Statistical Analysis Plan

INCB 18424-326

Topical Ruxolitinib Evaluation in Atopic Dermatitis Study (TRuE-AD4)

A Phase 3b, Double-Blind, Multicenter, Randomized, Vehicle-Controlled, Efficacy, and Safety Study of Ruxolitinib Cream in Adults With Moderate Atopic Dermatitis

Product:	**Ruxolitinib Cream**
IND Number	[redacted]
EU CT Number:	**2023-505433-27-00**
Sponsor:	**Incyte Corporation** **1801 Augustine Cut-Off** **Wilmington, DE 19803 USA**
Protocol Version:	**Protocol Amendment 2 dated 08 AUG 2024**
CRF Approval Date:	**24 MAR 2025**
SAP Version:	**Original**
SAP Author:	[redacted] [redacted]**, Biostatistics**
Date of Plan:	**06 MAY 2025**

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
AD	atopic dermatitis
AE	adverse event
ANOVA	analysis of variance
ASR	application site reaction
BID	twice daily
BSA	body surface area
CI	confidence interval
CMH	Cochran-Mantel-Haenszel (test for general association)
CRF	case report form
CTCAE	Common Terminology Criteria for Adverse Events
DB	double-blinded
DLQI	Dermatology Life Quality Index
EASI	Eczema Area and Severity Index
EASI50/75/90	≥ 50%, ≥ 75%, or ≥ 90% improvement in Eczema Area and Severity Index score
eCRF	electronic case report form
EOS	end of study
EOT	end of treatment
FDA	Food and Drug Administration
HADS	Hospital Anxiety and Depression Scale
IGA	Investigator's Global Assessment
IGA-TS	Investigator's Global Assessment Treatment Success (IGA score of 0 or 1 with ≥ 2-grade improvement from baseline)
ITT	intent-to-treat
MedDRA	Medical Dictionary for Regulatory Activities
MI	multiple imputation
MMRM	mixed-model repeated measures
NRI	nonresponder imputation
NRS	Numerical Rating Scale
PD	pharmacodynamic(s)
PK	pharmacokinetic(s)
POEM	Patient-Oriented Eczema Measure
PP	per Protocol
PRO	patient-reported outcome
PROMIS®	Patient-Reported Outcomes Measurement Information System
PT	preferred term

Abbreviation	Term
QoL	quality of life
ROW	rest of world
SAP	Statistical Analysis Plan
SCORAD	SCORing Atopic Dermatitis
SOC	system organ class
TCI	topical calcineurin inhibitor
TCS	topical corticosteroid
TEAE	treatment-emergent adverse event
WPAI-AD	Work Productivity and Activity Impairment – Atopic Dermatitis
WHO	World Health Organization
VAS	visual analog scale
VC	vehicle-controlled
VCE	vehicle-controlled extension

1. INTRODUCTION

This is a Phase 3b, multicenter, randomized, double-blind, VC study with an extension period in adult (aged ≥ 18 years) participants with moderate AD who had an inadequate response to, or are intolerant to, or contraindicated to TCSs and TCIs. The study will be conducted at approximately 100 sites in North America, Asia-Pacific, and Europe. Approximately 225 participants will be randomized 2:1 to blinded treatment with ruxolitinib 1.5% cream BID or vehicle cream BID, with stratification by baseline EASI score (< 16 or ≥ 16) and geographic region (ROW or Europe).

In this study, participants will apply blinded study treatment for 8 weeks. After completing 8 weeks of continuous treatment, all eligible participants with an adequate response, defined as achieving at least EASI50 from baseline, will continue blinded treatment allocation and will be evaluated during an additional 16-week VCE period for durability of response. Participants will be eligible to enter the ruxolitinib 1.5% cream open-label escape arm if EASI50 is not achieved at Week 8 or if an EASI50 response from baseline is lost during the VCE period (Note: < EASI50 must be observed at 2 consecutive visits, one of which can be an unscheduled visit, at least 1 week apart.)

The purpose of this SAP is to provide details of the statistical analyses that have been outlined in the INCB 18424-326 Protocol. The Department of Biostatistics or designee will execute the scope of this plan, the Department of Clinical Pharmacokinetics or designee will execute the analyses of PK, and the Department of Translational Science or designee will execute the analyses of PD.

2. STUDY INFORMATION, OBJECTIVES, AND ENDPOINTS

2.1. Protocol and Case Report Form Version

This SAP is based on INCB 18424-326 Protocol Amendment 2 dated 08 AUG 2024 and CRFs approved 24 MAR 2025. 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
Coprimary	
To establish the efficacy of ruxolitinib cream in participants with moderate AD who had an inadequate response to, or are intolerant to, or contraindicated to TCSs and TCIs.	• The binary response status of EASI75 from baseline at Week 8. (EASI75 is defined as achieving ≥ 75% improvement in EASI score.) • The binary response status of IGA-TS at Week 8. (IGA-TS is defined as achieving an IGA score of 0 or 1 with ≥ 2-grade improvement from baseline.)
Key Secondary	
To further assess the treatment effects of ruxolitinib cream in participants with moderate AD who had an inadequate response to, or are intolerant to, or contraindicated to TCSs and TCIs.	• The binary response status of ITCH4 from baseline to Week 8. (ITCH4 is defined as achieving ≥ 4-point improvement in Itch NRS score.) • The binary response status of ITCH4 from baseline to Day 7. • The binary response status of ITCH4 from baseline to Day 3. • The binary response status of ITCH4 from baseline to Day 2.
Secondary	
To evaluate the safety and tolerability of ruxolitinib cream in participants with moderate AD who had an inadequate response to, or are intolerant to, or contraindicated to TCSs and TCIs.	The type and severity of AEs as well as changes from baseline in vital signs and laboratory data for hematology and serum chemistry.
To further evaluate the efficacy of ruxolitinib cream in participants with moderate AD who had an inadequate response to, or are intolerant to, or contraindicated to TCSs and TCIs.	• The binary response status of EASI75 from baseline at each postbaseline visit except Week 8. • The binary response status of IGA-TS from baseline at each postbaseline visit except Week 8. • The binary response status of ITCH4 from baseline at each postbaseline visit except Week 8. • Time to achieve ITCH4 during the VC period. • Time to achieve ITCH2 during the VC period. (ITCH2 is defined as achieving ≥ 2-point improvement from baseline in Itch NRS score.)

