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

NCT Number: NCT05593445

Document Date: Statistical Analysis Plan (Original Version): 24-August-2023

Statistical Analysis Plan

INCB 18424-220

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

IND Number:	**77,101**
Sponsor:	**Incyte Corporation** **1801 Augustine Cut-Off** **Wilmington, DE 19803** **United States**
Protocol Version:	**Protocol Amendment 2 dated 07 MAR 2023**
CRF Approval Date:	**02 JUN 2023**
SAP Version:	**Original**
SAP Author:	**[redacted], PhD** **[redacted]**
Date of Plan:	**24 AUG 2023**

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

TABLE OF CONTENTS

TITLE PAGE 1
TABLE OF CONTENTS 2
LIST OF ABBREVIATIONS 5
1. INTRODUCTION 6
2. STUDY INFORMATION, OBJECTIVES, AND ENDPOINTS 7
2.1. Protocol and Case Report Form Version 7
2.2. Study Objectives and Endpoints 7
3. STUDY DESIGN 9
3.1. Randomization 9
3.2. Control of Type I Error 9
3.3. Sample Size Considerations 10
3.4. Schedule of Assessments 10
4. DATA HANDLING DEFINITIONS AND CONVENTIONS 11
4.1. Scheduled Study Evaluations and Study Periods 11
4.1.1. Day 1 11
4.1.2. Study Day 11
4.1.3. Baseline Value 11
4.1.4. Handling of Missing and Incomplete Dates 12
4.2. Variable Definitions 12
4.2.1. Body Mass Index 12
4.2.2. Prior and Concomitant Medication 12
5. STATISTICAL METHODOLOGY 13
5.1. General Methodology 13
5.2. Treatment Groups 13
5.3. Analysis Populations 13
5.3.1. Intent-to-Treat Population 13
5.3.2. Safety Population 13
5.3.3. Open-Label Extension Evaluable Population 13
[redacted] [redacted] 14
6. BASELINE, EXPOSURE, AND DISPOSITION 14
6.1. Demographics, Baseline Characteristics, and Disease History 14

6.1.1. Demographics and Baseline Characteristics 14
6.1.2. Baseline Disease Characteristics 14
6.1.3. Disease History 15
6.1.4. Medical History 15
6.2. Disposition of Participant 15
6.3. Protocol Deviations 15
6.4. Exposure 15
6.5. Study Drug Compliance 16
6.6. Prior and Concomitant Medication 16
7. EFFICACY 17
7.1. Efficacy Hypotheses 17
7.2. Efficacy Measures 17
[redacted] [redacted] 17
7.2.2. Clinical Lichen Sclerosus Score 17
[redacted] [redacted] 17
7.2.4. Body Surface Area 18
7.2.5. Patient-Reported Outcomes 18
7.2.5.1. Itch Numerical Rating Scale 18
7.2.5.2. Skin Pain Numerical Rating Scale 19
[redacted] [redacted] 19
[redacted] [redacted] 20
[redacted] [redacted] 20
[redacted] [redacted] 20
[redacted] [redacted] 21
7.3. Analysis of the Primary Efficacy Parameter 21
7.3.1. Primary Efficacy Analysis 21
7.3.2. Subgroup Analyses for Primary Endpoint 21
7.4. Analysis of the Secondary Efficacy Parameters 22
7.4.1. Continuous Efficacy Endpoints 22
7.4.2. Time-To-Event Efficacy Endpoint 22
[redacted] [redacted] 22
[redacted] [redacted] 24
[redacted] [redacted] 24

[redacted] 24
9. SAFETY AND TOLERABILITY 25
9.1. General Considerations 25
9.2. Adverse Events 25
9.2.1. Adverse Event Definitions 25
9.2.2. Adverse Event Summaries 26
9.3. Clinical Laboratory Tests 27
9.3.1. Laboratory Value Definitions 27
9.3.2. Laboratory Value Summaries 27
9.4. Vital Signs 28
10. PLANNED ANALYSES 29
11. CHANGES AND MODIFICATIONS TO THE ANALYSIS PLAN 29
11.1. Changes to Protocol-Defined Analyses 29
11.2. Changes to the Statistical Analysis Plan 29
12. REFERENCES 30
APPENDIX A. PLANNED TABLES, FIGURES, AND LISTINGS 32

