.	Safety analyses of "Period 2".

8. STATISTICAL METHODS

8.1 General Methodology

Unless otherwise specified, all hypothesis tests will be reported with 2-sided p-values. All confidence intervals (CIs) will be 2--sided with nominal 95% coverage. An adjustment of 0.0001 will be made to account for the unblinded data summaries reviewed by the Independent Data Monitoring Committee (IDMC). Hypotheses testing will be adjusted to control the error rate in the strong sense at $\alpha = 0.0499$. The hypotheses will therefore be tested at $\alpha = 0.0499$.

All data summaries after Week 16 are considered exploratory.

For binary efficacy endpoints, stratified tests will use the 4 strata from the randomization process. If a subject was incorrectly classified during the randomization process, the analysis will use the correct classification as recorded in EDC, not the classification used during randomization.

Tabulations will be produced for demographic, baseline, limited efficacy, and relevant safety parameters. For categorical variables, summary tabulations of the number (n) and percentage (%) of subjects with non-missing data per category will be the default summary presentation, and where appropriate and present, the number of missing values as a "Missing" category. For continuous variables, the number of subjects (n), mean, standard deviation (SD), median, minimum, and maximum will be presented. Coefficient of variation (%CV) may be presented for certain analyses. The number of subjects in the analysis population will be used as denominator for percentages calculation, unless stated otherwise in TFLs mock shell(s).

The treatment group labels, and presentation order are as follows:

- Placebo QW – Izokibep 160 mg QW
- Izokibep 160 mg QW

8.2 Adjustments for Covariates

Each instance of mixed model repeated measures (MMRM) will use the observed baseline of the parameter used in the response variable as well as stratification as covariates.

8.3 Handling of Dropouts or Missing Data

When considering AE and Concomitant Medications, partial start/end dates will be handled in a conservative fashion so that if there is any possibility that the available partial information could reasonably result in the AE being treatment-emergent or the medication being concomitant, that it will be imputed as such. Specifically:

- If the start day is missing but month and year are present:
    - If the known month/year are the same as the month/year of the first dose of treatment, then impute to the date of the first dose of treatment;
    - Otherwise, impute the day to 01.
- If the start day and month are both missing but the start year is present:
    - If the known year is the same as the year of the first dose of treatment, then impute to the date of the first dose of treatment;
    - Otherwise, impute the day and month to 01January.
- If the start date is completely missing, impute to the date of the first dose of treatment, unless the end date suggests it could have started prior to this, in which case impute to 01January of the same year as the end date.
- When imputing a start date, ensure that the new imputed date is on or prior to the end date of the AE or medication.
- If the end date is missing or partial, and the AE/medication is not ongoing (i.e. for AE, the outcome is not "Recovering/Resolving", "Not Recovered/Not Resolved", or "Unknown"), then impute the missing/partial end date as follows:
    - If the end day is missing but month and year are present:
        - If the known month/year are the same as the month/year of the first dose of treatment, then impute to the date of the first dose of treatment;
        - Otherwise, impute the day to the last day of the month.
    - If the end day and month are both missing but the end year is present:
        - If the known year is the same as the year of the first dose of treatment, then impute to the date of the first dose of treatment;
        - Otherwise, impute the day and month to 31December.
    - If the end date is completely missing, impute to the date of the first dose of treatment.
    - When imputing an end date, ensure that the new imputed date is on or after the start date of the AE or medication.
- Other unexpected combinations of partial date information (e.g., known month but missing day and year) will be handled as needed on a case-by-case basis.

The handling of missing efficacy data will be discussed throughout sections 6 and 9.

8.4 Handling of Plasma Concentrations that are Below the Lower Limit of Quantification

Plasma concentrations that are below the limit of quantification (BLQ) will be handled as follows for descriptive statistics:

- Pre-dose concentrations that are BLQ will be treated as zero.
- Post-dose concentrations that are BLQ will be treated as missing.
- Values below the limit of quantification (BLQ) that are set to 0 in the source dataset will be treated as 0 in the calculation of summary statistics.
- Missing values will be excluded from the calculation of means, and n will reflect the actual number of values used in the calculation.
- Arithmetic mean or median values that are BLQ will be presented as 0.
- If any pre-dose BLQ results (treated as 0) are in a series of summarized data, geometric mean and coefficient of variation (CV) % of geometric mean will be reported as not calculated (NC).

8.5 Multi-Center Studies and Pooling of Centers

Although this is a multi-center study, all subjects from all centers will be pooled together for analysis. Treatment-by-center interaction will not be assessed.

8.6 Multiple Comparisons/Multiplicity

Hypotheses tested will be adjusted to control the familywise error rate in the strong sense at $\alpha = 0.050$, 2-sided. An adjustment of 0.0001 will be made to account for the unblinded data summaries reviewed by the DMC. The hypotheses will therefore be tested at $\alpha = 0.0499$.

The statistical comparisons for the primary efficacy endpoint and the secondary endpoints, all at Week 12, will be carried out in sequential order. The primary endpoint, comparing izokibep to placebo, will be tested first, with significance concluded if $p < 0.0499$. Testing of secondary endpoints will only be carried out if all prior tests, including the test of the primary endpoint, first show significance with $p < 0.0499$. If all prior tests are significant, testing will proceed in the following order:

- The first secondary endpoint: proportion of subjects achieving HiSCR90 at Week 12.
- The second secondary endpoint: proportion of subjects achieving HiSCR100 at Week 12.
- The third secondary endpoint: proportion of subjects achieving HiSCR50 at Week 12.
- The fourth secondary endpoint: change in DLQI from baseline to Week 12.
- The fifth secondary endpoint: achieving AN count of 0, 1, or 2 at Week 12 among subjects with baseline Hurley Stage II.
- The sixth secondary endpoint: proportion of subjects who achieve a reduction in pain NRS of at least 3 points from baseline to Week 12, among subjects with a pain NRS score of at least 4 at baseline.
- The seventh secondary endpoint: proportion of subjects who experience flares through Week 12.

If a null hypothesis is not rejected, p-values for subsequent hypotheses in the sequence will be reported as nominal and will not be used to assess objectives or make determinations of efficacy.

At the time of the primary analysis of Week 12 data, some data was available for subjects at Week 16, but not all. A cutoff date was chosen and subjects randomized before that cutoff date were included in summaries and analyses of Week 16 data; subjects randomized after that cutoff date were excluded from summaries and analyses of Week 16 data. When all subjects have completed or discontinued the terminated study at Week 32, a final analysis will be reported, which will include Week 16 data. Since Week 16 data will not be used for any alpha-preserving analyses, there will be no multiplicity adjustment for analyzing Week 16 data twice.

8.7 Interim Analyses and Data Monitoring Committees

An independent Data Monitoring Committee (IDMC) is established to monitor data on an ongoing basis to ensure the continuing safety of the subjects enrolled in this study.

The DMC consists of at least 1 medical expert with the relevant therapeutic area and at least 1 statistician. The DMC has a minimum of 3 members, 1 of whom serves as the Chair. The DMC responsibilities, authorities, and procedures will be documented in the DMC charter. The committee will meet as needed to review significant safety findings. After each review, the DMC will make recommendations regarding the continuation of the study based on safety. The DMC will not be empowered to recommend stopping this study or changing the sample size due to a demonstration of positive efficacy.

