Official Title: A Phase 2, Double-Blind, Randomized, Vehicle Controlled, Efficacy, and Safety Study of Ruxolitinib Cream in Participants With Hidradenitis Suppurativa

NCT Number: NCT05635838

Document Date: Statistical Analysis Plan: 12 SEP 2023

Statistical Analysis Plan

INCB 18424-221

A Phase 2, Double-Blind, Randomized, Vehicle-Controlled, Efficacy, and Safety Study of Ruxolitinib Cream in Participants With Hidradenitis Suppurativa

IND Number:	**77,101**
Sponsor:	**Incyte Corporation** **1801 Augustine Cut-Off** **Wilmington, DE 19803** **United States**
Protocol Version:	**Protocol dated 19 AUG 2022**
CRF Approval Date:	**14 AUG 2023**
SAP Version:	**Original**
SAP Author:	[redacted] [redacted]**Biostatistician**
Date of Plan:	**12 SEP 2023**

This study is being conducted in compliance with Good Clinical Practice, including the archiving of essential documents.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Term
AE	adverse event
AN	abscess and inflammatory nodule
ASR	application site reaction
BID	twice daily
BMI	body mass index
BSA	body surface area
CRF	case report form
CTCAE	Common Terminology Criteria for Adverse Events
DBVC	double-blind, vehicle-controlled
[redacted]	[redacted]
eCRF	electronic case report form
[redacted]	[redacted]
EQ VAS	EuroQol visual analogue scale
HiSCR	hidradenitis suppurativa clinical response
[redacted]	[redacted]
HS	hidradenitis suppurativa
IHS4	International Hidradenitis Suppurativa Severity Score System
ITT	intent-to-treat
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	mixed model for repeated measures
NRS	numeric rating scale
OLE	open-label extension
[redacted]	[redacted]
[redacted]	[redacted]
PRO	patient-reported outcome
PT	preferred term
Q#	Question <number>
SAP	Statistical Analysis Plan
SI	International System of Units
SOC	system organ class
TEAE	treatment-emergent adverse event
ULN	upper limit of normal
WHO	World Health Organization
[redacted]	[redacted]

1. INTRODUCTION

INCB 18424-221 is a Phase 2, randomized, DBVC, efficacy, and safety study in participants aged ≥ 18 years with a diagnosis of HS. The study will consist of a 16-week DBVC period followed by a 16-week OLE period. A 30-day post-treatment safety follow-up visit will be conducted after the OLE period.

The purpose of this SAP is to provide details of the statistical analyses that have been outlined in the INCB 18424-221 Protocol.

2. STUDY INFORMATION, OBJECTIVES, AND ENDPOINTS

2.1. Protocol and Case Report Form Version

This SAP is based on INCB 18424-221 Protocol dated 19 AUG 2022 and CRFs approved 14 AUG 2023. Unless superseded by an amendment, this SAP will be effective for all subsequent Protocol amendments and eCRF versions.

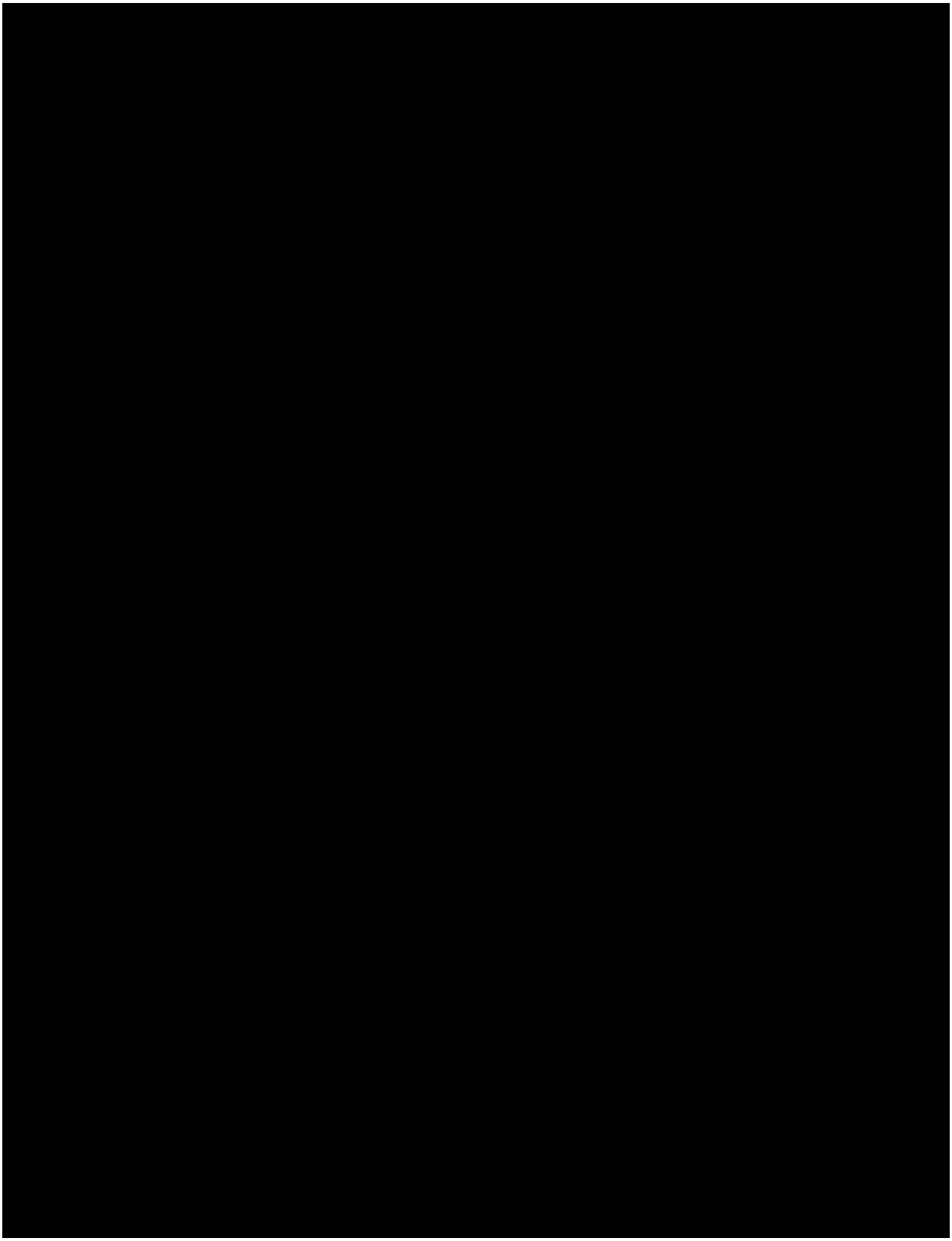
2.2. Study Objectives and Endpoints

Table 1 presents the objectives and endpoints.