LIST OF TABLES

Table 1: Objectives and Endpoints 7
Table 2: Criteria for Clinically Notable Vital Sign Abnormalities 28
Table 3: Statistical Analysis Plan Versions 29

LIST OF FIGURES

Figure 1: Study Design Schema 9

LIST OF ABBREVIATIONS

Abbreviation	Term
AE	adverse event
BID	twice daily/2 times a day
BMI	body mass index
BSA	body surface area
CI	confidence interval
CLISSCO	Clinical Lichen Sclerosus Score
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]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
ITCH4	≥ 4-point improvement in WI-NRS score
ITT	intent-to-treat
LS	lichen sclerosus
[redacted]	[redacted]
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed Model Repeated Measures
NCI	National Cancer Institute
NRS	numeric rating scale
OLE	open-label extension
[redacted]	[redacted]
[redacted]	[redacted]
PT	preferred term
SAP	Statistical Analysis Plan
SI	International System of Units
SOC	system organ class
TEAE	treatment-emergent adverse event
[redacted]	[redacted]
WHO	World Health Organization

1. INTRODUCTION

INCB 18424-220 is a Phase 2, randomized, DBVC study to evaluate the efficacy and safety of ruxolitinib 1.5% cream in participants with LS over a 12-week treatment period followed by a 12-week OLE period during which all participants receive active treatment. 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 18424-220 Protocol.

2. STUDY INFORMATION, OBJECTIVES, AND ENDPOINTS

2.1. Protocol and Case Report Form Version

This SAP is based on INCB 18424-220 Protocol Amendment 2 dated 07 MAR 2023 and CRFs approved 02 JUN 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 LS.	• ITCH4 response (defined as a ≥ 4-point improvement in Itch NRS score from baseline) at Week 12. (Proportion of participants with ITCH4 at Week 12.)
Secondary	
To further evaluate efficacy of ruxolitinib 1.5% cream BID in participants with LS.	• Change from baseline in CLISSCO score at Week 12. • Change from baseline in the Skin Pain NRS score at Week 12. • Time to achieve ITCH4.
To evaluate the safety and tolerability of ruxolitinib 1.5% cream BID.	• The type, frequency, and severity of AEs, including the changes in vital signs and clinical laboratory blood samples.
[redacted]	

Table 1: Objectives and Endpoints (Continued)

Objectives	Endpoints
[redacted]	

3. STUDY DESIGN

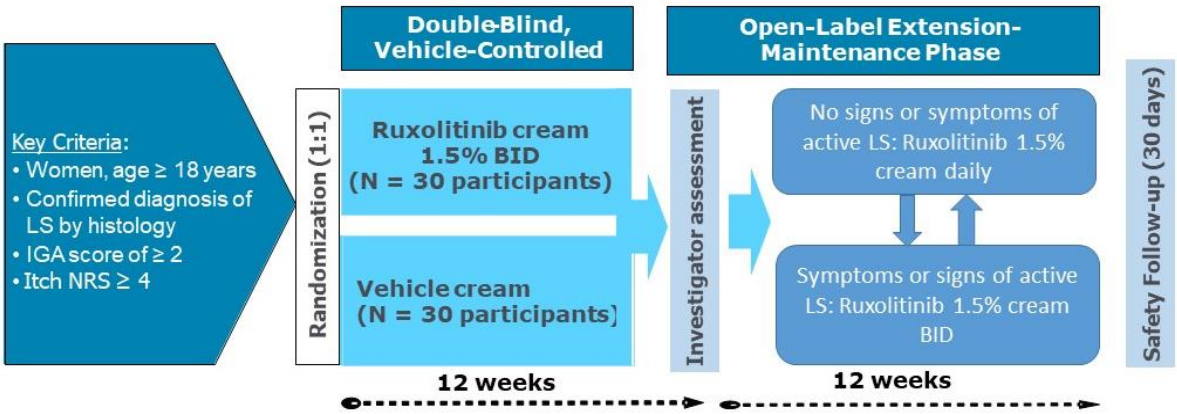
This is a Phase 2, randomized, DBVC study in participants aged ≥ 18 years with a biopsy-confirmed diagnosis of LS. The study will consist of a 12-week DBVC period followed by a 12-week OLE period.