Data may be summarized for a DMC at regular intervals during the study to ensure safety of study subjects. The type I error rate for hypothesis tests at the end of the study will be decreased by 0.0001 to account for each interim summary of unblinded data.

After all subjects have had an opportunity to complete the Week 12 visit (complete the visit, or complete a subsequent visit, or permanently discontinue the study before Week 12), the primary endpoint analyses will be conducted. Site staff, and all sponsor staff who interact with site staff, will remain blinded to individual subject treatment assignment until the final subject has completed the final visit and the database lock has occurred.

At various points after the primary analysis, analyses of data collected after Week 12 will be reported for regulatory submission purposes or other reasons. These analyses will be documented. No adjustment to the type I error rate will be applied due to these summaries, because the inferential statistics will already have been determined.

8.8 Examination of Subgroups

No subgroup analyses will be performed for the purposes of the abbreviated CSR.

9. STATISTICAL ANALYSIS

9.1 Subject Disposition and Analysis Sets

Study disposition will be summarized for each treatment group and overall for the All Screened Set. The following information will be included:

- The number of subjects Screened
- The number and percentage of subjects for the following categories:
    - Randomized

  - Treated
  - Not treated
  - Completed Week 12 Assessment
  - Completed Study at Week 12
  - Discontinued Study Before Week 12
  - Reasons for for Study Discontinuation before Week 12
  - Completed Week 16 Assessment
  - Completed Study at Week 16 (Period 1)
  - Discontinued Study before Week 16 (within Period 1)
  - Reasons for Study Discontinuation Prior to Completing Week 16 (Period 1)
  - Completed the Week 52 Assessment
  - Completed Study at Week 52 (Period 2)
  - Discontinued Study before Week 52 (within Period 2)
  - Reasons Study Discontinuation Prior to Completing Week 52 (within Period 2)
  - Discontinued from the study prior to Week 16 (Period 1)
  - Discontinued Study before Week 52
  - Reasons for Study Discontinuation Prior to Completing Week 52
  - Completed Study
  - Discontinued Study
  - Reasons for discontinuing study
- Number and percentage of subjects included in, and excluded from, each analysis set together

The denominator for the calculations above will be the number of randomized subjects in each group, except for the proportion of subjects randomized. For the proportion of subjects randomized, the denominator will be the number of subjects screened.

Treatment disposition will be summarized for each treatment group and overall for the FAS. The following information will also be included:

- The number and percentage of subjects for the following categories:
  - Completed treatment through Week 11
  - Discontinued from treatment before Week 11
  - Reasons for treatment discontinuation prior to the Week 11
  - Completed treatment through Week 15 (Period 1)
  - Discontinued from treatment prior to Week 15 (within Period 1)
  - Reasons for Treatment Discontinuation before Week 15 (within Period 1)
  - Completed treatment through Week 51 (Period 2)
  - Discontinued from treatment before Week 51 (within Period 2)
  - Reasons for Treatment Discontinuation before Week 51 (within Period 2)
  - Discontinued from treatment any time before Week 51
  - Reasons for Treatment Discontinuation any time before Week 51
  - Discontinuing treatment within the following time categories:
    - Day 1
    - Day 2 - Day 14 (within 2 weeks)
    - Day 15 - Day 28 (>2 - 4 weeks)
    - Day 29 - Day 56 (>4 - 8 weeks)

- Day 57 - Day 98 (>8 - 12 weeks)
- Day 99 - Day 112 (>12 - 16 weeks)
- Day 113 – Day 140 (>16 - 20 weeks)
- Day 141 – Day 168 (>20 – 24 weeks)
- Day 169 – Day 224 (>24 – 32 weeks)
- Day 225 – Day 364 (>32 – 52 weeks)
- > Day 364 (> 52 weeks)

9.2 Protocol Deviations

Protocol deviations will not be summarized or listed for the abbreviated CSR. The data will be available in the SDTM datasets.

9.3 Demographics and Baseline Characteristics

Demographic parameters and other baseline characteristics will be summarized by treatment group and overall for the FAS.

Descriptive statistics (n, mean, SD, median, minimum, and maximum) will be summarized for:

- Age (years) at Screening
- Height (cm)
- Weight (kg)
- Body Mass Index (kg/m^2)
- Body Temperature (C)
- Respiratory Rate (breaths per minute)
- SBP (mmHg)
- DBP (mmHg)
- Heart Rate (bpm)
- Time since diagnosis with HS (years)
- AN Count
- Abscess Count
- Draining Fistula Count
- Non-Draining Fistula Count
- Total Fistula Count
- Inflammatory Nodule Count
- Non-inflammatory Nodule Count
- HADS Total Score
- HADS-Anxiety
- HADS-Depression
- PGA
- DLQI
- EQ-5D (Domain Scores [Mobility, Self-Care, Pain/Discomfort, Usual Activities, and Anxiety/Depression], Index Score, Visual Analog Scale (VAS) Score)
- SF -12 (Mental Component Score and Physical Component Score)
- NRS Pain – At its worst
- NRS Pain – On average

- IHS4
- HiSQOL

The number and percentage of subjects in each category will be summarized by treatment group and overall for the FAS:

- Age groups: <65 years vs. ≥65 years
- Sex
- Race
- Ethnicity
- Hurley Stage (II, III) (Using the 'worst case' value collected on the baseline Lesion Counts eCRF page)
- Childbearing potential (Yes, No)
- Alcohol Use (Never, Current, Former)
- Nicotine Use (Never, Current, Former)
- C-SSRS – Lifetime Ideation
- C-SSRS – Past 6 Months Ideation
- Location of lesions

Medical History will be summarized by treatment group and overall for the FAS. The number and percentage of subjects with at least 1 medical history event will be presented. The number and percentage of subjects and in each system organ class (SOC) and preferred term (PT) will be summarized. Subjects will be counted once within each SOC and PT. The summary will be sorted using the internationally agreed order for SOC and using descending order of overall numerical counts for PT. Where terms tie, these will be sorted alphabetically.

9.4 Prior and Concomitant Medications

Prior and concomitant medications will be summarized separately for the FAS by treatment group and overall. The number and percentage of subjects with at least 1 prior/concomitant medication will be presented. The number and percentage of subjects with at least 1 prior/concomitant medication within each Anatomical Group (ATC Level 1), Therapeutic Subgroup (ATC Level 2), and preferred term will be presented. The summary will be sorted using numerical counts by descending order of Anatomical Group, then descending order of Therapeutic Subgroup, then descending order of preferred term in the total column. Where groups or terms tie these will be sorted alphabetically.

Concomitant medications will be summarized for Period 1, and, separately, for Period 2.

9.5 Analysis of Primary Efficacy Endpoint

The primary efficacy variable is defined as HiSCR75 at Week 12.

The statistical null and alternative hypotheses to be used to assess the primary objective are:

$$H_0: \pi_{ABY} - \pi_{PBO} = 0$$

$$H_A: \pi_{ABY} - \pi_{PBO} \neq 0$$

where π_{ABY} and π_{PBO} are the proportion of subjects achieving HiSCR75 at Week 12 among subjects randomly assigned to receive izokibep and placebo, respectively.

The treatment policy strategy approach for the estimand will be used in general, so subjects will be included using observed data at Week 12 regardless of treatment compliance, use of rescue medications or procedures except as noted below, or any other protocol deviations.