Table 1: Objectives and Endpoints (Continued)

Objectives	Endpoints
Secondary (continued)	
To further evaluate the efficacy of ruxolitinib cream in participants with moderate AD who had an inadequate response to, or are intolerant to, or contraindicated to TCSs and TCIs (continued).	• Change from baseline (pre–study cream application) in current Itch NRS score at 5, 15, 30, 45, and 60 minutes and 2, 4, and 6 hours post–initial dose on Day 1. • The binary response status of achieving at least a 2-point decrease from baseline in current Itch NRS score at 5, 15, 30, 45, and 60 minutes and 2, 4, and 6 hours post–initial dose on Day 1. • The binary response status of achieving at least a 4 point decrease from baseline in current Itch NRS score at 5, 15, 30, 45, and 60 minutes and 2, 4, and 6 hours post–initial dose on Day 1. • The binary response status of EASI50 from baseline at each postbaseline visit. (EASI50 is defined as achieving ≥ 50% improvement in EASI score.) • The binary response status of EASI90 from baseline at each postbaseline visit. (EASI90 is defined as achieving ≥ 90% improvement in EASI score.) • The binary response status of both EASI75 from baseline and IGA-TS at each postbaseline visit. • Change from baseline at each postbaseline visit for the following: – AD-affected %BSA – EASI score – SCORAD score – Itch NRS score – Skin Pain NRS score • Time to open-label escape arm (defined as not achieving 50% improvement in EASI score from baseline at 2 consecutive visits at least 1 week apart) • The binary response status of participants concurrently meeting all of the following criteria at each postbaseline visit: IGA score ≥ 3, EASI score ≥ 16, Itch NRS score ≥ 4, BSA ≥ 10%, and DLQI score > 10. • Time to concurrently meeting all of the following criteria: IGA score ≥ 3, EASI score ≥ 16, Itch NRS score ≥ 4, BSA ≥ 10%, and DLQI score > 10. • The binary response status of experience a relapse after study treatment discontinuation. (The binary response status of participants among EASI75 responders who are on study treatment at Week 24 who meet relapse criteria [loss of EASI50 from baseline] at the safety follow-up visit)

Table 1: Objectives and Endpoints (Continued)

Objectives	Endpoints
Secondary (continued)	
To further evaluate the efficacy of ruxolitinib cream in participants with moderate AD who had an inadequate response to, or are intolerant to, or contraindicated to TCSs and TCIs (continued).	• Time to first re-treatment during the VCE period. • Proportion of time off study treatment due to lesion clearance during the VCE period by visit. • Proportion of time on study treatment during the VCE period by visit.
To evaluate quality of life and other PROs in participants with moderate AD who had an inadequate response to, or are intolerant to, or contraindicated to TCSs and TCIs.	• The binary response status of achieving ≥ 4point improvement in DLQI from baseline at each postbaseline visit. • Change from baseline in the following scores at each postbaseline visit: – DLQI – POEM – EQ-5D-5L – HADS – PROMIS Short Form – Sleep-Related Impairment (8a – 7-day recall) – PROMIS Short Form – Sleep Disturbance (8b – 7-day recall) • Change from baseline score at Weeks 8 and 24 in WPAI-AD.
[redacted]	

3. STUDY DESIGN

This is a randomized, double-blind, VC study with an extension period in adult (aged ≥ 18 years) participants with moderate AD who had an inadequate response to, or are intolerant to, or contraindicated to TCSs and TCIs. For participants who meet all study inclusion criteria and none of the exclusion criteria, study drug will be assigned at the Day 1/baseline visit. Approximately 225 participants will be randomized 2:1 to blinded treatment with ruxolitinib 1.5% cream BID or vehicle cream BID, with stratification by baseline EASI score (< 16 or ≥ 16) and geographic region (ROW or Europe).

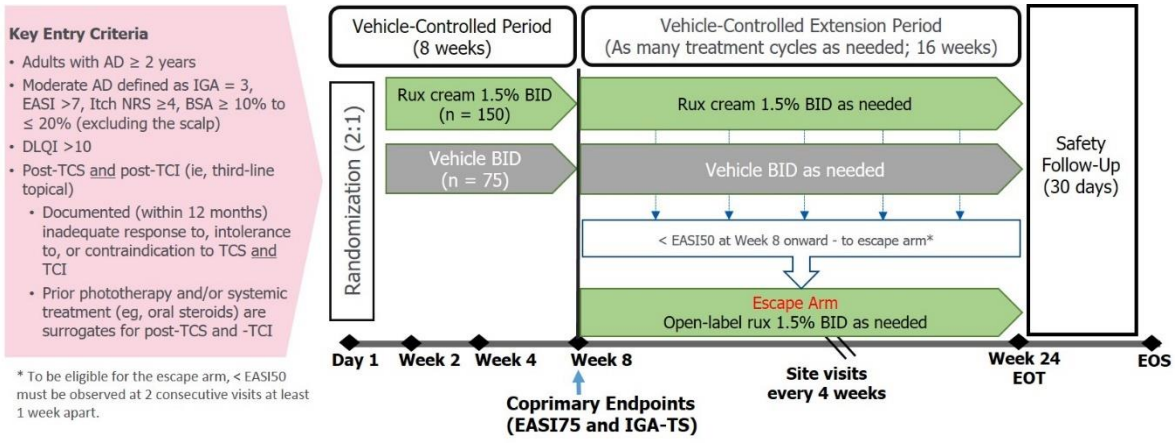
During the 8-week VC period, participants will apply study cream BID to all areas identified for treatment at the Day 1/baseline visit even if the AD begins to improve and lesions decrease in size. If there are new areas to be treated (including expansion of existing areas or development of new areas), after consultation with the investigator, study cream should be applied to these areas in addition to the areas identified at the baseline visit (up to a maximum total of 20% BSA) for the remainder of the VC period.

After completing 8 weeks of continuous treatment, all eligible participants with an adequate response, defined as achieving at least EASI50 from baseline, will continue blinded treatment allocation and will be evaluated during an additional 16-week VCE period for durability of response. Participants will be eligible to enter the ruxolitinib 1.5% cream open-label escape arm if EASI50 is not achieved at Week 8 or if an EASI50 response from baseline is lost during the VCE period (Note: < EASI50 must be observed at 2 consecutive visits, one of which can be an unscheduled visit, at least 1 week apart). During the VCE period, participants (including those in the escape arm) will have study visits every 4 weeks for up to 24 weeks total. At each of these visits, including Week 8, the participant's AD lesions will be evaluated by the investigator to determine whether the participant requires continuation of treatment or should (re)enter an observation/no treatment period. If the IGA score is ≥ 1, the participant will start or continue study cream BID. If the IGA score is 0 (clear), the participant will (re)enter an observation/no treatment period.

All participants will have a safety follow-up visit 30 days after their Week 24 or early termination visit. Refer to Protocol Amendment 2 for further details.

Figure 1 presents the study design schema.

Figure 1: Study Design Schema



3.1. Randomization

In the VC period, the interactive response technology system will assign approximately 225 participants in a 2:1 ratio to ruxolitinib 1.5% cream BID or vehicle cream BID. Additionally, participants will be stratified by baseline EASI score (< 16 or ≥ 16) and geographic region (ROW vs Europe).

All eligible participants with an adequate response, defined as achieving at least EASI50 from baseline, will continue into the 16-week VCE period with the same treatment regimen. Participants who did not achieve EASI50 at Week 8 or for whom an EASI50 response from baseline is lost during the VCE period will be eligible to enter the ruxolitinib 1.5% cream open-label escape arm.