Approximately 60 female participants will be randomized 1:1 to either ruxolitinib 1.5% cream or vehicle cream. Participants will apply either ruxolitinib 1.5% cream or vehicle cream (both BID) through Week 12 to all the affected areas identified at baseline.

At Week 12, participants who meet the criteria (compliant with the Protocol and without safety concerns) will enter the 12-week OLE period. Participants randomized initially to vehicle in the DBVC period will be crossed over to ruxolitinib 1.5% cream and participants randomized to ruxolitinib 1.5% cream at baseline will continue to apply ruxolitinib 1.5% cream through Week 24 in an open-label fashion. During the OLE period, participants who have no symptoms or signs of active LS (based on investigator's assessment) will apply ruxolitinib 1.5% cream once daily through Week 24. Ruxolitinib dose can be increased to BID for those participants who have symptoms or signs of active disease (based on investigator's assessment) until symptoms and signs of active LS have resolved in the opinion of the investigator.

The study schema is shown below in Figure 1.

Figure 1: Study Design Schema



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

3.1. Randomization

Approximately 60 participants will be randomized 1:1 to either ruxolitinib 1.5% cream or vehicle cream BID.

3.2. Control of Type I Error

The significance level for primary efficacy analysis will be 0.05 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 cream BID. Due to the lack of data with ruxolitinib cream in LS, 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 provide sufficient data to determine whether ruxolitinib 1.5% cream warrants further investigation in LS. Also, the sample size is considered sufficient to provide enough data for the evaluation of the ruxolitinib 1.5% cream safety profile in LS.

3.4. Schedule of Assessments

Refer to Protocol Amendment 2 dated 07 MAR 2023for 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 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.

4.1.2. Study Day

If a visit/reporting date is on or after Day 1, then the study day at the visit/reporting date will be calculated as

Day # = (visit/reporting date – Day 1 date + 1).

If the visit/reporting date is before Day 1, then the study day at the visit/reporting date will be calculated as

Day # = (visit/reporting date – Day 1 date).

A study day of –1 indicates 1 day before Day 1.

4.1.3. Baseline Value

Baseline is the last nonmissing measurement obtained before or on the day of the first application of ruxolitinib cream or vehicle cream for the DBVC period.

For randomized participants not treated with any study treatment, 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 LS diagnosis date will be handled as follows in the calculation:

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

4.2. Variable Definitions

4.2.1. Body Mass Index

Body mass index will be calculated as follows:

$$\text{BMI (kg/m}^2\text{)} = [\text{weight (kg)}] / [\text{height (m)}]^2$$

4.2.2. Prior and Concomitant Medication

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

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 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, DBVC study with an open-label extension period. Table summaries, unless otherwise indicated, will present data by treatment group.

For the DBVC period, the treatment groups will be vehicle cream and ruxolitinib 1.5% cream BID.

For the OLE period, the participants will be grouped into the following groups according to the treatment group they received during the DBVC period:

- Vehicle cream to ruxolitinib 1.5% cream
- Ruxolitinib 1.5% cream

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, and analyses of all efficacy data.

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.

All safety analyses will be conducted using the safety population.