Subjects who receive antibiotic therapy that could affect HS will be included in the primary analysis by imputing as non-responders (non-response imputation or NRI) at subsequent timepoints. A list of all oral antibiotic therapy that results in such imputation will be finalized after review of all oral antibiotic therapy received by any subject and before unblinding of the study, and will include tetracycline, clindamycin, and possibly other products. Similarly, subjects with missing HiSCR75 assessments at Week 12 due to discontinuing treatment due to an AE or lack of efficacy will also be imputed as NRI. These subjects will be identified by meeting any of the following criteria:

- Subject has adverse event with action taken = "Drug Withdrawn" and corresponding AE start date is before study day 84 and last dose of study drug is taken before study day 84 or
- Subject has last dose of study drug taken before study day 84 and reason for discontinuation of treatment is any of the following:
    - Adverse event
    - Death
    - Lack of efficacy
    - Disease progression

Other subjects with missing data will have efficacy data imputed using multiple imputation, as follows:

Multiple imputation will be implemented to replace missing endpoint data using multiple draws from a posterior predictive distribution estimated from the placebo group only. First, a monotone missingness pattern will be enforced using Markov Chain Monte Carlo (MCMC). The MCMC process will use predictor variables of:

- baseline abscess count (continuous)
- baseline inflammatory nodule count (continuous)
- baseline draining fistula count (continuous)
- age (continuous)
- baseline BMI (continuous)

for nonmonotone missingness at Week 2 (missing at Week 2 but available at a later timepoint). MCMC predictors in the regression model for missing values after Week 2 will include the predictor variables listed above, plus:

- counts of abscess (continuous)
- inflammatory nodule (continuous)
- and draining fistula (continuous)

collected at prior scheduled assessments. After monotone missingness is enforced, missing HiSCR75 at Week 2, 4, 8, 12, and 16 will be imputed, assuming monotone missingness pattern. Any missing HiSCR75 will be imputed by first imputing components of the AN count (abscess and inflammatory nodules) and draining fistula count, then assigning HiSCR75 status from the imputed values. Predictors in the regression model for missing values at Week 2 will be:

- baseline Hurley stage (II/III) ('worst case' value selected from the Lesion Count eCRF page at baseline)
- baseline abscess count (continuous)
- baseline inflammatory nodule count (continuous)
- baseline draining fistula count (continuous)
- sex (male/female)
- race (white/non-white)
- age (continuous)
- BMI (continuous)
- Prior Biologic/Janus Kinase (JAK) Inhibitor Use for HS (yes/no).
- baseline smoking/tobacco status (current, former, never)
- highest categorization of location of baseline HS abscesses (difficult to treat, easier to treat)
    - A subject can have HS abscesses in multiple locations around their body and will indicate each location at baseline. The locations are categorized based on how difficult they are to treat (in bullets below). The subject will be assigned to a category based on the location of their most difficult to treat abscess at baseline.
        - Difficult to treat: left/right axilla, left/right sub/inframammary region, left/right buttock, left/right inguinocrural fold, perianal, perineal, genitalia, gluteal cleft, labia majora, abdominal cavity, arm, flank, thigh, mammary, mons pubis, thigh, vulva, scrotum, sternum, stomach fold, suprapubic region, chest, sacrum, inframammary, labia majora, sternal area, face, jaw area, and breast
        - Easier to treat: all remaining locations
- highest categorization of location of baseline HS draining fistulas (difficult to treat, easier to treat, per logic described above)
- highest categorization of location of baseline HS inflammatory nodules (difficult to treat, easier to treat, per logic described above)

Predictors in the regression model for missing values after Week 2 will be all these variables, plus:

- counts of abscess (continuous)
- inflammatory nodule (continuous)
- and draining fistula (continuous)

at prior scheduled assessments.

Missing HiSCR75 values in both the placebo group and the izokibep group will be imputed using observed data from the placebo group only.

Data will be processed sequentially by repeatedly calling SAS® PROC MI to impute missing outcomes at Week 4, 8, 12, and 16. A total of 100 imputations will be imputed to form the complete data set with a pre-specified seed (891523) to ensure reproducibility. These multiple imputed data sets are then analyzed by using the same method for the primary analysis for complete data. The final inference on treatment difference is conducted from the multiple datasets using Rubin's combining rules. Imputed values without rounding will be used at each imputation step.

Example SAS code for the PROC MI procedure described above:

```
/*Imputation to monotone missing pattern*/

        proc mi data = data seed=891523 nimpute = 100 out = t2
                min = . . . . .
                0 0 0 0 0;
                var age bmibl bl_abscnt bl_infncnt bl_dfistcnt week_2 week_4 week_8 week_12
                    week_16;
                mcmc chain = multiple impute = monotone;
                EM MAXITER=1000000;
        run;
/*Complete imputation*/

        proc mi data = data seed=891523 out = out nimpute = 1 minmaxiter=600
                min = . . . . . .
                . . . . .;
                by _imputation_;
                class trt01p stratum1 sex race stratum2 locdiff_ab  locdiff_df locdiff_in smkstat;
                var bl_abscnt bl_infncnt bl_dfistcnt locdiff_ab locdiff_df locdiff_in smkstat sex race
                    stratum1 stratum2 age bmibl week_2 week_4 week_8 week_12 week_16;
                monotone reg (week_2 = bl_abscnt bl_infncnt bl_dfistcnt locdiff_ab locdiff_df locdiff_in
                    smkstat sex race age bmibl stratum1 stratum2);
                monotone reg (week_4 = bl_abscnt bl_infncnt bl_dfistcnt locdiff_ab locdiff_df locdiff_in
                    smkstat sex race age bmibl stratum1 stratum2 week_2);
                monotone reg (week_8 = bl_abscnt bl_infncnt bl_dfistcnt locdiff_ab locdiff_df locdiff_in
                    smkstat sex race age bmibl stratum1 stratum2 week_2 week_4);
                monotone reg (week_12 = bl_abscnt bl_infncnt bl_dfistcnt locdiff_ab locdiff_df
                    locdiff_in smkstat stratum1 stratum2 sex race age bmibl week_2  week_4
                    week_8);
                monotone reg (week_16 = bl_abscnt bl_infncnt bl_dfistcnt locdiff_ab locdiff_df
                    locdiff_in smkstat stratum1 stratum2 sex race age bmibl week_2 week_4
                    week_8 week_12);
        mnar model(week_2 week_4 week_8 week_12 week_16 / modelobs= (trt01p='Placebo QW'));
        run;
```

locdiff_ab = location difficulty of abscesses; locdiff_df = location difficulty of draining fistulas; locdiff_in = location difficulty of inflammatory nodules; trt = treatment group; smkstat = smoking status; strat = stratum; bl_abscnt = baseline abscess count; bl_infncnt = baseline inflammatory nodule count; bl_dfistcnt = baseline draining fistula count

Other imputation, fully conditional specification method may be implemented if the multiple imputation model does not converge. If imputation is not possible due to sparseness of data for certain predictor variables, that predictor variable will be collapsed into a single category, in essence removing that predictor from the predictor model. In the case that baseline predictors need to be removed from the model, the location difficulty of inflammatory nodules variable will be removed first, followed by location difficulty of abscesses, location difficulty of draining fistulas, smoking status, sex, race, age, BMI, baseline abscess count, baseline inflammatory nodule count, and baseline draining fistula count. The number of iterations may be increased, if the imputed values are outside the allowable specifications.

If negative values result from the multiple imputation procedure, then these will be set to 0.