3.2. Control of Type I Error

For the coprimary and key secondary analyses, an overall 2-sided Type I error of 0.05 will be used.

The following coprimary null hypotheses H1 and H2 will be tested:

Null hypotheses H1:

$$p_1 = p_2$$

Where p_1 and p_2 are the proportion of participants achieving EASI75 at Week 8 in the ruxolitinib 1.5% cream and vehicle cream group, respectively.

Null hypotheses H2:

$$p_3 = p_4$$

Where p_3 and p_4 are the proportion of participants achieving IGA-TS at Week 8 in the ruxolitinib 1.5% cream and vehicle cream group, respectively.

The key secondary endpoints will be tested in a fixed sequence as follows if the null hypotheses H1 and H2 for both coprimary endpoints (EASI75 and IGA-TS) are rejected:

- Proportion of participants with ITCH4 from baseline at Week 8
- Proportion of participants with ITCH4 from baseline on Day 7
- Proportion of participants with ITCH4 from baseline on Day 3
- Proportion of participants with ITCH4 from baseline on Day 2

3.3. Sample Size Considerations

Approximately 225 participants will be enrolled in this study and randomized 2:1 to ruxolitinib 1.5% cream or vehicle cream and stratified by baseline EASI score (< 16 or ≥ 16) and geographic region (ROW or Europe).

The sample size was calculated to provide sufficient power to detect a difference between the treatment groups in the coprimary and key secondary endpoints. The assumptions and powers for different endpoints are provided in Table 2. The assumed response rates are derived from the overall and subgroup analyses (population: baseline scores of Itch NRS ≥ 4, EASI > 7, post-TCS and -TCI, and IGA = 3) in 2 Phase 3 AD studies, INCB 18424-303 and INCB 18424-304. Fisher

exact test was used to provide a conservative evaluation of statistical power. Using a 2-sided α of 0.05, the sample size will have > 95% power to detect a difference in the coprimary endpoints, EASI75 and IGA-TS response rates between the treatment groups, and > 95% power for all key secondary endpoints.

In addition to providing sufficient power for efficacy variables, the sample size will provide an adequate database for safety evaluations.

Table 2: Powering for Coprimary and Key Secondary Endpoints

Variable	Response Rates in Ruxolitinib 1.5% BID	Response Rates in Vehicle	Power: Ruxolitinib vs Vehicle
EASI75	70%	15%	> 95%
IGA-TS	60%	10%	> 95%
ITCH4 at Week 8	50%	15%	> 95%
ITCH4 on Day 7	40%	12%	> 95%
ITCH4 on Day 3	35%	10%	> 95%
ITCH4 on Day 2	20%	3%	> 95%

3.4. Schedule of Assessments

Refer to Protocol Amendment 2 dated 08 AUG 2024 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 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 day 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 first application of ruxolitinib cream or vehicle cream for the VC and VCE DB periods.

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

For participants who enter the ruxolitinib 1.5% cream open-label escape arm in the VCE period, baseline for open-label escape arm is defined as follows:

- For efficacy evaluation, baseline is the last nonmissing measurement obtained before or on the day of first application of study treatment in the VC period.
- For safety evaluation, for participants who cross over from vehicle to ruxolitinib cream, baseline is the last nonmissing measurement obtained before or on the day of first application of ruxolitinib cream in the open-label escape arm; for participants on ruxolitinib cream in both periods, baseline is the last nonmissing measurement obtained before or on the day of first application of study treatment in the VC period.

When scheduled assessments and unscheduled assessments occur on the same day and time of the assessment or time of first application is 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 all scheduled assessments are missing on the day of the first application and an unscheduled assessment is available, use the unscheduled assessment as baseline.

4.1.4. Last Available Value

The last available value is the last nonmissing measurement obtained after starting ruxolitinib cream or vehicle cream and within 30 days after the last application of ruxolitinib cream or vehicle cream, or before the first application of ruxolitinib cream in the next period, whichever is earlier.

4.1.5. Handling of Missing and Incomplete Data

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.

For response endpoints, all nonresponders, as well as all participants who are missing postbaseline values at Week 8, will be defined as nonresponders for the NRI analysis.

For continuous endpoints, any participant who is missing postbaseline values may have missing data handled using MMRM or MI under the missing-at-random assumption. The MMRM model implicitly adjusts for missing data through a variance-covariance structure.

For other endpoints, missing observations will be handled as detailed in the specific sections addressing each analysis.

4.2. Variable Definitions

4.2.1. Body Mass Index

Body mass index will be calculated as follows:

Body mass index (kg/m^2) = (weight [kg] / (height [m])^2.

4.2.2. Prior and Concomitant Therapy

Prior therapy is defined as any nonstudy therapy started before the first application of study treatment.

Concomitant therapy is defined as any nonstudy therapy that is started accordingly:

- Before the date of first application of study treatment and is ongoing throughout the study or ends on/after the date of first application of study treatment
- On/after the date of first application of study treatment and is ongoing or ends during the course of study treatment

A prior therapy could also be classified as "both prior and concomitant therapy" if the end date is on or after the first application of study treatment. In the listing, it will be indicated whether a therapy is prior only, concomitant only, or both prior and concomitant.

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

5. STATISTICAL METHODOLOGY

5.1. General Methodology

Unless otherwise noted, SAS® software (SAS Institute Inc, Cary, NC; v9.1 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, maximum, first quartile, third quartile, and 95% CI. Descriptive summaries for categorical variables will include the number and percentage of participants in each category.

5.2. Treatment Groups

This is a randomized, double-blinded, VC study followed by a VCE period. Data will be summarized based on treatment regimen that was assigned (ITT) or that the participant actually applied (safety).

During the VC and VCE DB periods, the treatment groups will be 1.5% BID and vehicle.

For the escape arm in the VCE period, the treatment groups will be 1.5% BID and vehicle to 1.5% BID.

5.3. Analysis Populations

5.3.1. Intent-to-Treat Population

All participants who are randomized to the study 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 treatment the participant might apply during their participation in the study.

The ITT population will be used for the summary of all efficacy data in the VC and VCE DB periods.

5.3.2. Per Protocol Population

Participants in the ITT population who are considered to be sufficiently compliant with the Protocol compose the PP population, which is defined for supportive sensitivity analyses for efficacy endpoints in the treatment period. Participants with important Protocol deviations, as defined in Section 6.3, will be excluded from the PP population.

5.3.3. Safety Population

All randomized participants who applied ruxolitinib cream or vehicle cream at least once will constitute the safety population. Treatment groups for this population will be determined according to the actual treatment the participant applied on Day 1 regardless of assigned study treatment.

All safety analyses will be conducted using the safety population.

5.3.4. Pharmacokinetic/Pharmacodynamic-Evaluable Population

The PK-evaluable population includes participants who applied ruxolitinib 1.5% cream at least once and provided at least 1 postbaseline blood sample for PK analysis. The study pharmacokineticist will review data listings of participant administration and sample records to identify participants to be excluded from the analysis.