5.3.3. Open-Label Extension Evaluable Population

All analyses for the OLE period will be conducted with the OLE evaluable population, which includes all participants who applied ruxolitinib 1.5% cream at least once during the OLE period.

[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 evaluable population in the OLE period: age, sex, race, ethnicity, weight, height, and BMI.

6.1.2. Baseline Disease Characteristics

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

- Disease duration (years)
- LS diagnosis confirmed by biopsy (yes/no)
- Anatomic location (vulva/vagina, other anogenital area, other body areas)
- Complications with LS (yes/no)
- Number of flares/episodes of LS in the past year other than current flare
- Clinical manifestations related to LS (% participants with each of the manifestations)
- Prior treatment received for LS (yes/no)
- History of other skin disease (yes/no)
- Family history of LS (yes/no)
- History of skin infection related to LS requiring antibiotic treatment (yes/no)
- Days since onset of current episode of LS flare/activity
- Skin type (Type I/II/III/IV/V/VI)
- Total % BSA affected
- Baseline Itch NRS score
- Baseline total CLISSCO
- Baseline CLISSCO subscores (total score for each section)

- Baseline Skin Pain NRS
- Baseline total LSCA score
- Baseline IGA score

6.1.3. Disease History

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

Time since diagnosis (ie, disease duration [years]) will be calculated as follows:

> Disease duration (years) = (date of randomization – date of initial LS diagnosis +1) / 365.25

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 as part of 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 as part of the OLE evaluable population.

6.3. Protocol Deviations

Protocol deviations will be summarized and listed by treatment group and overall ITT and OLE populations.

6.4. Exposure

For participants in the safety population in the DBVC period, the OLE evaluable population in the OLE period, and those who applied ruxolitinib 1.5% cream BID throughout the study, study drug exposure will be summarized by treatment group, descriptively, as follows:

- Duration of treatment (days):

  Date of last application of ruxolitinib cream or vehicle cream in the specific period – date of first application of ruxolitinib cream or vehicle cream in the specific period + 1

- Total amount of cream applied (g):

  Total weight of ruxolitinib cream or vehicle cream dispensed in the specific period – total weight of ruxolitinib cream or vehicle cream returned in the specific period

- Average daily application:

  Total amount of cream applied in the specific period (g) / [duration of treatment with study cream in the specific period (days) – number of interrupted days with study cream in the specific period]

6.5. Study Drug Compliance

The overall compliance (%) for the application of ruxolitinib 1.5% cream BID or vehicle cream BID during the DBVC period will be calculated for all participants in the safety population as follows:

Cream application compliance (%) = 100 × [total number of nonmissing applications] / [total number of intended applications], where

Total number of nonmissing applications = number of applications that the participants actually applied during the study, and

Total number of intended applications = 2 × duration of treatment (days) – number of interrupted applications.

6.6. Prior and Concomitant Medication

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

7. EFFICACY

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

7.1. Efficacy Hypotheses

The primary hypothesis is the superiority of ruxolitinib 1.5% cream BID compared with vehicle cream BID in the proportions of participants achieving ITCH4 at Week 12.

7.2. Efficacy Measures

[redacted]

7.2.2. Clinical Lichen Sclerosus Score

The CLISSCO is a validated tool to assess disease severity in vulvar LS. Clinical Lichen Sclerosus Score consists of 12 items divided into 3 sections: symptoms (3 items; likely reversible [ie, itch, pain, dysuria]); signs (3 items; possibly reversible [ie, whitening, petechiae/ecchymosis, fissures]); and architectural changes (6 items; irreversible [ie, skin fusion, perianal involvement, etc]). All symptoms, signs, and architectural changes will be rated on a 4-point Likert scale: 0 (absent), 1 (mild), 2 (moderate), and 3 (severe) (Erni et al 2021). The investigator will be asked to document the score of each of the 12 items; the CLISSCO will be calculated by summing the score of each question with a maximum score of 36 and a minimum score of 0. The higher the score, the more severe the disease. Additionally, the total score for each of the 3 sections (ie, symptoms, signs, and architectural changes) will be summarized by summing the scores of the questions in each section. No imputation will be performed for missing values.