Results from all runs will be combined using Rubin's rule (Rubin, 1987), resulting in a single point estimate, a single CI, and a single p-value. If the imputed datasets do not contain sufficient variability to combine them together, then the first imputed dataset will be selected for analysis and statistical summaries will be run based on that data.

The number of subjects imputed with NRI as well as the number of subjects imputed with multiple imputation will be tabulated with reasons for imputation at Week 12.

A stratified test of response rates will be used for the primary analysis. Within each of the 4 strata used for randomization, the response rate for each treatment group and corresponding standard error will be calculated. The difference in response rates (risk difference) will be calculated for each stratum. The common risk difference among the 4 strata and associated standard error will be estimated by combining the observed risk differences using Cochran-Mantel-Haenszel weighting. The estimated risk difference divided by the standard error will be used as the test statistic and a p-value calculated assuming that the test statistic follows a standard normal distribution under the null hypothesis. Analyses at earlier timepoints when data to calculate the HiSCR75 are collected will also be presented using the same methodology. P-values from earlier timepoints and Week 16 will be presented for descriptive purposes, not part of the alpha-preserving multiple testing strategy.

Sample SAS code with stratification factors of comparing each dosing regimen of izokibep to the combined placebo group is given as following:

```
ods output commonpdiff = ;
proc freq data= ;
   tables stratum * treatment * response / cmh commonriskdiff (test=mh cl=mh);
   weight count / zeros;
run;
```

9.5.1 Sensitivity Analysis of the Primary Endpoint

Sensitivity analyses will not be performed for the purposes of the abbreviated CSR.

9.6 Analysis of Secondary Efficacy Endpoints

The secondary efficacy variables of HiSCR50, HiSCR90 and HiSCR100 at Week 12 will be analyzed using the same approach as the primary efficacy analysis.

9.6.1 Subjects Experiencing HS Flares through Week 12

The statistical null and alternative hypotheses to be used to assess HS Flare are:

$$H_0: \pi_{ABY} - \pi_{PBO} = 0$$

$$H_A: \pi_{ABY} - \pi_{PBO} \neq 0$$

where π_{ABY} and π_{PBO} are the proportion of subjects experiencing HS flares at any time from first dose of study treatment through Week 12 among subjects randomly assigned to receive izokibep and placebo, respectively.

Subjects who received an oral antibiotic that could impact HS at any point before the Week 12 assessment or subjects who discontinue treatment due to AE or lack of efficacy will be included in the analysis with NRI (did experience HS flare). All subjects with missing abscess, inflammatory nodule, or draining tunnel counts will be included in the multiple imputation. The multiple imputation process will be the same as described for HiSCR75, except that draining fistula counts do not play a role in flare and will therefore not be imputed when missing.

Subjects with missing AN counts at any scheduled visit (week 4, week 8, and/or week 12) will have their abscess and inflammatory nodules counts imputed (separately) for each visit with missing data. The imputed counts will be summed and compared to criteria for flare. If the observed or imputed value at any visit meets the definition of flare, the subject will be counted as having a flare; otherwise, the subject will be counted as not having a flare. A total of 100 imputations will be imputed to form the complete data set with a pre-specified seed (3617627) to ensure reproducibility. These multiple imputed data sets will then be analyzed by using the same method for the primary analysis for complete data. The final inference on treatment difference is conducted from the multiple datasets using Rubin's combining rules. Other imputation, fully conditional specification method may be implemented if the multiple imputation model does not converge. Troubleshooting (removing predictor variables or increasing the number of iterations) will be done in the same manner as for the primary endpoint.

The number and proportion of subjects for subjects who experience a HS flare through Week 12 in each treatment group will be reported analogous as the primary endpoint.

9.6.1.1 Sensitivity Analysis of Subjects Experiencing HS Flares through Week 12

Sensitivity analyses will not be done for the purposes of the abbreviated CSR.

9.6.2 Change from Baseline in DLQI at Week 12

DLQI mean values, mean changes and associated descriptive statistics along with 95% CI (for change from baseline) through Week 12 for each treatment group will be provided.

Treatment policy estimand strategy will be used, in general, with all subjects included in the analysis regardless of treatment discontinuation, concomitant medication/rescue medication, protocol deviations, or other actions.

DLQI change from baseline will be analyzed using a mixed model repeated measure (MMRM). The model for the comparison will include treatment, baseline DLQI, stratification factors, visit week, and treatment by visit week interaction as covariates. LS (least square) mean changes and associated SEs with associated 95% CI and p-values will be presented. Model will include all data up to Week 12. An unstructured (UN) covariance matrix will be used. Other variance-covariance matrices, such as autoregressive (AR), and compound-symmetric (CS) will be considered and specified in the footnote if the model does not converge. Sample code is as follows:

```
proc mixed ;
  class week stratum1 stratum2 subjectid treatment(ref = "placebo")
  model change_dlqi = base_dlqi stratum1 stratum2 week treatment treatment*week
  repeated week / subject = subjected (treatment) TYPE = UN;
  lsmeans treatment*week /slice = week cl;
  slice treatment*week/ sliceby = week pdiff cl;
run;
```

9.6.3 Baseline Hurley Stage II Subjects who Achieve AN count of 0, 1, or 2 at Week 12

The secondary endpoints of percentage of subjects with baseline actual Hurley Stage II who achieved AN count of 0, 1, or 2 at Week 12 will be analyzed analogously to the primary endpoint. The statistical null and alternative hypotheses to be used to assess these objectives are:

$$H_0: \pi_{ABY} - \pi_{PBO} = 0$$

$$H_A: \pi_{ABY} - \pi_{PBO} \neq 0$$

where π_{ABY} and π_{PBO} are the proportion achieving AN count of 0, 1, or 2 at Week 12 among subjects randomly assigned to receive izokibep and placebo, respectively.

Subjects who have missing data at Week 12 due to discontinuation of treatment due to adverse event or lack of efficacy or who received an oral antibiotic that could impact HS will be included in the analysis with NRI (did not achieve AN count of 0, 1, or 2). Other missing data will be handled using multiple imputation analogously to the primary endpoint. Only subjects with Hurley Stage II at baseline will be included in the hypothesis test for AN count, while other subjects will be summarized, and a p-value reported for descriptive use only.

Hurley stage collected in the eCRF ('worst case' Hurley stage from the baseline Lesion Counts eCRF page) will be used.

Methodology will be identical to that used for the primary endpoint with the exception of stratification factors. For the secondary endpoint, achieve AN count of 0, 1, or 2 at Week 12, which uses only subjects with Hurley Stage II at baseline, the stratification factor of baseline Hurley Stage will be omitted from the analysis. For the analysis using all subjects, the stratification factor of baseline Hurley Stage will be used, as in the primary endpoint analysis.

9.6.4 NRS of Skin Pain at Week 12

The secondary endpoint of percentage of subjects achieving at least 3 points reduction from baseline in NRS in Patient Global Assessment of Skin Pain at its worst at Week 12 among participants with baseline NRS ≥ 3 will be analyzed analogously to the primary endpoint.

The number and proportion of subjects achieving at least 3 points reduction from baseline in NRS in Patient Global Assessment of Skin Pain at its worst at Week 12 among participants with baseline NRS ≥ 4 in each treatment group will be reported.