Table 1: Objectives and Endpoints

Objectives	Endpoints
Primary	
To establish the efficacy of ruxolitinib 1.5% cream BID in participants with HS.	Change from baseline in AN count at Week 16.
Secondary	
To further evaluate the treatment effect of ruxolitinib 1.5% cream BID in participants with HS.	• Change from baseline to Week 16 in total AN count in anatomical areas with pre-existing AN at baseline. Note: Pre-existing AN at baseline are defined as abscesses and/or inflammatory nodules present at baseline. • Proportion of participants achieving AN50, AN75, AN90, and AN100 (at least 50%, 75%, 90%, and 100% reduction, respectively, in AN count relative to baseline) at Week 16. • Change from baseline in the Skin Pain NRS score at Week 16. • Change from baseline in the Itch NRS score at Week 16. • Proportion of participants who achieve HiSCR at Week 16. Note: HiSCR is defined as at least 50% reduction in AN count with no increase in either abscess or draining fistula counts, relative to baseline. • Change from baseline in the IHS4 score at Week 16. • Note: IHS4 score is calculated by the number of inflammatory nodules (multiplied by 1) plus the number of abscesses (multiplied by 2) plus the number of draining tunnels (multiplied by 4).
To evaluate the safety and tolerability of ruxolitinib 1.5% cream BID in participants with HS.	The type, frequency, and severity of AEs, and changes in vital signs and hematology and serum chemistry parameters.

Table 1: Objectives and Endpoints (Continued)



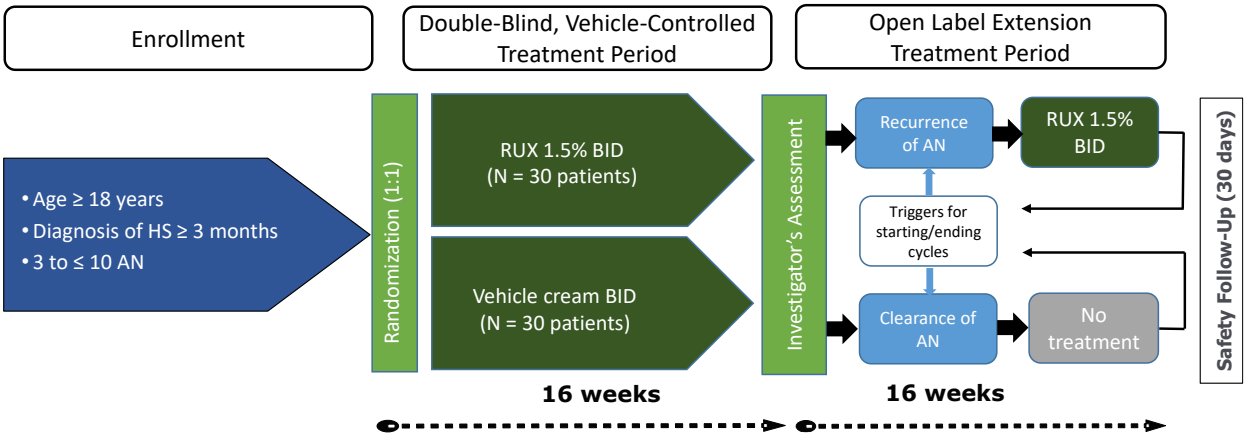
3. STUDY DESIGN

This is a Phase 2, randomized, DBVC study in participants aged ≥ 18 years with confirmed diagnosis of HS. The study will consist of a 16-week DBVC period followed by a 16-week OLE period.

Participants will be screened for up to 28 days prior to the first application of ruxolitinib 1.5% cream or vehicle cream. Approximately 60 eligible participants aged ≥ 18 years will be randomized 1:1 to either ruxolitinib 1.5% cream or vehicle cream. Participants will be stratified by baseline AN count (ie, ≥ 3 to 4 or ≥ 5 to 10). Participants will apply either ruxolitinib 1.5% cream BID or vehicle cream BID through Week 16 to all affected areas identified at baseline. At Week 16, participants who meet the criteria (compliant with the Protocol and without safety concerns) will enter the 16-week OLE period. Participants randomized to vehicle BID in the DBVC period will be crossed over to ruxolitinib 1.5% cream BID and participants randomized to ruxolitinib 1.5% cream BID at baseline will remain on ruxolitinib 1.5% cream BID through Week 32 in an open-label fashion. Participants will enter the 30-day safety follow-up period after the last application of study treatment.

The study schema is shown below in Figure 1.

Figure 1: Study Design Schema



Abbreviations: Rux = ruxolitinib.

The primary analysis will occur after the primary database lock, when all participants have completed or discontinued from the DBVC period. The final analysis will occur when all participants have completed or withdrawn from the study, whichever comes first.

3.1. Randomization

Approximately 60 participants will be randomized 1:1 to either ruxolitinib 1.5% cream BID or vehicle cream. Participants will be stratified by baseline AN count (ie, ≥ 3 to 4 or ≥ 5 to 10).

3.2. Control of Type I Error

The significance level for primary efficacy analysis will be 0.1 for a 2-sided test.

3.3. Sample Size Considerations

Approximately 60 participants will be randomized 1:1 to ruxolitinib 1.5% cream BID or vehicle BID. The sample size was not calculated based on statistical power calculations, but for demonstration of preliminary findings of clinical response. It is anticipated that a sample size of approximately 60 participants will 1) generate sufficient data to assess whether ruxolitinib 1.5% cream warrants further investigation in HS, and 2) be sufficient to provide enough data for an initial evaluation of the safety profile of ruxolitinib 1.5% cream in HS.

3.4. Schedule of Assessments

Refer to Protocol dated 19 AUG 2022 for a full description of all study procedures and assessment schedules for this study.

4. DATA HANDLING DEFINITIONS AND CONVENTIONS

4.1. Scheduled Study Evaluations and Study Periods

4.1.1. Day 1

Day 1 is the date that the first application of ruxolitinib 1.5% cream or vehicle cream is administered to the participants in the specific period.

For randomized participants not treated with any study drug, Day 1 is defined as the date of randomization.

[redacted]

4.1.3. Baseline Value

For the DBVC period, baseline is the last nonmissing measurement obtained before or on the day of the first application of ruxolitinib 1.5% cream or vehicle cream in the DBVC period, unless otherwise defined.

For randomized participants not treated with any study drug, baseline is defined as the last nonmissing assessment before or on the date of randomization.

For participants who continue in the OLE period, baseline is defined as follows:

- For efficacy evaluation in the OLE period, baseline is the last nonmissing measurement obtained before or on the day of first application of study treatment in the DBVC period.
- For safety evaluation in the OLE period, baseline is the last nonmissing measurement obtained before or on the day of first application of ruxolitinib 1.5% cream in the study.