[redacted]

[redacted]

7.2.4. Body Surface Area

Total % BSA affected will be estimated at each visit. Body surface area assessment will be approximated to the nearest 0.1% using the Palmar Method as a guide. The Palmar Method is the palm plus 5 digits, with fingers tucked together and thumb tucked to the side (handprint) as 1% BSA and the thumb as 0.1% BSA.

7.2.5. Patient-Reported Outcomes

Patient-reported outcomes will be collected and assessed. For all patient-reported outcome assessments conducted at the study site, in order to avoid bias in the participants' responses to the questionnaires, assessments should be completed before any other evaluations or study procedures on the day of the study visit and prior to treatment-related discussions with the investigator or study site staff.

7.2.5.1. Itch Numerical Rating Scale

The Itch NRS is a daily participant-reported measure (24-hour recall) of the worst level of itch intensity (Kimball et al 2016, Silverberg et al 2021). Participants will be instructed to complete and record their Itch NRS in a diary each evening beginning on the day of screening through Week 24 or treatment discontinuation. Participants will rate itch severity of their LS by selecting a number from 0 (no itch) to 10 (worst imaginable itch) that best describes the worst level of itch they experienced in the past 24 hours.

The participant's baseline Itch NRS score will be determined by averaging the 7 daily NRS scores before Day 1 (ie, Days –7 to –1) for all the 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 (out of the 7) are missing, the by-visit Itch NRS score will be set to missing. For all daily itch-related analyses, including time to achieve Itch NRS score improvement of at least 4 points analysis, baseline will be defined as the last available Itch NRS score during the week prior to Day 1 (ie, Days −7 to −1). Nonresponder imputation will be used to handle missing postbaseline values in the DBVC period.

A participant is said to have an ITCH4 response if they achieved a ≥ 4-point improvement in Itch NRS score from baseline. For determining ITCH4 response at a visit, the by-visit Itch NRS score will be compared with the baseline Itch NRS score. For the endpoint time to achieve ITCH4, the daily Itch NRS score will be compared with the baseline Itch NRS score.

7.2.5.2. 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 24 or treatment discontinuation. Participants will rate their pain, which will include all types of pain (eg, burning, tearing, pulling, stabbing, etc) severity of the LS by selecting a number from 0 (no pain) to 10 (worst imaginable pain) that best describes the worst level of pain they experienced in the past 24 hours.

The participant's baseline Skin Pain NRS score will be determined by averaging the 7 daily NRS scores before Day 1 (ie, Days –7 to –1) for all the 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 (out of the 7) are missing, the Skin Pain NRS score at the visit will be set to missing. For all daily skin pain-related analyses, baseline will be defined as the last available Skin Pain NRS score during the week prior to Day 1 (ie, Days −7 to −1). No imputation will be performed for missing values.

[redacted]

[redacted]

7.3. Analysis of the Primary Efficacy Parameter

7.3.1. Primary Efficacy Analysis

The primary efficacy analysis will be based on the ITT population. The primary efficacy endpoint is the ITCH4 response, defined as a ≥ 4-point improvement in Itch NRS score from baseline, at Week 12. The comparison of the proportions of participants with ITCH4 at Week 12 will be made between the ruxolitinib 1.5% cream BID arm and the vehicle cream BID arm. The primary hypothesis (the superiority of ruxolitinib 1.5% cream BID compared with vehicle cream in participants with LS) will be tested using a Chi-squared test at a 2-sided $\alpha = 0.05$ level; the p-value will be provided. The Wald 95% CI will be provided for the ITCH4 response rate. A Fisher's exact test will be performed for the comparison if either of the treatment groups has an expected cell count less than 5. Summary of ITCH4 rates will be reported for each treatment group.

All nonresponders in the DBVC period, as well as all participants who are missing Week 12 values, will be defined as nonresponders for the primary analysis.