Subjects who have missing data at Week 12 due to discontinuation of treatment due to adverse event or lack of efficacy will be included in the analysis with NRI (did not achieve least 3 points reduction from baseline in NRS in Patient Global Assessment of Skin Pain at its worst). If a subject received certain prohibited analgesic therapy within 28 days of the visit (including the day of the visit) for HS indication, then the visit within the 28 days of the start date of the prohibited analgesic will be imputed as NRI.

If a subject received certain prohibited analgesic therapy within 28 days of the visit (including the day of the visit) for non-HS indication, the subject will have the pain scores after such use of prohibited pain medication omitted from the dataset due to the ICE and replaced via multiple imputation in a model analogous to the one used for the primary analysis. This process will only be done for the visit(s) within

28 days of the start date of the prohibited analgesic therapy for non-HS indication. For the multiple imputation, the model used will be analogous to the one for the primary endpoint, however the baseline NRS (at its worst) score will be added as a baseline covariate. In addition, the imputation will be done for the average NRS value for the visit.

Other missing data will be handled using multiple imputation analogously to the primary endpoint (including baseline NRS [at its worst] score as a covariate in the model).

Allowable and prohibited analgesics are pre-specified in sections 6.8.2.3 (Analgesic Therapy) and 6.8.3 (Prohibited Medications) of the protocol. A final list of all subjects who took prohibited analgesic therapy which results in NRI or data omission and subsequent multiple imputation will be created after the data is final and before unblinding and will include justification for choosing the subjects based on specific analgesic, timing relative to the visit, and other factors considered.

Since the NRS in Patient Global Assessment of Skin Pain is collected as a daily diary, the score used for analysis will be the average of the scores recorded 6 days prior to and the day of the visit (7 days total for the average). If scores are recorded on fewer than 3 of these 7 days, the pain score will be treated as missing; otherwise, the mean of the observed values (up to 7) will be used. The same approach to calculating the average value will be used for any visit occurring through Week 16 visit, including End of Treatment visits windowed to a visit on or before Week 16.

9.7 Analysis of Exploratory Endpoints

Limited exploratory endpoints will be analyzed analogously to primary and secondary endpoints with the same strategies for ICEs.

Binary endpoints in Period 1 will be analyzed with a stratified test of risk difference and continuous endpoints will be analyzed with an MMRM model in general. Risk differences, differences in LSMeans, and p-values will be calculated for time-points in Period 2, but will be used for descriptive purposes only.

Continuous data at timepoints before and including the primary timepoint will use all data from planned assessments up to and including the primary timepoint; continuous data after the primary timepoint will use all data from all planned assessments. Baseline parameter value, Stratification variables, visit week, treatment group and interaction of treatment and visit week will be included in the MMRM model.

Analyses of binary endpoints in Period 1 or Period 1 + Period 2 will use the full FAS as the analysis set. Analyses of binary endpoints of Period 2 data only will use the subjects in FAS who have at least one dose of treatment in Period 2.

Results for continuous endpoints will presented as LS mean changes and associated SEs with associated 95% CI and p-values. Exploratory endpoint results will also be listed.

The following binary endpoints will be analyzed analogously to the primary endpoint of HiSCR75 at Week 12:

- HiSCR50, HiSCR75, HiSCR90 and HiSCR100 at Weeks 2, 4, 8, 16, 24, and 32;
- AN count of 0, 1, or 2 at Weeks 4, 8, 16, 24, and 32 among subjects with baseline Hurley Stage II;

- Proportion with at least 3 points reduction from baseline in NRS in Patient Global Assessment of Skin Pain at its worst, among subjects with baseline NRS ≥ 4, at weeks 4, 8, 16, 24, and 32. NRI will be handled analogously to the corresponding secondary endpoint analysis.
- HS flares through Weeks 4, 16, and 52;

The binary endpoints listed above (excluding Flares) at the Week 24, 32, and/or Week 52 timepoint will also be analyzed for just the subgroup of FAS who had at least one dose of study drug in Period 2. The ICE handling strategy corresponding to NRI due to antibiotic use will be slightly modified for the analyses of responder rates corresponding to Period 2 visits: subjects originally randomized to placebo treatment and who receive antibiotic therapy that could affect HS *after the first dose of study treatment in Period 2* will be included in the analysis by imputing as non-responders (non-response imputation or NRI) at subsequent timepoints. If these subjects took antibiotics that could affect HS *prior to the first dose of study treatment in Period 2*, they will not be automatically assigned as non-responders unless the forbidden antibiotic was taken in Period 2.

The following continuous endpoint will be analyzed with an MMRM model. LS mean changes and associated SEs with associated 95% CI and p-values will be presented. For time-points on or before Week 16, the model will include only data up to Week 16. For time-points after Week 16, the model will include all available time-points. Mean values, mean changes and associated descriptive statistics along with 95% CI (for change from baseline) according to the treatment group:

- Change from baseline in DLQI at Weeks 4, 8, 16, and 24;

All other exploratory endpoints will not be summarized for the purposes of the abbreviated CSR.

9.8 Analysis of Safety Endpoints

9.8.1 Study Drug Exposure and Compliance

Subject exposure to treatment will be summarized by treatment group and overall for the SAF.

Number of doses, duration of treatment, and percent compliance will be summarized by treatment group and overall through Week 16 and, separately, through Week 52 with descriptive statistics (n, mean, SD, median, minimum, and maximum). Compliance categories will be summarized by treatment group and overall for Week 16 and Week 52 with the number and percentage of subjects in each category. For the analyses through Week 52, a separate summary will be done for period 2, and separately, for the total study exposure.

9.8.2 Adverse Events

The number and percentage of subjects with AEs as well as the number of events will be summarized by MedDRA-coded SOCs and PTs.

The following AE summaries occurring through Week 16 will be presented by treatment group and overall:

- Overall summary of TEAEs
- TEAEs by SOC and PT
- TEAEs by Cumulative Severity Category by SOC, and PT
- Related TEAEs by SOC and PT

- Related TEAEs by Cumulative Severity Category by SOC, and PT
- Serious TEAEs by SOC and PT
- Treatment-related Serious TEAEs by SOC and PT
- TEAEs of Special Interest by SOC and PT
- TEAEs leading to discontinuation by SOC and PT
- TEAEs Experienced by >= 5% Subjects in Izokibep Group by SOC and PT
- Summary of ISRs
    - Total Number of subjects having at least 1 ISR (including count of events)
    - Subjects and Event ISRs Count by severity in the following categories:
        - Grade 1 or above
        - Grade 2 or above
        - Grade 3 or above
        - Grade 4 or above
    - Summary of signs and symptoms (subject counts/percentages and count of events)
    - ISRs (Subjects) Count by Injection Visit (if 1 subject has an ISR corresponding to multiple visits, the subject is counted in both)

The AE summaries listed above will also be produced for Period 2 for subjects with at least one dose of study drug in Period 2 (SAFP2). These will also be presented for Period 1 + Period 2 (Total study) combined for the safety analysis set (SAF).

Within each summary, a subject will only be counted once within each SOC and PT.

Summaries by SOCs and PTs will be sorted by SOCs by descending order of total incidence and PTs within SOC by descending order of total incidence. Where PTs tie, PTs will be sorted alphabetically. Summaries by PTs only will be sorted by descending order of total incidence.

All AE data will be listed and AEs and TEAEs will be presented together. Treatment-emergence status as well as period will be indicated in the listing. The listing will present the relative start and stop day of the AE calculated relative to the first dose of treatment and will be presented for those subjects who received at least 1 dose of treatment. If the AE is "Ongoing" it will be indicated as such in the listing and the relative stop day will not be calculated. Only original dates will be presented in the listing even though the relative day may be based on an imputed date.