The PD-evaluable population includes participants who applied study cream at least once and provided at least 1 postbaseline sample for PD analysis. The study translational scientist will review data listings of participant administration and sample records to identify participants to be excluded from the analysis.

The PK/PD-evaluable population includes all participants who are in both PK-evaluable and PD-evaluable populations, as well as participants who are PD-evaluable while receiving vehicle cream during the VC and VCE periods.

5.3.5. Vehicle-Controlled Extension Primary Population

The VCE primary population includes all participants who entered the VCE period and applied study cream at least once during the VCE DB period. Participants will be analyzed according to the treatment assigned at randomization. For participants who enter the escape arm, only information prior to entry to the escape arm will be presented.

5.3.6. Vehicle-Controlled Extension Escape Population

The VCE escape population includes all participants who moved to the ruxolitinib 1.5% cream BID open-label escape arm. All efficacy and safety analyses for the escape arm will be conducted with the VCE escape population.

5.3.7. Vehicle-Controlled Extension Re-Treatment Evaluable Population

The VCE re-treatment evaluable population includes all participants who entered the VCE DB period.

6. BASELINE, EXPOSURE, AND DISPOSITION VARIABLES AND ANALYSES

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

6.1. Baseline and Demographics, Physical Characteristics, and Disease History

6.1.1. Demographics

The following demographics will be summarized for the ITT population, the VCE primary population, and the VCE escape population: age, age group, sex, race, ethnicity, weight, height, and body mass index.

6.1.2. Baseline Disease Characteristics

Baseline disease characteristics summarized for the ITT population, the VCE primary population, and the VCE escape population include but are not limited to the following:

- Time since first onset of AD
- Prior history of asthma (no/yes)
- Prior allergies (food, environmental) (no/yes)
- History of contact dermatitis (no/yes)
- Common complications of AD
- Time since onset of current AD episode
- Facial and/or neck involvement (no/yes) during past episodes
- Number of AD episodes/flare-ups over the last 12 months
- Average duration of episodes/flare-ups over the last 12 months
- Baseline EASI score
- Baseline %BSA
- Baseline DLQI score
- Baseline Itch NRS score

- Baseline Skin Pain NRS score
- Inadequate response to TCS (no/yes)
- If inadequate response to TCS is no, prior systemic therapy over the last 12 months (no/yes)
- If inadequate response to TCS is no, prior phototherapy over the last 12 months (no/yes)
- Intolerance to TCS (no/yes)
- Contraindication to TCS (no/yes)
- Inadequate response to TCI (no/yes); if inadequate response to TCI is no, prior systemic therapy over the last 12 months (no/yes)
- If inadequate response to TCI is no, prior phototherapy over the last 12 months (no/yes)
- Intolerance to TCI (no/yes)
- Contraindication to TCI (no/yes)
- Prior systemic therapy (no/yes) over the last 12 months
- Prior phototherapy (no/yes) over the last 12 months
- Intertriginous involvement at baseline (no/yes)

6.1.3. Prior Therapies for Atopic Dermatitis

Prior medications for AD will be coded using the WHO Drug Dictionary and summarized by treatment group for the ITT population. The type of treatment and reason for discontinuation will be summarized as well. Prior nondrug therapies for AD will be listed.

6.1.4. Medical History

For participants in the ITT population during the VC 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 in the eCRF.

6.2. Disposition of Participants

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

The number and percentage of participants who were treated and completed the VCE DB period and escape arm, were ongoing treatment during the VCE DB period and escape arm, and discontinued treatment or withdrew from the study during the VCE DB period and escape arm with a primary reason for discontinuation will be summarized for the participants who entered VCE period. The number and percentage of participants who entered the escape arm during the VCE DB period will be summarized as well.

6.3. Protocol Deviations

In general, the following are important Protocol deviations that may significantly affect the primary analysis:

- Missing data for the coprimary endpoint
- Overall application compliance less than 80% during the VC period
- Did not apply any study cream in the VC period
- IGA assessment outside the 14-day window for the Week 8 (Day 57 ± 7) visit.

Participants with any of the deviations listed above will be excluded from the PP population. In addition, important Protocol deviations, such as those related to inclusion/exclusion criteria, discontinuation criteria, and use of prohibited concomitant medications, will be reviewed. Decisions as to whether any of these deviations warrant exclusion from the PP population will be made prior to unblinding.

Protocol deviations will be summarized by treatment groups in the VC, VCE DB, and escape arm in the VCE periods separately.

6.4. Exposure

For participants in the safety population during the VC period, VCE primary population in the VCE DB period, and VCE escape population for the participants who entered escape arm, descriptive statistics will be provided (ie, by treatment group for duration of treatment, average daily amount of cream applied [g] and total amount of cream applied [g]). Duration of treatment with ruxolitinib cream or vehicle cream is defined as the number of days from Day 1 to the last record of ruxolitinib cream or vehicle cream application in the specific period.

6.5. Study Drug Compliance

Overall compliance (%) for the application of ruxolitinib cream or vehicle cream during the VC period will be calculated for all participants in the safety population as follows:

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

where the total number of nonmissing applications is the total number of applications that the participant actually applied during the study. The total number of intended applications is the number of planned applications minus the number of interrupted applications.

6.6. Prior and Concomitant Medications

For participants in the ITT population during the VC period, prior and concomitant medications will be coded using the WHO Drug Dictionary and summarized as number and percentage of participants with prior and concomitant therapies by WHO drug class and WHO drug term. For participants in the VCE period, only concomitant medications will be summarized. Concomitant medications for AD will also be summarized by treatment groups in the VC and VCE periods.

7. EFFICACY

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

7.1. General Considerations

For all continuous variables, both the actual value and change and/or percentage from baseline (if available) will be analyzed.

All by-visit analyses will include the safety follow-up period if the data are available.

7.2. Efficacy Hypotheses

See Section 3.2 for the efficacy hypotheses that will be tested in this study.

7.3. Analysis of the Primary Efficacy Parameters

7.3.1. Primary Efficacy Measures

The coprimary endpoints are:

- Proportion of participants with EASI75 from baseline at Week 8.
- Proportion of participants with IGA-TS at Week 8.

7.3.1.1. Eczema Area and Severity Index Score

Atopic dermatitis will be assessed using the EASI scoring system, which is a validated scoring system that grades the physical signs of AD to provide a measure of AD severity (ranging from 0 to 72).

The weight of the body regions for the EASI scoring is as follows:

head/neck (H) = 0.1, upper limbs (UL) = 0.2, trunk (T) = 0.3, and lower limbs (LL) = 0.4.

Each of the 4 body regions is assessed separately for erythema (E), induration/papulation/edema (I), excoriations (Ex), and lichenification (L) for an average degree of severity of each sign in each region with 0 = none, 1 = mild, 2 = moderate, and 3= severe, with half-steps allowed.

The disease severity strata for the EASI are as follows: 0 = clear; 0.1 to 1.0 = almost clear; 1.1 to 7.0 = mild; 7.1 to 21.0 = moderate; 21.1 to 50.0 = severe; 50.1 to 72.0 = very severe.