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 Itch NRS score: ≥ 4 and < 7 versus ≥ 7
- Baseline IGA score: 2, mild; 3, moderate; and 4, severe
- Baseline categorical age: ≥ 18 and < 65 versus ≥ 65
- Baseline duration of disease: ≤ 1, > 1 and ≤ 5, >5
- Childbearing potential: pre versus postmenopausal status

The primary efficacy endpoint will be summarized using descriptive statistics based on the ITT population for the above defined 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

Summary statistics for the following continuous measurements, including change and percentage change from baseline, if applicable, will be presented for each treatment arm:

- CLISSCO (total and subscores) at Week 12
- Skin Pain by-visit NRS score at Week 12

The summary statistics includes sample size, mean, median, standard deviation, minimum, maximum, first quartile, third quartile, and 95% CI. An MMRM may be fit for the comparison between ruxolitinib 1.5% cream BID and vehicle cream BID at Week 12. The MMRM will include the fixed effect of treatment, visit, and treatment by visit interaction. The variance-covariance matrix of the within-participant errors in MMRM will be modeled as unstructured.

7.4.2. Time-To-Event Efficacy Endpoint

The time to achieve Itch NRS score improvement of at least 4 points from baseline (ie, ITCH4) will use the daily Itch NRS score in comparison to baseline Itch NRS score in the DBVC period. A log-rank test will be used for between-treatment group comparisons. The hazard ratio and its 95% CI will be estimated based on the Cox regression model using Efron's method accounting for ties. Kaplan-Meier curves will be presented by treatment groups. The number of participants, number of events, and number of censoring will be summarized by treatment groups. The Kaplan-Meier estimate of median time will be presented with its 95% CI. The 95% CI will be calculated using the method by Brookmeyer and Crowley (1982).

The last available value for time to ITCH4 during the DBVC period is the last nonmissing measurement obtained after the first application of study cream and within 30 days after the last application of study cream in DBVC period, or before the first application of study cream in the OLE period, whichever is earlier.

[redacted]

[redacted]

9. SAFETY AND TOLERABILITY

Appendix A provides a list of data displays.

9.1. General Considerations

The analyses in this section will be provided for the safety population in the DBVC period and OLE evaluable population in the OLE period, 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 until 30 days of the last application of study drug. For participants who cross over treatments, the first application date is period-specific; however, 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. For purposes of analysis, all AEs will be considered TEAEs unless the AE can unequivocally be defined as not treatment-emergent.

Adverse events will be tabulated by MedDRA PT and SOC. Severity of AEs will be graded using the NCI CTCAE v5. 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.

Any missing onset date of a TEAE will be handled according the following rules:

- If completely missing, then Day 1 will be used.
- If only the day is missing, then the first day of the month or Day 1, whichever is later, will be used.
- If both the month and day are missing, then 01 JAN of the year or Day 1, whichever is later, will be used.

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 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 Application Site Reactions by MedDRA PT in decreasing order of frequency

9.3. Clinical Laboratory Tests

9.3.1. Laboratory Value Definitions

All laboratory assessments will be performed using a central laboratory with the exception of urine pregnancy tests. 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

Clinical laboratory tests, including hematology and serum chemistry, will be performed at the Protocol-specified visits. If specific safety issues arise, additional unscheduled laboratory tests/analyses may be performed at the discretion of the investigator.

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. 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

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 2. 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 2: 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	> 38ºC	< 35.5ºC
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 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 3.

Table 3: Statistical Analysis Plan Versions

SAP Version	Date
Original	24 AUG 2023

11.1. Changes to Protocol-Defined Analyses

Not applicable.

11.2. Changes to the Statistical Analysis Plan

Not applicable.

12. REFERENCES

[redacted]

Brookmeyer R, Crowley J. A confidence interval for the median survival time. Biometrics 1982;38:29-41.