The following listings will be presented for the SAF:

- Listing of all AEs (flag will be included for TEAE and safety follow-up AE)
- Listing of Serious TEAEs
- Listing of AEs Leading to Discontinuation of Study Treatment
- Listing of AEs of Special Interest
- Listing of ISRs
- Listing of Deaths

9.8.3 Clinical Laboratory Assessments

Clinical lab data will not be summarized or listed for the purposes of abbreviated CSR.

9.8.3.1 Liver Function Tests

A summary table will present the count/proportion of subjects who meet each of the 6 criteria for consideration of potential DILI and subjects meeting Hy's Law at any time through Week 16, and, separately, for any time through Week 52. A similar summary table will be created for Period 2 only, for just the subjects in the SAF with at least one dose of study treatment in Period 2. Baseline results will not be included in the table; only post-baseline liver function test values are flagged.

Potential liver toxicity will be assessed by listing subjects who met potential drug induced liver injury (pDILI) criteria as described in the protocol and by summarizing subjects who met Hy's Law for the SAF.

For subjects who meet any of the 6 criteria for consideration of potential DILI and subjects being classified as meeting Hy's Law or not meeting Hy's Law, all LFT results at any time through Week 52 will be listed, sorted by subject and date. Listing will include subject, treatment assignment, ALT, AST, ALP, bilirubin and INR, along with specific criterion met for DILI as well as met/not met Hy's Law. Subjects who met the criteria prior to or at Week 16, and anytime in Period 2 will be flagged.

9.8.4 Electrocardiograms

ECGs results will not be summarized or listed for the purposes of abbreviated CSR.

9.8.5 Vital Signs

Vital signs will not be summarized or listed for the purposes of abbreviated CSR.

9.8.6 Physical Examination

Physical examination findings will not be summarized or listed for the purposes of abbreviated CSR.

9.8.7 Other Endpoints

No additional endpoint analyses are planned.

9.9 Pharmacokinetic Assessments

Pharmacokinetic data will be summarized for the safety analysis set by treatment group and scheduled visit, including N (number of non-missing data), N BLQ (for concentrations only), arithmetic mean, standard deviation (SD), arithmetic coefficient of variation (CV%), geometric mean, geometric CV%, median, minimum, and maximum. Only trough collections are planned, so no calculations of PK parameters (maximum concentration, area under the concentration curve, etc.) will be provided.

For the calculation of summary statistics, pre-dose concentrations that are below the limit of quantification (BLQ) will be treated as zero and post-dose concentrations that are BLQ will be treated as missing.

Pharmacokinetic data may be combined with data from other studies of izokibep for modeling. If this is done, details will be described in a separate document.

9.10 Anti-Drug Antibodies

Samples tested for ADAs will be reported as positive or negative based on a confirmatory assay. Any continuous values of concentration (titers) will be summarized using observed data.

Subjects will be summarized at each scheduled visit according to the confirmatory test, with number (%) positive.

9.11 Subgroup Analyses

Subgroup analyses of the efficacy data will be not be done, due to the decision to write an abbreviated CSR.

10. COMPUTER SOFTWARE

Statistical analyses will be performed using SAS® version 9.4 or higher in a Windows environment, or R version 4.1.1 or later.

11. BIBLIOGRAPHY

Daoud Mathieu, S. M. (2023, Apr 17). Overview and comparison of the clinical scores in hidradenitis suppurativa: A real-life clinical data. *Frontiers in medicine,, 10*, 1145152. doi:10.3389/fmed.2023.1145152

Kirby, J. S. (2020). The Hidradenitis Suppurativa Quality of Life (HiSQOL) score: development and validation of a measure for clinical trials. *The British journal of dermatology, 183*(2), 340–348. doi:10.1111/bjd.18692

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys.* New York: John Wiley & Sons, Inc.

Sartorius K, E. L. (2009). Objective scoring of hidradenitis suppurativa reflecting the role of tobacco smoking and obesity. *The British journal of dermatology, 161*(4), 831–839. doi:10.1111/j.1365-2133.2009.09198.x

Ware, J. E.-B. (2002). How to score version 2 of the SF-12® health survey (with a supplement documenting version 1).

Zigmond AS, S. R. (1983). The Hospital Anxiety and Depression Scale. *Acta psychiatrica Scandinavica, 67*(6), 361–370. doi:10.1111/j.1600-0447.1983.tb09716.x

Zouboulis CC, T. T. (2017). Development and validation of the International Hidradenitis Suppurativa Severity Score System (IHS4), a novel dynamic scoring system to assess HS severity. *British journal of dermatology, 177*(5), 1401–1409. doi:10.1111/bjd.15748

12. TABLE/FIGURE/LISTING SHELLS

The mock shells for the tables/figures/listings will be provided in a separate document.

13. APPENDICES

13.1 End Points Definitions and Scoring Method

13.1.2 Modified Sartorius Score

The modified Sartorius score includes an assessment of the anatomical regions involved, the numbers and scores of lesions for each region, the longest distance between 2 relevant regions (or size of a single lesion), and whether all lesions are separated by normal skin (yes or no). The detailed scoring is as following (Sartorius K, 2009):

Each region will score as defined in Table 7:

Table 7: Scoring Modified Sartorius Score

Item No	Items	Scoring Rules	Points	Score:
1	Anatomical region	If the region is affected	3	
		If the region is not affected, the score for that region is 0, and no further calculations are needed for the region. If the region is affected, the following are added in addition to the 3 points for the region in line 1 above.		
	For each affected region:			
2	Nodules (Inflammatory)	Count from this region	1	1 × count
3	Fistulae (Draining only)	Count from this region	6	6 × count
4	Longest distance between 2 relevant lesions within a region or size of the lesion if single	< 5 cm	1	
		5 - 10 cm	3	
		> 10 cm	9	
5	Lesions separated by normal skin?*	Yes	0	
		No	9	
	Score for the region			Sum of above

* If Lesions separated by normal skin = "No" and/or Hurley Stage = "STAGE III" then 9 points is added for the region.

Score for each region is (if the region involved) = 3 + (number of inflammatory nodule) × 1 + {(number of draining fistula) × 6 +{0 < (The longest distance between 2 relevant lesions [or size of lesion if single]) <5} × 1+ {5≤(The longest distance between 2 relevant lesions [or size of lesion if single]) ≤ 10} × 3 + {(The longest distance between 2 relevant lesions [or size of lesion if single]) > 10 } × 9 + (all lesions separated by normal skin = Yes or Hurley Stage = "STAGE III") × 9

Modified Sartorius score is the sum of the score of each region. Any region with no involvement will have a score of zero for purposes of calculating the modified Sartorius score.

A region is considered "affected" if any of the following conditions are true:

- Abscess Count > 0 OR
- Any "Interventions needed DATE" question is not missing.
- Any "How many lesions had intervention" question > 0
- Non-draining fistula count > 0
- Draining fistula count > 0
- Non-inflammatory nodule count > 0
- Inflammatory nodule count > 0
- Hypertrophic scar > 0
- Longest distance > 0
- Normal appearance tissue = NO
  Hurley Stage contains "STAGE"

13.1.3 EQ-5D

The EQ-5D descriptive system comprises the following 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has 3 levels: no problems, some problems, and extreme problems. A unique health state is defined by combining 1 level from each of the 5 dimensions. There are 243 possible health states defined in this way. Each state is referred to in terms of a 5-digit code. No health state value will be calculated if any of the dimensions are missing.