When a scheduled assessment and an unscheduled assessment occur on the same day and the times of the assessments are not available, use the following convention to determine baseline:

- If both a scheduled and an unscheduled visit are available on the day of the first application and the time is missing, use the scheduled assessment as baseline.
- If the scheduled assessment is missing on the day of the first application and an unscheduled assessment is available, use the unscheduled assessment as baseline.

4.1.4. Handling of Missing and Incomplete Dates

In general, values for missing dates will not be handled unless methods for handling missing dates are specified in this section or relevant sections. The original reported dates collected on the eCRF should be used in all relevant listings. The following rules will be used for handling partial dates for analyses requiring dates.

Partial HS diagnosis date will be handled, as follows, in the calculation:

- If only the day is missing, then the first day of the month will be used.
- If both the month and day are missing, then 01 JAN of the year will be used.
- If the diagnosis date is completely missing, then the time since diagnosis will not be calculated.

When the date of the last application is used in deriving variables, such as duration of treatment or TEAE flag, a missing date of the last application in the specified treatment period will be handled as follows:

- If the last application date is missing, then the end date of the specified treatment period will be used as the date of the last application.
- If the last application date and the end date of the specified treatment period are both missing and the participant is lost to follow-up, then the date of last contact will be used as the date of the last application.

4.2. Variable Definitions

4.2.1. Body Mass Index

Body mass index will be calculated as follows:

BMI (kg/m^2) = [weight (kg)] / [height (m)]^2.

4.2.2. Prior and Concomitant Medications

A prior medication is defined as any nonstudy medication started before the first application of ruxolitinib 1.5% cream or vehicle cream.

A concomitant medication is defined as any nonstudy medication that is started accordingly:

- Before the date of first application of ruxolitinib 1.5% cream or vehicle cream and is ongoing throughout the study or ends on/after the date of first study drug application.
- On/after the date of first application of ruxolitinib 1.5% cream or vehicle cream and is ongoing or ends during the course of the study.

A prior medication could also be classified as "both prior and concomitant medication" if the end date is on or after first application of ruxolitinib 1.5% cream or vehicle cream. In the listing, it will be indicated whether a medication is only prior, only concomitant, or both prior and concomitant.

For the purposes of analysis, all medications will be considered concomitant medications unless the medications can unequivocally be defined as not concomitant.

5. STATISTICAL METHODOLOGY

5.1. General Methodology

Unless otherwise noted, SAS® software (SAS Institute Inc, Cary, NC; v9.4 or later) will be used for the generation of all tables, graphs, and statistical analyses. Descriptive summaries for continuous variables will include, but not be limited to, the number of observations, mean, standard deviation, median, minimum, and maximum. Descriptive summaries for categorical variables will include the number and percentage of participants in each category.

5.2. Treatment Groups

This is a randomized study with DBVC and OLE periods.

Efficacy data will be presented by treatment group assigned at study entry for the ITT population. The treatment groups will be 1.5% BID (ie, ruxolitinib 1.5% cream BID) and vehicle BID.

Safety data will be presented by treatment group that participants actually applied.

- For safety data in the DBVC period, the treatment groups will be 1.5% BID and vehicle BID.
- For safety data in the OLE period, the treatment groups will be 1.5% BID to 1.5% BID and vehicle BID to 1.5% BID.
- For safety data throughout the study, the only treatment group will be 1.5% BID, including participants who applied ruxolitinib 1.5% cream BID from Day 1 through Week 32 or end of study treatment.

5.3. Analysis Populations

5.3.1. Intent-to-Treat Population

All participants who are randomized will constitute the ITT population. Treatment groups for this population will be defined according to the treatment assignment at the time of randomization regardless of the actual study drug the participant might take during their participation in the DBVC period.

The ITT population will be used for the summary of demographics, baseline characteristics, participant disposition, medical history, prior and concomitant medications, and analyses of all efficacy data, unless otherwise specified.

5.3.2. Safety Population

The safety population will include all participants who applied ruxolitinib 1.5% cream or vehicle cream at least once. Treatment groups for this population will be determined according to the actual treatment the participant applied on Day 1 regardless of assigned treatment group.

The safety population will be used for safety summaries in the DBVC period.

5.3.2.1. Open-Label Extension Safety Population

The OLE safety population includes all participants who applied ruxolitinib 1.5% cream BID at least once during the OLE period.

All safety analyses conducted specifically for the OLE period will use the OLE safety population.

[redacted]

6. BASELINE, EXPOSURE, AND DISPOSITION

Appendix A provides a list of planned tables, figures, and listings. Sample data displays are included in a separate document.

6.1. Demographics, Baseline Characteristics, and Disease History

6.1.1. Demographics and Baseline Characteristics

The following demographics and baseline characteristics will be summarized by treatment group and overall for the ITT population in the DBVC period and for the OLE safety population in the OLE period: age, sex, race, ethnicity, weight, height, and BMI.

6.1.2. Baseline Disease Characteristics

For the ITT population, the baseline disease characteristics will be summarized and analyzed by treatment group and will include, but not be limited to the following:

- Disease duration (months)
- Prior surgical treatment for HS (yes [incision and drainage, deroofing, or other]/no)
- Prior therapy received for HS
- HS family history (yes/no)
- Selected comorbidities
- Number of anatomical areas with HS
- Days since onset of current episode of HS flare/activity
- Number of HS flares since diagnosis
- Hurley stage
- AN count category (≥ 3 to 4 or ≥ 5 to 10)
- AN count
- Inflammatory nodule count
- Abscess count
- Nondraining tunnel count
- Draining tunnel count
- Skin Pain NRS score
- Itch NRS score
- Total BSA involvement

6.1.3. Disease History

The time since diagnosis will be summarized for all participants in the ITT population.

Time since diagnosis will be calculated as follows:

Disease duration (months) = (date of randomization – date of initial HS diagnosis +1) / 30.4375.

6.1.4. Medical History

For participants in the ITT population during the DBVC period, medical history will be summarized by assigned treatment groups. This summary will include the number and percentage of participants with medical history for each body system/organ class as documented on the eCRF.