[redacted]

Erni B, Navarini AA, Huang D, Schoetzau A, Kind A, Mueller SM. Proposition of a severity scale for lichen sclerosus: The "Clinical Lichen Sclerosus Score." Dermatol Ther 2021;34:e14773.

[redacted]

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.

Opzelura (ruxolitinib) [prescribing information]. Incyte Corporation; July 2022.

Powell JJ, Wojnarowska F. Lichen sclerosus. Lancet 1999;353:1777-1783.

Ruxolitinib Cream Investigator's Brochure. Wilmington, DE: Incyte Corporation.

[redacted]

Silverberg JI, DeLozier A, Sun L, et al. Psychometric properties of the itch numeric rating scale, skin pain numeric rating scale, and atopic dermatitis sleep scale in adult patients with moderate-to-severe atopic dermatitis. Health Qual Life Outcomes 2021;19:247.

[redacted]

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	ITT	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			
1.3.1	Summary of Baseline Disease Characteristics	ITT	X
1.3.2	Summary of Baseline Disease Characteristics	OLE	X
1.4 Prior and Concomitant Medication			
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 in the DBVC Period	ITT	X
Efficacy			
2.1 Itch NRS Score			
2.1.1	Summary and Analysis of Participants Achieving ITCH4 in Treatment Periods	ITT	
2.1.2	Summary and Analysis of Participants Achieving ITCH4 by Baseline ITCH NRS Score in Treatment Periods	ITT	
2.1.3	Summary and Analysis of Participants Achieving ITCH4 by Baseline IGA Score in Treatment Periods	ITT	
2.1.4	Summary and Analysis of Participants Achieving ITCH4 by Baseline Age Group in Treatment Periods	ITT	
2.1.5	Summary and Analysis of Participants Achieving ITCH4 by Baseline Duration of Disease in Treatment Periods	ITT	
2.1.6	Summary and Analysis of Participants Achieving ITCH4 by Childbearing Potential Status in Treatment Periods	ITT	
2.1.7	Summary and Analysis of By-Visit Itch NRS Score in Treatment Periods	ITT	
2.1.8	Summary and Analysis of Daily Itch NRS Score from Day 1 to Day 28	ITT	
2.1.9	Summary and Analysis of Time to ITCH4 in the DBVC Period	ITT	
2.2 CLISSCO			
2.2.1	Summary and Analysis of CLISSCO Total Score and Subscores in Treatment Periods	ITT	

Table No.	Title	Population	Standard
2.3 Skin Pain NRS Score			
2.3.1	Summary and Analysis of By-Visit Skin Pain NRS Score in Treatment Periods	ITT	
[redacted]			
2.11 BSA			
2.11.1	Summary of By-Visit Total % BSA in Treatment Periods	ITT	
Safety			
3.1 Exposure & 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 in Treatment Periods	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	

Table No.	Title	Population	Standard
3.2.13.x	Summary of Treatment-Related Serious Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term	Safety	X
3.2.14.x	Summary of Treatment-Emergent Adverse Events With a Fatal Outcome by MedDRA System Organ Class and Preferred Term	Safety	X
3.2.15.x	Summary of Treatment-Emergent Adverse Events Leading to Dose Interruption by MedDRA System Organ Class and Preferred Term	Safety	X
3.2.16.x	Summary of Treatment-Emergent Adverse Events Leading to Discontinuation of Study Drug by MedDRA System Organ Class and Preferred Term	Safety	X
3.2.17.x	Summary of Application Site Reactions by MedDRA Preferred Term in Decreasing Order of Frequency	Safety	X
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	X
3.3.2.2.x	Shift Summary of Chemistry Laboratory Values – to the Worst Abnormal Value	Safety	X
3.3.3.1x	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, OLE period, and the treatment periods (ie, DBVC and OLE periods). For laboratory and vital sign tables ending with "x," separate tables will be provided for the DBVC and the OLE periods.