Table 8: Domain Scores of EQ-5D-3L Descriptive System

Response Category	Score
No problems	1
Moderate problems	2
Severe problems	3

Note: Missing values are coded as '.'. Ambiguous values (e.g., 2 boxes are ticked for a single dimension) should be treated as missing values.

EQ-5D-3L = sum(non-missing question score)/(number of non-missing questions) × 5. If the number of questions with a missing answer is ≥ 3 (≥ 50%) then EQ-5D-3L is missing.

For example, if a patient provides the following responses: Mobility = No problems, Self-care = No problems, Usual Activities = Moderate problems, Pain/Discomfort = Moderate problems, Anxiety/Depression = Extreme problems, his response sequence is 11223. The 243 theoretical possible sequences can then be mapped to an index value to provide a summary across all dimensions. For the

calculation of the index value, the UK value set based on Time Trade-Off method will be used. The scoring algorithm is defined in Table 9:

Table 9: Scoring Algorithm for EQ-5D-3L

Full health (11111) =	1.000
At least 1 domain at 2 or 3 (N2):	-0.081
At least 1 domain at 3 (N3):	-0.269
Mobility level 1:	0
Mobility level 2:	-0.069
Mobility level 3:	-0.314
Self-care level 1:	0
Self-care level 2:	-0.104
Self-care level 3:	-0.214
Usual activities 1:	0
Usual activities 2:	-0.036
Usual activities 3:	-0.094
Pain/discomfort 1:	0
Pain/discomfort 2:	-0.123
Pain/discomfort 3:	-0.386
Anxiety/depression 1:	0
Anxiety/depression 2:	-0.071
Anxiety depression 3:	-0.236

The index for the patient with example sequence 11223 = 0.255, calculated as follows:
Full health = 1
Minus N2: -0.081
Minus N3: -0.269
Minus mobility level 1: 0
Minus self-care level 1: 0
Minus usual activities level 2: -0.036
Minus pain/discomfort level 2: -0.123
Minus anxiety/depression level 3: -0.236
where level 1 corresponds to no problems, level 2 to some problems and level 3 to extreme problems.

If any question is missing at an assessment, then the index will be missing.

13.1.4 DLQI

DLQI measures the health-related quality of life of adult patients suffering from a skin disease. It comprises 10 items arranged in 6 categories: symptoms and feelings (questions 1 and 2), daily activity (3 and 4), leisure (5 and 6), work or study (7), interpersonal relationships (8 and 9), and treatment (10). The total score can vary from 0 (no impact) to 30 (maximum impact). The DLQI score is calculated by summing the score of each question resulting in a maximum of 30 and a minimum of 0. The highest score represents the highest impairment of the health-related quality of life.

The scoring of each item is as per Table 10:

Table 10: Scoring of Individual DLQI Items

• Very much	Scored 3
• A lot	Scored 2
• A little	Scored 1
• Not at all	Scored 0
• Not relevant^	Scored 0
• Question unanswered or missing	Scored 0
• If leading Question 7 response is Yes*	Scored 3
• If leading Question 7 response is not Yes*	Scored as above using follow-up Question 7

* Questions 7 and 7a substitute each other. Only 1 of them may contribute to overall DLQI both in terms of score and number of questions. I.e. if Q7 = "Yes" and Q7a = "A little" – only Q7 is scored as "3" and +1 non missing-question, while Q7a is completely ignored.

^ if a question is answered as "Not relevant" then this should be considered a non-missing response, and is assigned a score of 0.

DLQI = sum(non-missing question score) / (number of non-missing questions) × 10. If the number of questions with a missing answer is ≥ 5 (≥ 50%) then DLQI is missing.

13.1.5 HADS

The HADS is a self-report rating scale of 14 items on a 4-point Likert scale (range 0–3, with 0 least severe and 3 most severe) (Zigmond AS, 1983). It is designed to measure anxiety and depression (7 items for each subscale). The total score is the sum of the 14 items, and for each subscale, the score is the sum of the respective 7 items (ranging from 0–21). HADS-Anxiety includes questions 1, 3, 5, 7, 9, 11, and 13. HADS-Depression includes questions 2, 4, 6, 8, 10, 12, and 14.

HADS will be present as anxiety, depression, and overall respectively. If the number of questions with a missing answer is ≥ 4, ≥ 4, and ≥ 7 (≥ 50%) for anxiety, depression, and overall respectively, then the score is missing.

13.1.6 International Hidradenitis Suppurativa Severity Score System

The IHS4 is a validated tool to dynamically assess HS severity to be used both in real-life and clinical study setting. The determination of IHS4 requires counting the nodules, abscesses, and draining fistulas/sinus tracts. The IHS4 score (points) = (number of nodules × 1) + (number of abscesses × 2) + (number of draining tunnels [fistulae/sinuses] × 4).

A score of ≤ 3 signifies mild HS, a score of 4 to 10 signifies moderate HS, and a score of ≥ 11 signifies severe (Zouboulis CC, 2017).

For the analysis of the IHS4 score change from baseline, if a location field is left blank for the count of abscess, nodules, or draining fistula, then the IHS4 score will be missing.

For the analysis of IHS4-55, if a location field is left blank for the count of abscess, nodules, or draining fistula, then the sum of all locations will be calculated assuming that the missing location has zero counts.

13.1.7 Hidradenitis Suppurativa Quality of Life (HiSQOL) Score

The 17-item HiSQOL questionnaire includes 4 symptom items, 8 activity-adaptation items, and 5 psychosocial items. The item scores are summed to create a total ranging from 0 to 68, with higher scores indicating a more severe impact on health-related quality of life (Kirby, 2020).

If two or more items are unanswered, the questionnaire is not scored, and the total score will be considered missing.

13.1.8 Hidradenitis Suppurativa – Physician Global Assessment (HS-PGA)

Although this value is collected directly on the eCRF, for analysis purposes it will be derived programmatically due to inconsistent and unreliable entry on the eCRF.

Derivation will be as follows:

- Clear: No inflammatory or non-inflammatory nodules
- Minimal: Only the presence of non-inflammatory nodules
- Mild: [< 5 inflammatory nodules and no abscess/draining fistula] or [(1 abscess or 1 draining fistula) and no inflammatory nodules]
- Moderate: [≥ 5 inflammatory nodules and no abscess/draining fistula] or [(1 abscess or 1 draining fistula) and ≥ 1 inflammatory nodules] or [(2 to 5 abscesses or 2-5 draining fistulas) and (< 10 inflammatory nodules)]
- Severe: [2 to 5 abscesses or 2 to 5 draining fistulas] and ≥ 10 or more inflammatory nodules
- Very severe: > 5 abscesses or draining fistulas

The derivation above follows Table 11 from (Daoud Mathieu, 2023):

Table 11: HS-PGA Scoring

	Abscesses + draining fistulas	Inflammatory nodules	Non-inflammatory nodules
Clear (1)	0	0	0
Minimal (2)	0	0	>0
Mild (3)	0	1 - 4	
	1	0	
Moderate (4)	0	≥ 5	
	1	≥ 1	
	2 – 5	< 10	
Severe (5)	2 - 5	≥ 10	
Very Severe (6)	> 5		

If the count of abscesses, draining fistulas, or non-inflammatory nodules is missing for a region, then the HS-PGA score will not be calculated.