The categorical variable EASI75 to indicate the EASI responder will be set to 1 for percentage improvement from baseline in EASI score of 75% or greater and will be equal to 0 for percentage improvement of less than 75%. The EASI50 and EASI90 are defined with the same pattern.

7.3.1.2. Investigator's Global Assessment

The IGA is an overall eczema severity rating on a 0 to 4 scale that will be assessed during site visits. The grades for the IGA are shown in Table 3.

Table 3: Investigator's Global Assessment

Grade	Severity	Status
0	Clear	No erythema or induration/papulation, no oozing/crusting; there may be minor residual discoloration
1	Almost clear	There may be trace faint pink erythema, with almost no induration/papulation, and no oozing/crusting
2	Mild	There may be faint pink erythema, with mild induration/papulation and no oozing/crusting
3	Moderate	There may be pink-red erythema with moderate induration/papulation and there may be some oozing/crusting
4	Severe	There may be deep or bright red erythema with severe induration/papulation and with oozing/crusting

Source: FDA 2022.

The IGA-TS is defined as an IGA score of 0 or 1 with ≥ 2-grade improvement from baseline.

7.3.2. Primary Efficacy Analyses

The primary analysis will be based on the ITT population in the VC period. The primary alternative hypothesis (superiority of ruxolitinib 1.5% cream BID compared with vehicle cream BID) in both of the coprimary endpoints will be tested using a CMH test stratified by baseline EASI score and geographic region. The p-values and overall odds ratio with 95% CI between the ruxolitinib 1.5% cream BID and vehicle cream BID treatment groups will be provided. The p-values will be compared with the 2-sided α level of 0.05 for each of the coprimary endpoints. A summary of the coprimary endpoints EASI75 and IGA-TS rates will be reported for each treatment group. Stratum-adjusted rate differences (ruxolitinib 1.5% cream BID vs vehicle cream BID) on EASI75 and IGA-TS and the 95% CI will be computed using Mantel-Haenszel weights (Mantel and Haenszel 1959).

All nonresponders during the VC period, as well as all participants missing the Week 8 assessment and who discontinue study treatment at any time before Week 8, or discontinue from the study for any reason, will be defined as nonresponders for the nonresponder imputation analysis in both coprimary endpoints.

The coprimary endpoints will also be examined for the PP population using the same model as the primary analysis.

7.3.2.1. Primary Estimand

The proportion of participants within each treatment group who achieve the coprimary endpoints, assuming participants with missing Week 8 assessments due to any reasons are nonresponders, will be the primary estimand. The 5 attributes and the strategies associated with the defined intercurrent events are detailed in Table 4.

Table 4: Primary Estimand

Estimand Attribute	**Definition**		
Treatment effect	Ruxolitinib 1.5% cream BID vs vehicle cream BID		
Population	Participants with moderate AD who had an inadequate response to, or are intolerant to, or contraindicated to TCSs and TCIs (as defined by study inclusion/exclusion criteria)		
Variable	Apply to the coprimary endpoints		
Intercurrent events	**Events**	**Strategy**	**Rationale (if needed)**
	Treatment discontinuation due to any reason	Composite strategy	Participants who discontinue treatment due to any reason prior to Week 8 will be considered as treatment failures (ie, nonresponders) after the intercurrent event. Therefore, composite strategy is used for this type of intercurrent event.
Population-level summary	Stratum-adjusted response proportions difference between–treatment groups		

7.3.3. Subgroup Analyses for the Coprimary Endpoints

Subgroup analysis will be performed for the coprimary endpoints based on the following participant demographics and baseline disease characteristics variables for those participants whose data are available:

- Baseline EASI score (≥ 16, < 16)
- Age (18 to < 65 years, >= 65 years)
- Sex (male, female)
- Race
- Geographic region (Europe vs ROW)
- Prior systemic or phototherapy (yes or no)
- Season the participant was enrolled

Season the participant was enrolled will be defined as below:

- Spring: March to May (Northern Hemisphere) and September to November (Southern Hemisphere).
- Summer: June to August (Northern Hemisphere) and December to February (Southern Hemisphere).
- Fall: September to November (Northern Hemisphere) and March to May (Southern Hemisphere).
- Winter: December to February (Northern Hemisphere) and June to August (Southern Hemisphere).

7.3.4. Sensitivity and Supportive Analyses for the Coprimary Endpoints

7.3.4.1. Longitudinal Logistic Regression With Repeated Measures

To adjust for the dependence underlying the hierarchical multilevel data structure (visit, participant, and site), a longitudinal logistic regression with repeated measures will be applied in both of the coprimary endpoints. The 3 level structures in the model are as follows:

- Level 1: visit
- Level 2: participant
- Level 3: site

where visits are nested within participants, which are further nested within sites.

The binary response (IGA-TS or EASI75) of each participant at Week 2, Week 4, and Week 8 will be included as the dependent variable. Treatment (1.5% BID and vehicle BID), the randomization stratification factors (baseline EASI and geographic region), visit, and treatment-by-visit interaction will be included as fixed effects. Site level intercept and participant nested in site level intercept will be included as random effects. The within-participant and within-site errors will be modeled by an unstructured variance-covariance matrix. The Kenward-Roger approximation will be used to estimate denominator degrees of freedom for this model.

7.3.4.2. Multiple Imputation

Multiple imputation will be used as an alternative method to handle missing data. A full conditional specification method (van Buuren 2007) that assumes the existence of a joint distribution for all variables will be used to impute the IGA or EASI score. A regression model including treatment group, stratification factors and baseline and scheduled postbaseline scores up to Week 8 will be specified for the fully conditional specification method. The imputation will be repeated a number of times to generate corresponding complete datasets in order to reflect the uncertainty around the true values. An example of the multiple imputation SAS code is provided in Appendix B. The coprimary endpoints binary response will be derived for each of the imputed datasets per the definition in Sections 7.3.1.1 and 7.3.1.2. Regardless of MI imputed values, subjects with a missing Week 8 IGA and EASI assessments after discontinuation from treatment due to lack of efficacy or adverse event will be counted as nonresponders. The CMH test in Section 7.3.2 will be applied to each of the imputed datasets. The results will then be combined for the inference using Rubin's rule for each of the coprimary endpoints.

7.4. Analysis of the Key Secondary Efficacy Parameters

7.4.1. Key Secondary Efficacy Measures

7.4.1.1. Itch Numerical Rating Scale Score

The Itch NRS is a once-daily patient-reported measure of itch intensity assessed using an 11-point scale (0 = no itch to 10 = worst imaginable itch) that best describes their 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 directly before Day 1 (Day −7 to Day −1). The by-visit Itch NRS score for postbaseline visits will be determined by averaging the 7 daily NRS scores directly 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 daily itch related analyses, including time to achieve Itch NRS score improvement of at least 2 or 4 points analyses, baseline will be defined as the last available Itch NRS score during the week prior to Day 1 (Day −7 to Day −1).