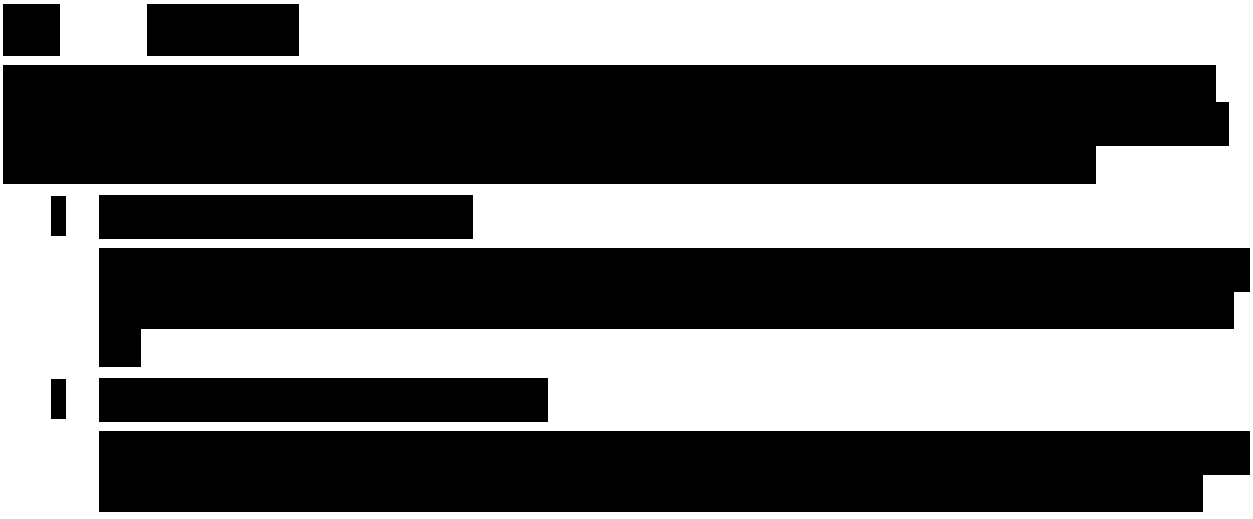
6.2. Disposition of Participant

The number and percentage of participants who were randomized, treated, and completed the DBVC period, as well as the number and percentage of participants who discontinued the treatment or withdrew from the study during the DBVC period with a primary reason for discontinuation, will be summarized for the ITT population.

The number and percentage of participants who completed the OLE period, in addition to those participants who discontinued treatment or withdrew from the study during the OLE period with a primary reason for discontinuation, will be summarized for the OLE safety population.

6.3. Protocol Deviations

For participants in the ITT population in the DBVC period, as well as the OLE safety population in the OLE period, Protocol deviations recorded on the eCRF will be summarized and listed by category (critical, major, or minor) for each treatment group and overall.



[redacted]

6.6. Prior and Concomitant Medications

Prior medications and concomitant medications will be coded using the WHO Drug Dictionary. The number and percentage of participants with prior and concomitant medications will be summarized by treatment group and overall for the ITT population during the DBVC period by WHO drug class and the WHO drug preferred term. For the OLE period, only concomitant medications will be summarized based on OLE safety population.

7. EFFICACY

Appendix A provides a list of data displays. Sample data displays are included in a separate document.

7.1. Efficacy Hypothesis

The primary hypothesis is that ruxolitinib 1.5% cream BID, compared with vehicle cream BID, will improve change in AN count from baseline to Week 16 in participants with HS.

7.2. Efficacy Measures

7.2.1. Lesion Counts

The HS lesion counts will be assessed at each study visit by the investigator and used for calculation of efficacy parameters as follows: AN counts, HiSCR, IHS4, and flare incidence.

Lesion count and assessment will be recorded in the lesion count worksheet; the assessment of anatomical regions includes, but is not limited to the left and right axilla and the left and right inguinocrural fold or inframammary areas.

7.2.1.1. Abscess and Inflammatory Nodule Counts

The AN counts will be recorded at all visits. The AN results will be used to calculate change in AN count relative to baseline, as well as AN50, AN75, AN90, and AN100, which are defined as, at least, a 50%, 75%, 90%, and 100% decrease, respectively, in AN count relative to baseline.

7.2.1.2. Hidradenitis Suppurativa Clinical Response

The HiSCR was originally developed based on the underlying Phase 2 trial of adalimumab and validated against meaningful changes in pain score and DLQI (Kimball et al 2016, Sabat et al 2020). The achievement of HiSCR is defined as at least 50% reduction in AN count with no increase in either abscess or draining fistula counts, relative to baseline. In this study, participants with draining fistulas (tunnels) will be excluded from the study. Should a randomized participant develop a draining tunnel during the study, the participant will be discontinued from the study.

7.2.1.3. International Hidradenitis Suppurativa Severity Score System

The IHS4 (Zouboulis et al 2017) is a composite, dynamic score, and validated tool, used to determine HS severity. It employs a weighted scale using the number of inflammatory nodules, the number of abscesses, and the number of draining tunnels (eg, fistulas or sinuses), with respective weight factors of 1, 2, and 4. For example, IHS4 score equals the number of inflammatory nodules (multiplied by 1) plus the number of abscesses (multiplied by 2) plus the number of draining tunnels (multiplied by 4).

[redacted]

7.2.2. Hurley Stages of Hidradenitis Suppurativa

The Hurley classification is a static score and was originally designed for selection of the appropriate treatment modality in a certain body location (Zouboulis et al 2017): medical therapy for Stage I, local surgery for Stage II, and wide surgical excision for Stage III (see Table 2). Participants who have been diagnosed with HS, Hurley Stage I or II, as per inclusion criterion, will be enrolled into the study. The investigator (or designee) will determine the Global Hurley Stage in each affected anatomical region at the designated study visits. For each participant, an overall Hurley Stage will be derived as the stage of the worst involved anatomical region. The number and proportion of participants in each Hurley Stage will be summarized at baseline and each postbaseline visit.

Table 2: Hurley Stages of Hidradenitis Suppurativa

Hurley Stage	Description
I	Abscess formation, single or multiple, without sinus tracts and cicatrization/scarring.
II	One or more widely separated recurrent abscesses with tract formation and cicatrization/scarring.
III	Multiple interconnected tracts and abscesses across the entire area, with diffuse or near diffuse involvement.

7.2.3. Body Surface Area

Total %BSA affected by HS will be used to determine the number of tubes of study drug dispensed at each visit. Total %BSA affected will be estimated at each visit as outlined in the Protocol. The BSA assessment will be approximated to the nearest 0.1% using the Palmar Method as a guide, with the palm plus 5 digits, with fingers tucked together and thumb tucked to the side (handprint), considered as 1% BSA and the thumb as 0.1% BSA.

7.2.4. Patient-Reported Outcomes

Patient-reported outcomes will be assessed at designated visits as outlined in the Protocol.

For those PROs scheduled to be completed at study visits, on Day 1, the participant will complete their PRO after eligibility is met. For subsequent visits after randomization, questionnaires will be completed by the participant before site personnel perform any clinical assessments to prevent participant bias in their response.