13.1.9 Erythema Score

At every visit, for each anatomic region affected by HS, the investigator will assess the overall degree of erythema using a 4-point ordinal scale ranging between 0 and 3.

- Left axilla
- Right axilla
- Left sub/inframammary area
- Right sub/inframammary area
- Intermammary area
- Left buttock
- Right buttock
- Left inguino-crural fold
- Right inguino-crural fold
- Perianal
- Perineal
- Upper Back and Neck
- Other

The scoring rule as following:

- 0 = no redness
- 1 = faint but discernible pink coloration
- 2 = moderate red coloration
- 3 = very red or bright red coloration

Erythema score = 12 × {sum (non-missing score of each region [total 12 regions + Other regions])/ (number of non-missing region)}. The multiplier of 12 will not change even if the number of non-missing + Other regions changes. If the number of questions with a missing answer is ≥ 6 (≥ 50%) then the Erythema score is missing.

A region is considered "affected" if it has a > 0 abscess/nodule/fistula count or a non-zero response to any of the other quantitative questions on the Lesion Count/Hurley Stage eCRF form for the region or "intervening normal appearance tissue" = "No" or Hurley Stage not equal to "Not Applicable".

13.1.10 SF-12v2

The SF-12v2 (Ware, 2002) is a self-reported outcome measure assessing the impact of health on an individual's everyday life. Two component scores are derived from the SF-12 including the SF-12 Physical Component Score and the SF-12 Mental Component Score. SF-12 summary score can be calculated only when all components are non-missing (i.e. all answers provided), otherwise should be set to missing. If all answers are available then calculate as per following steps.

The instructions for computing the component scores are as follows:

1. Compute raw scale scores based on the table from precoded item value to final item value.
2. Transform raw scale scores to 0-100 scale.
3. Standardization of SF-12v2 scales (Z-scores), Standard Form
4. Norm-Based Transformation of SF-12v2 scales Z-scores, Standard Form
5. Norm-Based Scoring SF-12v2 Physical and Mental Summary Measures

Table 12: SF-12v2 Pre-coded and Final Values by Items:

Item: 1

Response choices	Precoded Item Value	Final Item Value
Excellent	1	5.0
Very good	2	4.4
Good	3	3.4
Fair	4	2.0
Poor	5	1.0

Items: 2a, 2b

Response choices	Precoded Item Value	Final Item Value
Yes, limited a lot	1	1
Yes, limited a little	2	2
No, not limited at all	3	3

Items: 3a, 3b, 4a, 4b, 7

Response choices	Precoded Item Value	Final Item Value
All of the time	1	1
Most of the time	2	2
Some of the time	3	3
A little of the time	4	4
None of the time	5	5

Item: 5

Response choices	Precoded Item Value	Final Item Value
Not at all	1	5
A little bit	2	4
Moderately	3	3
Quite a bit	4	2
Extremely	5	1

Items: 6a, 6b

Response choices	Precoded Item Value	Final Item Value
All of the time	1	5
Most of the time	2	4
Some of the time	3	3
A little of the time	4	2
None of the time	5	1

Items: 6c

Response choices	Precoded Item Value	Final Item Value
All of the time	1	1
Most of the time	2	2
Some of the time	3	3
A little of the time	4	4

None of the time	5	5

1. Transformation of scale scores:

The next step involves transforming each raw score to 0-100 scale using the formula shown below:

$$Transformed\ scale = \left[\frac{(Actual\ raw\ score - lowest\ possible\ raw\ score)}{Possible\ raw\ score\ range}\right] \times 100$$

This transformation converts the lowest and highest possible scores to zero and 100, respectively. Scores between these values represent the percentage of the total possible score achieved.

Scale items aggregated and range of possible scores:

SF-12v2 scale	Sum final Item Values (after recoding items)	Lowest and highest possible raw scores	Possible raw score range
Physical Functioning (PF)	Items 2a + 2b	2, 6	4
Role Physical (RP)	Items 3a + 3b	2, 10	8
Bodily Pain (BP)	Item 5	1, 5	4
General Health (GH)	Item 1	1, 5	4
Vitality (VT)	Item 6b	1, 5	4
Social Functioning (SF)	Item 7	1, 5	4
Role Emotional (RE)	Items 4a + 4b	2, 10	8
Mental Health (MH)	Items 6a + 6c	2, 10	8

Standardization of SF-12v2 scales (Z-scores), Standard Form:

Formulas for z-score standardization of SF-12v2 scales, Standard Form:

SF-12v2	Z-scores
Physical Functioning (PF_Z)	PF_Z = (PF – 81.18122) / 29.10558
Role Physical (RP_Z)	RP_Z = (RP – 80.52856) / 27.13526
Bodily Pain (BP_Z)	BP_Z = (BP – 81.74015) / 24.53019
General Health (GH_Z)	GH_Z = (GH – 72.19795) / 23.19041
Vitality (VT_Z)	VT_Z = (VT – 55.59090) / 24.84380
Social Functioning (SF_Z)	SF_Z = (SF – 83.73973) / 24.75775
Role Emotional (RE_Z)	RE_Z = (RE – 86.41051) / 22.35543
Mental Health (MH_Z)	MH_Z = (MH – 70.18217) / 20.50597

Norm-Based Transformation of SF-12v2 scales Z-scores, Standard Form:

SF-12v2	Norm-Based scores
Physical Functioning (Norm-Based PF)	Norm-Based PF = 50 + (PF_Z * 10)
Role Physical (Norm-Based RP)	Norm-Based RP = 50 + (RP_Z * 10)
Bodily Pain (Norm-Based BP)	Norm-Based BP = 50 + (BP_Z * 10)
General Health (Norm-Based GH)	Norm-Based GH = 50 + (GH_Z * 10)
Vitality (Norm-Based VT)	Norm-Based VT = 50 + (VT_Z * 10)

Social Functioning (Norm-Based SF)	Norm-Based SF = 50 + (SF_Z * 10)
Role Emotional (Norm-Based RE)	Norm-Based RE = 50 + (RE_Z * 10)
Mental Health (Norm-Based MH)	Norm-Based MH = 50 + (MH_Z * 10)

2. Norm-Based Scoring SF-12v2 Physical and Mental Summary Measures (Standard 4-week recall)

SF-12v2	Aggregated scores
Aggregate Physical Summary Score (AGG_PHYS)	AGG_PHYS = (PF_Z * 0.42402) + (RP_Z * 0.35119) + (BP_Z * 0.31754) + (GH_Z * 0.24954) + (VT_Z * 0.02877) + (SF_Z * -0.00753) + (RE_Z * -0.19206) + (MH_Z * -0.22069)
Aggregate Mental Summary Score (AGG_MENT)	AGG_MENT = (PF_Z * -0.22999) + (RP_Z * -0.12329) + (BP_Z * -0.09731) + (GH_Z * -0.01571) + (VT_Z * 0.23534) + (SF_Z * 0.26876) + (RE_Z * 0.43407) + (MH_Z * 0.48581)

3. Transformation of Summary Scores (Standard Form)

SF-12v2	Z-scores
Transformed Physical Component Score (PCS)	50 + (AGG_PHYS * 10)
Transformed Mental Component Score (MCS)	50 + (AGG_MENT * 10)