The proportion of participants with a clinically relevant change in itch, defined as a ≥ 4-point improvement in Itch NRS score from baseline to Week 8, will be summarized by treatment group for participants with baseline Itch NRS score ≥ 4.

7.4.2. Key Secondary Efficacy Analysis

Key secondary efficacy analyses will be conducted in the ITT population in the VC period. If the coprimary endpoints are both achieved, the statistical comparisons for key secondary endpoints will be tested with the procedures specified in Section 3.2. The same nonresponder imputation method as specified in the primary analysis will be used to handle missing Week 8 itch data, and the CMH test described in Section 7.3.2 will be applied to compare 1.5% BID arm and the vehicle cream group for the key secondary endpoint of ≥ 4-point improvement in Itch NRS score response at Week 8. For the key secondary endpoints of Itch NRS score responses on Day 7, Day 3 and Day 2, all participants who are missing Day 7, Day 3, and/or Day 2 daily Itch NRS scores will be imputed using multiple imputation by fully conditional specification method. The variables to be included in the imputation regression model are treatment group, stratification factors, baseline itch score, and postbaseline daily itch score on Day 1 to Day 7. The corresponding binary response on Day 7, Day 3, and Day 2 will be derived for each of the imputed datasets per the definition in Section 7.4.1.1. The CMH test specified in Section 7.3.2 will be applied to each imputed dataset and then the results will be combined for the inference using Rubin's rule.

7.5. Analysis of Secondary Efficacy Parameters

7.5.1. Secondary Efficacy Measures

7.5.1.1. Body Surface Area

Total %BSA affected by AD will be estimated at each visit in the VC and VCE period. Body surface area assessment will be approximated to the nearest 0.1% using the handprint (Palmar) method as a guide. The approximate size of the participant's entire palmar surface (ie, the palm plus 5 digits, with fingers tucked together and thumb tucked to the side) should be considered as 1% BSA, and the approximate size of the participant's thumb should be considered as 0.1% BSA.

7.5.1.2. SCORing Atopic Dermatitis

The SCORAD is a tool to assess the extent and severity (ie, intensity) of eczema and will be completed before, during, and after treatment has begun to determine whether the treatment has been effective. This will be performed during all VC and VCE period study visits, starting at baseline.

- To determine extent, the rule of 9 or handprint method is used to calculate the eczema affected area (A) as a percentage of the whole body. Scores are added up to give a possible maximum of 100%.
- To determine intensity, a representative area of eczema is selected. The intensity of redness, swelling, oozing/crusting, scratch marks, skin thickening (lichenification), dryness (this is assessed in an area where there is no inflammation) is assessed individually as follows:
  - None (0)
  - Mild (1)
  - Moderate (2)
  - Severe (3)

  Intensity scores are added together to give "B" (maximum score of 18).
- Subjective symptoms, that is, itch and sleeplessness, are scored by the participant using a visual analog scale where "0" is no itch (or no sleeplessness) and "10" is the worst imaginable itch (or sleeplessness).

  These scores are added to give "C" (maximum score of 20).

Total score gives approximate weights of 60% to intensity and 20% each to extent and subjective signs (ie, insomnia) for the participant and will be calculated as follows: A/5 + 7B/2 + C.

7.5.1.3. Current Itch Numeric Rating Scale

On the morning of Day 1 (if the participant qualifies for randomization into the study) visit only, participants will be asked to evaluate the current intensity of their itch at the following assessment times: directly before initial study cream application and 5, 15, 30, 45, and 60 minutes as well as 2, 4, and 6 hours post–study cream application. The 6-hour assessment should be performed prior to the evening study cream application. Participants will be asked to rate the current itch severity of their AD by selecting a number from on a scale from 0 to 10, with 0 indicating no itch and 10 indicating the worst itch imaginable. Baseline of the current Itch NRS will be defined as the current itch score right before the initial study cream application.

7.5.1.4. Itch Numeric Rating Scale (7-Day Recall)

The 7-day recall Itch NRS is a daily participant-reported measure of the worst level of itch intensity. Participants will be asked to rate the itch severity of their AD by selecting a number from 0 (no itch) to 10 (worst itch imaginable) that best describes their worst level of itching in the past 7 days via the tablet during Protocol-defined clinic visits during the VCE period.

7.5.1.5. Skin Pain Numeric Rating Scale

The Skin Pain NRS is a daily participant-reported measure (24-hour or 7-day recall) of the worst level of pain intensity from 0 (no pain) to 10 (worst pain imaginable). Participants will be asked to rate the pain intensity by selecting a number that best describes the worst level of pain in the past 24 hours via the eDiary during the VC period and in the past 7 days via tablet at the site during the VCE period.

The baseline Skin Pain NRS score for VC period will be determined by averaging the 7 daily NRS scores directly before Day 1 (Day –7 to Day –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 directly 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 pain NRS score during the last week prior to Day 1 (from Day −7 to Day −1).

7.5.1.6. Dermatology Life Quality Index

The DLQI is a simple, 10-question validated questionnaire to measure how much the skin problem has affected the participant over the previous 7 days.

The scoring of each question is as follows: Very much = 3; A lot = 2; A little = 1; Not at all = 0; Not relevant = 0; Question 7, 'prevented work or studying' = 3.

The following imputation will be applied for incorrectly completed questionnaires:

- If one question is left unanswered, this is scored 0 and the scores are summed and expressed as usual out of a maximum of 30.
- If 2 or more questions are left unanswered the questionnaire is not scored.
- If Question 7 is answered 'yes' this is scored 3. If Question 7 is answered 'no' but then either 'a lot' or 'a little' is ticked, this is then scored 2 or 1. If "not relevant" is ticked, the score for Question 7 is 0. If it is answered 'no', but the second half is left incomplete, the score will remain 0.
- For DLQI 6 subscales, if the answer to one question in a subscale is missing, that subscale should not be scored.

The DLQI is calculated by summing the score of each question resulting in a maximum of 30 and a minimum of 0. The higher the score, the more quality of life is impaired.

The meaning of total DLQI scores can be categorized as follows:

- 0-1 = No effect
- 2-5 = Small effect
- 6-10 = Moderate effect
- 11-20 = Very large effect
- 21-30 = extremely large effect

A change from baseline in DLQI score of at least 4 points is considered clinically important (Basra et al 2015). The questionnaire is also analyzed under 6 subscales as follows:

- Symptoms and feelings (Questions 1 and 2)
- Daily activities (Questions 3 and 4)
- Leisure (Questions 5 and 6)
- Work and school (Question 7)
- Personal relations (Questions 8 and 9)
- Treatment (Question 10)

7.5.1.7. Patient-Oriented Eczema Measure

The POEM is a 7-question QoL assessment used for monitoring atopic eczema severity, focusing on the illness as experienced by the participant. The questionnaire asks how many days the participant has been bothered by various aspects of their skin condition during the past week. The response options and corresponding scoring questionnaires are no days (0), 1 to 2 days (1), 3 to 4 days (2), 5 to 6 days (3), and every day (4), with a range in total score of 0 to 28. If 1 question is missing, this question is scored 0 and the total score is calculated as summing the rest of the questions and expressed as usual out of a maximum of 28. If 2 or more questions are missing, the questionnaire is not scored (Charman et al 2004).