Figures

Figure No.	Title	Population
Efficacy		
4.1.1.1	Proportion of Participants Achieving ITCH4 in Treatment Periods	ITT
4.1.1.2	Proportion of Participants Achieving ITCH4 by Baseline ITCH NRS Score in Treatment Periods	ITT
4.1.1.3	Proportion of Participants Achieving ITCH4 by Baseline IGA Score in Treatment Periods	ITT
4.1.1.4	Proportion of Participants Achieving ITCH4 by Baseline Age Group in Treatment Periods	ITT
4.1.1.5	Proportion of Participants Achieving ITCH4 by Baseline Duration of Disease in Treatment Periods	ITT
4.1.1.6	Proportion of Participants Achieving ITCH4 by Pre and Postmenopausal Status in Treatment Periods	ITT
4.1.1.7	Forest Plot of Response Rate Difference in Achieving ITCH4 at Week 12	ITT
4.1.2.1	Mean and Standard Error Plot of By-Visit Itch NRS Score in Treatment Periods	ITT
4.1.2.2	Mean and Standard Error Plot of Change From Baseline in By-Visit Itch NRS Score in Treatment Periods	ITT

Figure No.	Title	Population
4.1.3.1	Mean and Standard Error Plot of Daily Itch NRS Score From Day 1 to Day 28	ITT
4.1.3.2	Mean and Standard Error Plot of Change From Baseline in Daily Itch NRS Score From Day 1 to Day 28	ITT
4.1.4	Kaplan-Meier Curve of Time to ITCH4 in the DBVC Period	ITT
4.2.1.1	Mean and Standard Error Plot of Total CLISSCO in Treatment Periods	ITT
4.2.1.2	Mean and Standard Error Plot of Change From Baseline in Total CLISSCO in Treatment Periods	ITT
4.2.1.3	Mean and Standard Error Plot of Percentage Change From Baseline in Total CLISSCO Score in Treatment Periods	ITT
4.3.1.1	Mean and Standard Error Plot of By-Visit Skin Pain NRS Score in Treatment Periods	ITT
4.3.1.2	Mean and Standard Error Plot of Change From Baseline in By-Visit Skin Pain NRS Score in Treatment Periods	ITT
4.3.1.3	Mean and Standard Error Plot of Percentage Change From Baseline in By-Visit Skin Pain NRS Score in Treatment Periods	ITT
[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]	[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]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
4.10.1.1	Mean and Standard Error Plot of Total % BSA in Treatment Periods	ITT
4.10.1.2	Mean and Standard Error Plot of Change From Baseline in Total % BSA in Treatment Periods	ITT
4.10.1.3	Mean and Standard Error Plot of Percentage Change From Baseline in Total % BSA in Treatment Periods	ITT

Listings

Listing No.	Title
2.1 Demographic and Baseline Characteristics	
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 and Violations
2.3 Data Excluded From [redacted] Efficacy, and/or Safety Analyses	
2.3	Analysis Populations
2.4 Demographic and Baseline Characteristics (Including Prior and Concomitant Medications)	
2.4.1	Demographic and Baseline Characteristics
2.4.2	Baseline Disease Characteristics
2.4.3	Medical History
2.4.4	Prior and Concomitant Medications
2.4.5	Prior Medications for Lichen Sclerosus
2.5 Drug Exposure and Compliance	
2.5.1	Study Drug Exposure and Compliance in the DBVC Period
2.5.2	Study Drug Exposure in the OLE Period
2.6 Efficacy	
2.6.1.1	By-Visit Itch NRS Score
2.6.1.2	Daily Itch NRS Score
2.6.2	CLISSCO
2.6.3.1	By-Visit Skin Pain NRS Score
2.6.3.2	Daily Skin Pain NRS Score
[redacted]	[redacted]
[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 Interruption of Study Drug
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
2.8.1.4	Abnormal Clinical Laboratory Values – Hematology
2.8.1.5	Abnormal Clinical Laboratory Values – Chemistry

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