7.2.4.1. Skin Pain Numerical Rating Scale

Participants will be instructed to complete and record the Skin Pain NRS in a diary each evening, beginning on the day of screening through Week 32 or treatment discontinuation, whichever comes first. The participants will rate the severity of their pain associated with HS (ie, burning,

tearing, pulling, stabbing, etc) by selecting a number from 0 (no pain) to 10 (worst imaginable pain) that best describes their worst level of pain in the past 24 hours.

The Skin Pain NRS score for baseline will be determined by averaging the 7 daily NRS scores before Day 1 (ie, Days –7 to –1) for all by-visit summaries. The by-visit Skin Pain NRS score for postbaseline visits will be determined by averaging the 7 daily NRS scores before the visit day. If 4 or more daily scores are missing (out of the 7), the Skin Pain NRS score at the visit will be set to missing. For all the daily pain-related analyses, baseline will be defined as the last available daily pain NRS score during the last week prior to Day 1 (ie, Days −7 to −1).

7.2.4.2. Itch Numerical Rating Scale

The participant will be instructed to complete and record the Itch NRS in a diary each evening beginning on the day of screening through Week 32 or treatment discontinuation, whichever comes first. The participant will rate the severity of their itch associated with HS by selecting a number from 0 (no itch) to 10 (worst imaginable itch) that best describes the worst level of itching in the past 24 hours.

The Itch NRS score for baseline will be determined by averaging the 7 daily NRS scores before Day 1 (ie, Days –7 to –1) for all by-visit summaries. The by-visit Itch NRS score for postbaseline visits will be determined by averaging the 7 daily NRS scores before the visit day. If 4 or more daily scores are missing (out of the 7), the Itch NRS score at the visit will be set to missing. For all the daily itch-related analyses, baseline will be defined as the last available daily Itch NRS score during the last week prior to Day 1 (ie, Days −7 to −1).

[redacted]

[redacted]

[redacted]

7.3. Analysis of the Primary Efficacy Parameter

7.3.1. Primary Efficacy Analysis

The primary analysis will be based on the ITT population. The summary of primary endpoint analysis is provided in Table 3.

Table 3: Summary of Primary Analysis

Parameter	Definition
Treatment effect	Ruxolitinib 1.5% cream compared with vehicle cream
Population	ITT population
Variable	Change from baseline in AN count at Week 16
Population-level summary	Mean change from baseline in AN count at Week 16

Note: All randomized participants, including those with missing Week 16 data, and those who discontinue study treatment at any time before the timepoint of interest, or discontinue from the study for any reason, will be included for the primary analysis. No rescue therapy or treatment switch is allowed in this study.

The primary efficacy analysis will compare ruxolitinib 1.5% cream versus vehicle cream in the mean change from baseline in AN count using an MMRM. The MMRM will include the fixed effect of the treatment group (ruxolitinib 1.5% and vehicle cream), stratification factor (AN count $\geq$ 3 to 4 or AN count $\geq$ 5 to 10 at baseline), visit, and visit by treatment interaction. The variance-covariance matrix of the within-participant errors in MMRM will be modeled as unstructured.

The primary alternative hypothesis (superiority of active ruxolitinib 1.5% cream group compared with vehicle cream) will be tested at a 2-sided $\alpha = 0.1$ level using the least square mean estimate of the change from baseline in AN count at Week 16 from the aforementioned MMRM specified.

7.3.2. Subgroup Analyses for Primary Endpoint

Subgroups will be formed based on the following participant characteristics and baseline variables for those participants whose data are available:

- Baseline AN count: ≥ 3 to 4 or ≥ 5 to 10
- Age group: 18 to 29 years or ≥ 30 years

The primary efficacy endpoint will be summarized and analyzed using descriptive statistics based on the ITT population for the aforementioned subgroups.

7.4. Analysis of the Secondary Efficacy Parameters

All secondary efficacy analyses will be conducted for the ITT population.

7.4.1. Continuous Efficacy Endpoints

The summary and analysis of statistics for the following continuous measurements, including change from baseline, will be presented for each treatment group:

- Total AN count in anatomical areas with pre-existing AN at Week 16
- By-visit Skin Pain NRS score at Week 16
- By-visit Itch NRS score at Week 16
- IHS4 score at Week 16

The summary statistics include sample size, mean, median, standard deviation, minimum, maximum, first quartile, third quartile, and 95% confidence interval. A similar MMRM, as specified in the primary efficacy analysis, may be fit for the comparison between ruxolitinib 1.5% cream and vehicle cream at Week 16.

For the endpoint change from baseline in total AN count in anatomical areas with pre-existing AN (ie, not including new AN identified after baseline), only the anatomical areas with an AN count > 0 will be considered for each participant. At baseline, the anatomical areas with abscesses and/or inflammatory nodules will be identified and AN counts in these areas will be summed. At each postbaseline visit, the sum of AN counts in only the pre-existing areas (that were identified at baseline) will be used to calculate the change relative to baseline.

7.4.2. Categorical Efficacy Endpoints

For the following categorical parameters, summary statistics, including sample size, frequency, and percentages will be presented.

- Proportion of participants achieving AN50/75/90/100 at Week 16
- Proportion of participants achieving HiSCR at Week 16

[redacted]

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

[redacted]

- [redacted]
- [redacted]

[redacted]

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

9. SAFETY AND TOLERABILITY

Appendix A provides a list of data displays. Sample data displays are included in a separate document.

9.1. General Considerations

The analyses in this section will be provided for the safety population in the DBVC period, OLE safety population in the OLE period, and those who applied ruxolitinib 1.5% cream throughout the study, unless otherwise specified. Summary tables may be replaced with listings when appropriate. For instance, an AE frequency table may be replaced with a listing if it only contains a few unique preferred terms reported on relatively few participants.

9.2. Adverse Events

9.2.1. Adverse Event Definitions

A TEAE is any AE either reported for the first time or worsening of a pre-existing event after first application of study drug and within 30 days of the last application of study drug. For participants who cross over treatments, the first application date is period-specific; the end date is 30 days after the last application date in this period, or the first application date in the next period, whichever comes first. Analysis of AEs (as discussed below) will be limited to TEAEs, but data listings will include all AEs regardless of their timing in relation to study drug application.

Adverse events will be tabulated by MedDRA PT and SOC. Severity of AEs will be graded using the National Cancer Institute CTCAE v5.0. The CTCAE reporting guidelines and grading details are available on the Cancer Therapy Evaluation Program website.

The subset of AEs considered by the investigator to be related to study drug will be considered to be treatment-related AEs. If the investigator does not specify the relationship of the AE to study drug, the AE will be considered to be treatment-related. The incidence of AEs and treatment-related AEs will be tabulated. In addition, serious TEAEs will also be tabulated.

Any missing onset date, causality, or severity must be queried for resolution. Unresolved missing causality and severity will be handled according to the following rules:

- An unresolved missing causality will be considered treatment-related.
- An unresolved missing severity will be identified as an unknown severity.