The meaning of POEM scores can be categorized as follows: 0 to 2 (clear/almost clear), 3 to 7 (mild eczema), 8 to 16 (moderate eczema), 17 to 24 (severe eczema), and 25 to 28 (very severe eczema; Charman et al 2013).

7.5.1.8. EQ-5D-5L

The EQ-5D is a validated, self-administered, generic utility questionnaire wherein participants rate their current health state based on the following criteria (dimensions): mobility, self-care, usual activities, pain/discomfort, and anxiety/depression.

The 5L indicates that for each dimension, there are 5 levels, which are as follows: no problems, slight problems, moderate problems, severe problems, and extreme problems.

During all VC period study visits (starting at Day 1) and at specific VCE visits, the participant will be asked to indicate his/her health state over the past 7 days.

Missing values will not be imputed. The categorical outcomes for the 5 dimensions (mobility, self-care, usual activity, pain/discomfort, and anxiety/depression) will be summarized. The change from baseline in EQ-5D VAS score will be summarized.

7.5.1.9. Work Productivity and Activity Impairment Questionnaire: Atopic Dermatitis

The WPAI-AD questionnaire is a validated 6-item instrument that measures the effect of overall health and specific symptoms on productivity at work and regular activities outside of it during the past 7 days (Reilly et al 1993).

The WPAI outcomes are expressed as impairment percentages, with higher numbers indicating greater impairment and less productivity (ie, worse outcomes) as follows:

Questions:

1 = currently employed

2 = hours missed due to AD

3 = hours missed other reasons

4 = hours actually worked

5 = degree AD-affected productivity while working

6 = degree AD-affected regular activities

Scores:

- Multiply scores by 100 to express in percentages.
  - Percent work time missed due to AD: Q2 / (Q2 + Q4)
  - Percent impairment while working due to AD: Q5 / 10
  - Percent overall work impairment due to AD: Q2 / (Q2 + Q4) + [(1 − (Q2 / (Q2+Q4)) × (Q5/10)]
  - Percent activity impairment due to AD: Q6 / 10

These impairment percentages will be summarized by visit and by treatment for those patients who are currently employed. Missing values will not be imputed.

7.5.1.10. PROMIS Sleep Questionnaires

The PROMIS Short Form – Sleep-Related Impairment(8a)/Sleep Disturbance (8b) are both 8-item questionnaires. Each item is rated on a 5-point scale (never, almost never, sometimes, almost always, and always) with a range in score from 8 to 40, with higher scores indicating greater severity of sleep-related impairment/disturbance. Both questionnaires are collected at each study visit starting at Day 1 and through the VC and VCE periods. The recall period will be 7 days. The raw scores on the 8 items should be summed to obtain a total raw score. If more than 4 of the items on the measure are missing, the total score should be set to missing.

After confirming that enough responses were provided, the raw score will be calculated by the below formula:

$$\frac{(Raw\ sum\ \times number\ of\ items\ on\ the\ short\ form)}{Number\ of\ items\ that\ were\ actually\ answered}$$

If the result is a fraction, then the result will be rounded up to the nearest whole number. The total score will be summarized by treatment and visit.

7.5.1.11. Hospital Anxiety and Depression Scale

The HADS is a 14-item questionnaire that assesses the levels of anxiety and depression a person is currently experiencing. There are 7 questions each for measuring anxiety and for measuring depression, with 4 possible responses to each question (responses are scored as 0, 1, 2, or 3).

Separate scores are calculated for anxiety and depression. The recall period will be the past 7 days for both the VC and VCE periods.

The score for the anxiety and depression subscale will be calculated separately as follows: anxiety score = sum of items 1*, 3*, 5*, 7, 9, 11*, 13*; and depression score = sum of items 2, 4, 6*, 8*, 10*, 12, 14 where starred items are reverse scored. Missing values will not be imputed. The change from baseline in anxiety and depression score will be summarized by treatment and visit.

7.5.2. Secondary Efficacy Analysis

All secondary efficacy analyses will be conducted in the VC and VCE DB periods, with the exception that proportion of participants achieving IGA-TS and EASI75 which will also be summarized by visit in the escape arm for participants that entered escape arm at Week 8.

7.5.2.1. Continuous Efficacy Endpoints

By-visit or by-timepoint summary statistics for the following continuous measurements, including change from baseline and percentage change from baseline, will be presented:

- EASI score
- Itch NRS score
- Current itch score on Day 1
- Skin Pain NRS score
- PROMIS Short Form – Sleep-Related Impairment 8a score
- PROMIS Short Form – Sleep Disturbance 8a score
- AD-affected %BSA
- Total POEM score
- Total DLQI score
- EQ-5D-5L score
- SCORAD score
- HADS score
- WPAI-AD score

Summary statistics including sample size, mean, median, standard deviation, minimum, maximum, first quartile, and third quartile will be presented by visit or timepoint. An MMRM may be fit for the comparisons between ruxolitinib 1.5% cream BID and the vehicle cream group at Week 8. The MMRM will include the fixed effect of baseline, treatment, stratification factors, the visit, and treatment-by-visit interaction. The variance-covariance matrix of the within-participant errors in MMRM will be modeled as unstructured.

For the above continuous measurements during the VCE DB period, only summary statistics will be presented.

7.5.2.2. Categorical Efficacy Endpoints

For the following categorical parameters, summary statistics including sample size, frequency, and percentages will be presented by visit or timepoint:

- Proportion of participants achieving EASI75
- Proportion of participants achieving an IGA-TS
- Proportion of participants with a ≥ 4-point improvement in Itch NRS score
- Proportion of participants achieving with a ≥ 2-point improvement in current Itch NRS score at 5, 15, 30, 45, and 60 minutes and 2, 4, and 6 hours post–initial dose on Day 1
- Proportion of participants achieving with a ≥ 4-point improvement in current Itch NRS score at 5, 15, 30, 45, and 60 minutes and 2, 4, and 6 hours post–initial dose on Day 1
- Proportion of participants achieving EASI50
- Proportion of participants achieving EASI90
- Proportion of participants achieving both EASI75 and IGA-TS
- Proportion of participants concurrently meeting all of the following criteria:

  IGA score ≥ 3, EASI score ≥ 16, Itch NRS score ≥ 4, BSA ≥ 10%, and DLQI score > 10
- Proportion of participants who achieve ≥ 4-point improvement in DLQI from baseline

A CMH test stratified by stratification factors may be performed at Week 8. Stratum-adjusted rate differences (ruxolitinib 1.5% cream BID vs vehicle cream BID) may be provided. The NRI may be used to impute postbaseline missing values in VC period for binary outcomes based on IGA, Itch NRS, and EASI scores.