For purposes of analysis, all AEs will be considered TEAEs unless the AE can unequivocally be defined as not treatment-emergent.

9.2.2. Adverse Event Summaries

An overall summary of AEs by treatment group will include the following:

- Number (%) of participants reporting any TEAEs
- Number (%) of participants reporting any serious TEAEs
- Number (%) of participants reporting any Grade 3 or higher TEAEs

- Number (%) of participants reporting any treatment-related TEAEs
- Number(%) of participants reporting any ASRs
- Number (%) of participants who temporarily interrupted study treatment because of TEAEs
- Number (%) of participants who permanently discontinued study treatment because of TEAEs
- Number (%) of participants who had any fatal TEAEs

The following summaries will be produced by MedDRA term (if 2 or fewer participants appear in a table, a listing may be appropriate):

- Summary of TEAEs by MedDRA SOC and PT
- Summary of TEAEs by MedDRA PT in decreasing order of frequency
- Summary of TEAEs by MedDRA SOC, PT, and maximum severity
- Summary of TEAEs by MedDRA SOC, PT, and CTCAE grade category
- Summary of Grade 3 or higher TEAEs by MedDRA SOC and PT
- Summary of Grade 3 or higher TEAEs by MedDRA PT in decreasing order of frequency
- Summary of serious TEAEs by MedDRA SOC and PT
- Summary of serious TEAEs by MedDRA PT in decreasing order of frequency
- Summary of treatment-related TEAEs by MedDRA SOC and PT
- Summary of treatment-related TEAEs by MedDRA PT in decreasing order of frequency
- Summary of Grade 3 or higher treatment-related TEAEs by MedDRA SOC and PT
- Summary of treatment-related serious TEAEs by MedDRA SOC and PT
- Summary of TEAEs with a fatal outcome by MedDRA SOC and PT
- Summary of TEAEs leading to dose interruption by MedDRA SOC and PT
- Summary of TEAEs leading to discontinuation of study drug by MedDRA SOC and PT
- Summary of ASR by MedDRA PT in decreasing order of frequency

9.3. Clinical Laboratory Tests

The analyses in this section will be provided for the safety population in the DBVC period, as well as the OLE safety population in the OLE period, unless otherwise specified.

9.3.1. Laboratory Value Definitions

All laboratory assessments will be performed using a central laboratory, with the exception of urine pregnancy tests (as applicable). Laboratory values and change from baseline values will be summarized descriptively by visit; non-numeric test values will be tabulated when necessary. Baseline will be determined according to Section 4.1.3, using the last nonmissing value collected before the first application, prioritizing scheduled assessments for baseline identification over unscheduled visits. The last record before application in the highest priority will be considered the baseline record. For baseline laboratory candidates with the same date and time in the same priority category, additional rules may be provided after consultation with the medical monitor to delineate which value will be defined as baseline.

9.3.2. Laboratory Value Summaries

All test results and associated normal ranges from central laboratories will be reported in SI units. All tests with numeric values will have a unique unit per test. Any laboratory test results and associated normal ranges from local laboratories will be converted to SI units.

When there are multiple laboratory nonmissing values for a participant's particular test at a scheduled visit, central laboratory values have higher priority over local laboratory values. If a tie still exists, the laboratory value with the smallest laboratory sequence number will be used in by visit summaries.

For test results that will be summarized with available normal ranges, the number and percentage of participants with the laboratory values being low (but never high), normal, high (but never low) and both low and high will be calculated for each test. A shift summary will be produced for each test. The denominator for the percentage calculation will use the number of participants in the baseline category (ie, low, high, normal, or missing) as the denominator for the percentage in each of the categories during the study.

Severity grades will be assigned to laboratory test values based on the numerical component of CTCAE v5.0. Shift tables will also be presented showing change in CTCAE grade from baseline to the worst grade postbaseline. Separate summaries for abnormally high and abnormally low laboratory values will be provided when the laboratory parameter has both high and low grading criteria. The denominator for the percentage calculation will be the number of participants in the baseline category. The number of participants who experienced worsening of laboratory abnormalities will be summarized by maximum severity.

9.4. Vital Signs

The analyses in this section will be provided for the safety population in the DBVC period, as well as the OLE safety population in the OLE period, unless otherwise specified. Values at each scheduled visit, change, and percentage change from baseline for vital signs, including systolic blood pressure, diastolic blood pressure, pulse, temperature, and respiratory rate will be summarized descriptively.

Normal ranges for vital sign values are defined in Table 4. For participants exhibiting vital sign abnormalities, the abnormal values will be listed along with their assigned treatment group. Alert vital signs are defined as an absolute value outside the defined normal range and percentage change greater than 25%. The abnormal values for participants exhibiting alert vital sign abnormalities will be listed.

Table 4: Criteria for Clinically Notable Vital Sign Abnormalities

Parameter	High Threshold	Low Threshold
Systolic blood pressure	> 155 mmHg	< 85 mmHg
Diastolic blood pressure	> 100 mmHg	< 40 mmHg
Pulse	> 100 bpm	< 45 bpm
Temperature	> 38°C	< 35.5°C
Respiratory rate	> 20 breaths/min	< 8 breaths/min

10. PLANNED ANALYSES

No formal interim analysis is planned in this study.

There are 2 formal planned analyses:

- The primary analysis will occur after the primary database lock, when all participants have completed or discontinued from the DBVC period. The sponsor will be unblinded after the primary database lock; however, investigators and participants will remain blinded to the individual study treatment assignment after the primary database lock.
- The final analysis will occur when all participants have completed or withdrawn from the study.

11. CHANGES AND MODIFICATIONS TO THE ANALYSIS PLAN

All versions of the SAP are listed in Table 5.

Table 5: Statistical Analysis Plan Versions

SAP Version	Date
Original	12 SEP 2023

11.1. Changes to Protocol-Defined Analyses

1. **Section 2.2, Study Objectives and Endpoints**

   Description of change: Added the secondary endpoint, "Change from baseline to Week 16 in total AN count in anatomical areas with pre-existing AN at baseline."

   Rationale for change: To further evaluate the treatment effect of ruxolitinib 1.5% cream BID in participants with HS.

[redacted]

11.2. Changes to the Statistical Analysis Plan

Not applicable.

12. REFERENCES

[redacted]

Hurst H, Bolton J. Assessing the clinical significance of change scores recorded on subjective outcome measures. J Manipulative Physiol Ther 2004;27:26-35.

Kimball AB, Naegeli AN, Edson-Heredia E, et al. Psychometric properties of the Itch Numeric Rating Scale in patients with moderate-to-severe plaque psoriasis. Br J Dermatol 2016;175:157-162.

[redacted]

Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. Pharmacoeconomics 1993;4:353-365.

Sabat R, Jemec GBE, Lukasz M, Kimball AB, et al. Hidradenitis suppurativa. Nat Rev Dis Prim 2020;6:1-20.

Zouboulis CC, Tzellos T, Kyrgidis A, et al. Development and validation of the International Hidradenitis Suppurativa Severity Score System (IHS4), a novel dynamic scoring system to assess HS severity. Br J Dermatol 2017;177:1401-1409.

APPENDIX A. PLANNED TABLES, FIGURES, AND LISTINGS

This appendix provides a list of the planned tables, figures, and listings for the Clinical Study Report. Shells are provided in a separate document for tables that are not in the most current Standard Safety Tables v1.13.

The lists of tables, figures, and listings are to be used as guidelines. Modifications of the lists that do not otherwise affect the nature of the analysis will not warrant an amendment to the SAP.

Tables

Table No.	Title	Population	Standard
Baseline and Demographics Characteristics			
1.1 Disposition			
1.1.1	Analysis Populations	All randomized	X
1.1.2.1	Summary of Participant Disposition in the DBVC Period	ITT	X
1.1.2.2	Summary of Participant Disposition in the OLE Period	OLE	X
1.1.3	Summary of Number of Participants Enrolled by Country and Site	ITT	X
1.1.4.1	Summary of Protocol Deviations in the DBVC Period	ITT	X
1.1.4.2	Summary of Protocol Deviations in the OLE Period	OLE	X
1.2 Demographics and Baseline Characteristics			
1.2.1	Summary of Demographics and Baseline Characteristics	ITT	X
1.2.2	Summary of Demographics and Baseline Characteristics	OLE	X
1.3 Baseline Disease Characteristics			
1.3	Summary of Baseline Disease Characteristics	ITT	X
1.4 Prior and Concomitant Medications			
1.4.1	Summary of Prior Medications	ITT	X
1.4.2.1	Summary of Concomitant Medications in the DBVC Period	ITT	X
1.4.2.2	Summary of Concomitant Medications in the OLE Period	OLE	X
1.5 Others			
1.5	Summary of General Medical History	ITT	X
Efficacy			
2.1 AN Count			
2.1.1	Summary and Analysis of AN Count During the Study Period	ITT	
2.1.1.2	Summary and Analysis of AN Count by Baseline AN Count During the Study Period	ITT	
2.1.1.3	Summary and Analysis of AN Count by Age Group During the Study Period	ITT	
2.1.2	Summary and Analysis of AN Count in Anatomical Areas With Pre-Existing AN at Baseline During the Study Period	ITT	
2.1.3	Summary and Analysis of Participants Achieving AN50 During the Study Period	ITT	
2.1.4	Summary and Analysis of Participants Achieving AN75 During the Study Period	ITT	
2.1.5	Summary and Analysis of Participants Achieving AN90 During the Study Period	ITT	

Table No.	Title	Population	Standard
2.1.6	Summary and Analysis of Participants Achieving AN100 During the Study Period	ITT	
2.2 Skin Pain NRS			
2.2.1	Summary and Analysis of By-Visit Skin Pain NRS Score During the Study Period	ITT	
2.2.2	Summary and Analysis of Daily Skin Pain NRS Score From Baseline to Day 28	ITT	
2.2.3	Summary of Participants Achieving ≥ 2-Point Reduction From Baseline in By-Visit Skin Pain NRS Score During the Study Period	ITT (baseline Skin Pain NRS score ≥ 2)	
2.2.4	Summary of Participants Achieving ≥ 4-Point Reduction From Baseline in By-Visit Skin Pain NRS Score During the Study Period	ITT (baseline Skin Pain NRS score ≥ 4)	
2.3 Itch NRS			
2.3.1	Summary and Analysis of By-Visit Itch NRS Score During the Study Period	ITT	
2.3.2	Summary and Analysis of Daily Itch NRS Score From Baseline to Day 28	ITT	
2.3.3	Summary of Participants Achieving ≥ 4-Point Reduction From Baseline in By-Visit Itch NRS Score During the Study Period	ITT (baseline Itch NRS score ≥ 4)	
2.4 HiSCR			
2.4.1	Summary and Analysis of Participants Achieving HiSCR During the Study Period	ITT	
2.5 IHS4			
2.5.1	Summary and Analysis of IHS4 During the Study Period	ITT	
[redacted]			

Table No.	Title	Population	Standard
[redacted]			
2.13 Hurley Stage			
2.13.1	Summary of Participants in Each Hurley Stage During the Study Period	ITT	
Safety			
3.1 Exposure and Compliance			
3.1.1.1	Summary of Exposure in the DBVC Period	Safety	X
3.1.1.2	Summary of Exposure in the OLE Period	OLE	X
3.1.1.3	Summary of Exposure During Study Period	Safety	X
3.1.2.1	Summary of Study Drug Compliance in the DBVC Period	Safety	
3.2 Adverse Events			
3.2.1.X	Overall Summary of Treatment-Emergent Adverse Events	Safety	X
3.2.2.X	Summary of Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term	Safety	X
3.2.3.X	Summary of Treatment-Emergent Adverse Events by MedDRA Preferred Term in Decreasing Order of Frequency	Safety	X
3.2.4.X	Summary of Treatment-Emergent Adverse Events by MedDRA System Organ Class, Preferred Term, and Maximum Severity	Safety	X
3.2.5.X	Summary of Treatment-Emergent Adverse Events by MedDRA System Organ Class, Preferred Term, and CTCAE Grade Category	Safety	
3.2.6.X	Summary of Grade 3 or Higher Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term	Safety	X
3.2.7.X	Summary of Grade 3 or Higher Treatment-Emergent Adverse Events by MedDRA Preferred Term in Decreasing Order of Frequency	Safety	X
3.2.8.X	Summary of Serious Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term	Safety	X
3.2.9.X	Summary of Serious Treatment-Emergent Adverse Events by MedDRA Preferred Term in Decreasing Order of Frequency	Safety	X
3.2.10.X	Summary of Treatment-Related Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term	Safety	X
3.2.11.X	Summary of Treatment-Related Treatment-Emergent Adverse Events by MedDRA Preferred Term in Decreasing Order of Frequency	Safety	X
3.2.12.X	Summary of Grade 3 or Higher Treatment-Related Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term	Safety	
3.2.13.X	Summary of Treatment-Related Serious Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term	Safety	
3.2.14.X	Summary of Treatment-Emergent Adverse Events With a Fatal Outcome by MedDRA System Organ Class and Preferred Term	Safety	
3.2.16.X	Summary of Treatment-Emergent Adverse Events Leading to Dose Interruption in Some or All Treated Areas by MedDRA System Organ Class and Preferred Term	Safety	X

Table No.	Title	Population	Standard
3.2.17.X	Summary of Treatment-Emergent Adverse Events Leading to Dose Interruption in All Treated Areas by MedDRA System Organ Class and Preferred Term	Safety	
3.2.18.X	Summary of Treatment-Emergent Adverse Events Leading to Discontinuation of Study Drug by MedDRA System Organ Class and Preferred Term	Safety	
3.2.19.X	Summary of Application Site Reactions by MedDRA Preferred Term in Decreasing Order of Frequency	Safety	
3.3 Laboratory			
3.3.1.1.X	Summary of Laboratory Values – Hematology	Safety	X
3.3.1.2.X	Summary of Laboratory Values – Chemistry	Safety	X
3.3.2.1.X	Shift Summary of Hematology Laboratory Values – to the Worst Abnormal Value	Safety	
3.3.2.2.X	Shift Summary of Chemistry Laboratory Values – to the Worst Abnormal Value	Safety	
3.3.3.1.X	Shift Summary of Hematology Laboratory Values in CTCAE Grade – to the Worst Grade Abnormal Value	Safety	X
3.3.3.2.X	Shift Summary of Chemistry Laboratory Values in CTCAE Grade – to the Worst Grade Abnormal Value	Safety	X
3.4 Vital Signs			
3.4.1.x	Summary of Systolic Blood Pressure	Safety	X
3.4.2.x	Summary of Diastolic Blood Pressure	Safety	X
3.4.3.x	Summary of Pulse	Safety	X
3.4.4.x	Summary of Respiratory Rate	Safety	X
3.4.5.x	Summary of Body Temperature	Safety	X

Note: For AE tables ending with "x," separate tables will be provided for the DBVC period, the OLE period, and the entire study period. For Laboratory and Vital Sign tables ending with "x," separate tables will be provided for the DBVC period and the OLE period.

Figures

Figure No.	Title	Population
Efficacy		
4.1.1.1	Mean and Standard Error Plot of AN Count During the Treatment Period	ITT
4.1.1.2	Mean and Standard Error Plot of Change From Baseline in AN Count During the Treatment Period	ITT
4.1.1.3	Mean and Standard Error Plot of Percent Change From Baseline in AN Count During the Treatment Period	ITT
4.1.2.1	Mean and Standard Error Plot of AN Count in Anatomical Areas With Pre-Existing AN at Baseline During the Treatment Period	ITT
4.1.2.2	Mean and Standard Error Plot of Change From Baseline in AN Count in Anatomical Areas with Pre-Existing AN During the Treatment Period	ITT
4.1.2.3	Mean and Standard Error Plot of Percent Change From Baseline in AN Count in Anatomical Areas with Pre-Existing AN During the Treatment Period	ITT
4.1.3	Proportion of Participants Achieving AN50 During the Treatment Period	ITT
4.1.4	Proportion of Participants Achieving AN75 During the Treatment Period	ITT
4.1.5	Proportion of Participants Achieving AN90 During the Treatment Period	ITT
4.1.6	Proportion of Participants Achieving AN100 During the Treatment Period	ITT

Figure No.	Title	Population
4.2.1.1	Mean and Standard Error Plot of By-Visit Skin Pain NRS Score During the Treatment Period	ITT
4.2.1.2	Mean and Standard Error Plot of Change From Baseline in By-Visit Skin Pain NRS Score During the Treatment Period	ITT
4.2.3.1	Mean and Standard Error Plot of Daily Skin Pain NRS Score From Baseline to Day 28	ITT
4.2.3.2	Mean and Standard Error Plot of Change From Baseline in Daily Skin Pain NRS Score From Baseline to Day 28	ITT
4.3.1.1	Mean and Standard Error Plot of By-Visit Itch NRS Score During the Treatment Period	ITT
4.3.1.2	Mean and Standard Error Plot of Change From Baseline in By-Visit Itch NRS Score During the Treatment Period	ITT
4.3.2	Proportion of Participants Achieving ≥ 4-Point Improvement in By-Visit Itch NRS Score During the Treatment Period	ITT (baseline Itch NRS score ≥ 4)
4.3.3.1	Mean and Standard Error Plot of Daily Itch NRS Score From Baseline to Day 28	ITT
4.3.3.2	Mean and Standard Error Plot of Change From Baseline in Daily Itch NRS Score From Baseline to Day 28	ITT
4.4.1	Proportion of Participants Achieving HiSCR During the Treatment Period	ITT
4.5.1	Mean and Standard Error Plot of IHS4 Score During the Treatment Period	ITT
4.5.2	Mean and Standard Error Plot of Change From Baseline in IHS4 Score During the Treatment Period	ITT
[redacted]		
4.12.1	Box Plot of Selected Laboratory Values in the DBVC Period	ITT
4.12.2	Box Plot of Change From Baseline in Selected Laboratory Values During the DBVC Period	ITT
4.12.3	Box Plot of Percent Change From Baseline in Selected Laboratory Values During the DBVC Period	ITT

Listings

Listing No.	Title
2.1 Enrollment and Disposition	
2.1.1	Participant Enrollment and Disposition Status
2.1.2	Participant Inclusion and Exclusion Criteria Violations
2.2 Protocol Deviations	
2.2.1	Protocol Deviations
2.3 Analysis Populations	
2.3	Analysis Populations
2.4 Demographic and Baseline Characteristics (Including Prior and Concomitant Medications)	
2.4.1	Demographic Characteristics
2.4.2	Baseline Disease Characteristics
2.4.3	Disease History
2.4.4	Medical History
2.4.5	Prior and Concomitant Medications
2.4.6	Prior Medications for HS
2.5 Drug Compliance	
2.5.1	Study Drug Exposure and Compliance
2.6 Efficacy	
2.6.1.1	Lesion Count
2.6.1.2	AN Count
2.6.2	Skin Pain NRS Score
2.6.3	Itch NRS Score
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
2.7 Adverse Events	
2.7.1	Adverse Events
2.7.2	Adverse Events Leading to Study Drug Discontinuation
2.7.3	Serious Adverse Events
2.7.4	Treatment-Related Adverse Events
2.7.5	Adverse Events With a Fatal Outcome
2.7.6	Adverse Events Leading to Dose Interruption
2.7.8	Grade 3 or Higher Adverse Events
2.7.9	Application Site Reactions
2.8 Laboratory Data	
2.8.1.1	Clinical Laboratory Values – Hematology
2.8.1.2	Clinical Laboratory Values – Chemistry
2.8.1.3	Abnormal Clinical Laboratory Values – Hematology
2.8.1.4	Abnormal Clinical Laboratory Values – Chemistry
Vital Signs	
2.8.2.1	Vital Signs

Listing No.	Title
2.8.2.2	Abnormal Vital Sign Values
2.8.2.3	Alert Vital Sign Values