7.5.2.3. Time-to-Event Efficacy Endpoints

The following time-to-event endpoints will be analyzed:

- Time to achieve ITCH4 during the VC period.
- Time to achieve ITCH2 during the VC period.
- Time to open-label escape arm

  (defined as not achieving 50% improvement in EASI score from baseline at 2 consecutive visits at least 1 week apart, with the earliest possible time to enter escape arm being at the Week 8 visit)
- Time to concurrently meeting all of the following criteria: IGA score ≥ 3, EASI score ≥ 16, Itch NRS score ≥ 4, BSA ≥ 10%, and DLQI score > 10.

For the time to open-label escape arm analysis, the start time will be set to the Week 8 date and only the participants who entered VCE period will be considered as at risk. Participants who were incorrectly assigned to escape arm without meeting the criteria will be censored at the date of their first dose in the escape arm.

A log-rank test stratified by randomization stratification factors will be used for between treatment group comparisons. The hazard ratio and its 95% CI will be estimated based on the stratified Cox regression model using Efron's method accounting for ties. Kaplan-Meier curves will be presented by treatment group. The number of participants, number of events, and number of censoring will be summarized by treatment group. 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).

7.5.2.4. Analysis of Desensitization and Rebound in the VCE DB Period

The following categorical parameters will be summarized by treatment to assess rebound upon treatment cessation:

- Proportion of participants who experience a relapse after study treatment discontinuation. (ie, proportion of participants among EASI75 responders who are on study treatment at Week 24 who meet relapse criteria [loss of EASI50 from baseline] at the safety follow-up visit)

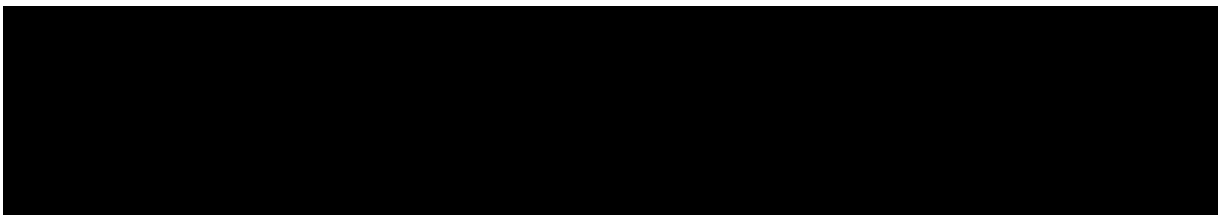
The following endpoints will be summarized by treatment and visit to assess desensitization upon treatment cessation:

- Proportion of time off study treatment due to lesion clearance during the VCE period by visit
- Proportion of time on study treatment during the VCE period by visit

Time to first re-treatment is defined as the interval between the first interruption due to lesion clearance starting from Week 8 to the end of the VCE period and the time of starting the re-treatment. This analysis will be performed in the VCE re-treatment evaluable population as defined in Section 5.3.7. The Kaplan-Meier estimate of median time to first re-treatment will be presented with its 95% CI. The 95% CI will be calculated using the method by Brookmeyer and Crowley (1982).

7.6. Analysis of Exploratory Efficacy Parameters

7.6.1. Exploratory Efficacy Analysis



[redacted]

[redacted]

9. SAFETY AND TOLERABILITY

Appendix A provides a list of planned tables, figures, and listings. 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 VC period, VCE primary population as well as VCE escape population in the VCE period. Cumulative TEAEs across the treatment periods (VC and VCE periods) will be summarized. 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 PTs 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 after the last dose of study drug. For participants who were initially on vehicle cream and crossed over to the ruxolitinib cream escape arm, the first application date is period specific, and 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 administration.

Adverse events will be tabulated by MedDRA PT and SOC. Severity of AEs will be graded using the National Cancer Institute CTCAE. The CTCAE v5.0 (Grade 1 to 5) is used for this study. 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 treatment-related. The incidence of AEs and treatment-related AEs will be tabulated. In addition, serious AEs 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.

Application site reactions are AEs that occur at the site of drug application. A summary of ASRs will be provided. Exposure-adjusted summaries will be provided for the safety population, including cumulative TEAEs and cumulative ASRs.

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 drug because of TEAEs
- Number (%) of participants who permanently discontinued study drug because of TEAEs
- Number (%) of participants who had a fatal TEAE

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 SOC and PT
- Summary of TEAEs by PT in decreasing order of frequency
- Summary of Grade 3 or higher TEAEs by SOC and PT
- Summary of Grade 3 or higher TEAEs by PT in decreasing order of frequency
- Summary of serious TEAEs by SOC and PT
- Summary of serious TEAEs by PT in decreasing order of frequency

- Summary of treatment-related TEAEs by SOC and PT
- Summary of treatment-related TEAEs by PT in decreasing order of frequency
- Summary of Grade 3 or higher treatment-related TEAEs by SOC and PT
- Summary of treatment-related serious TEAEs by SOC and PT
- Summary of TEAEs leading to temporarily interrupted study drug
- Summary of TEAEs with a fatal outcome by SOC and PT
- Summary of TEAEs leading to interruption of study drug by SOC and PT
- Summary of TEAEs leading to discontinuation of study drug by SOC and PT

Subgroup analysis for TEAEs by age, sex, ethnicity, and region will also be provided.

Adverse events of interest will be summarized in the following categories:

- Cytopenias
  - Anemia
  - Thrombocytopenia
  - Neutropenia
- Herpes zoster
- Viral skin infections
- Nonmelanoma skin neoplasms
- Liver function test elevations
- Malignancies
- Major adverse cardiovascular events
- Venous and arterial thromboembolic events
- Thrombocytosis and elevated mean platelet volume

9.3. Clinical Laboratory Tests

9.3.1. Laboratory Value Definitions

All laboratory assessments will be performed using a central laboratory except for urine pregnancy tests (as applicable). Laboratory values and change from baseline values will be summarized descriptively by visit, and non-numeric test values will be tabulated when necessary.

The baseline value will be determined using the last nonmissing value collected before the first application, prioritizing scheduled assessments for baseline identification over unscheduled visits if they were on the same date. The last record before administration with 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 numerical 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 nonmissing laboratory 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. This shift summary will be produced for each test for the safety population in the VC and VCE period. Shift tables will be presented showing change in CTCAE grade from baseline to worst grade postbaseline as well. The denominator for the percentage calculation will use the number of participants in the baseline category.

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, respiratory rate, and body temperature, will be summarized descriptively.

Criteria for clinically notable vital sign abnormalities are defined in Table 5. The abnormal values for participants exhibiting clinically notable vital sign abnormalities will be listed along with their assigned treatment group. Alert vital signs are defined as an absolute value outside the defined range and percentage change from baseline > 25%. The abnormal values for participants exhibiting alert vital sign abnormalities will be listed.

Table 5: Criteria for Clinically Notable Vital Sign Abnormalities

Parameter	High Threshold	Low Threshold
Systolic blood pressure	> 155 mm Hg	< 85 mm Hg
Diastolic blood pressure	> 100 mm Hg	< 40 mm Hg
Pulse	> 100 bpm	< 45 bpm
Temperature	> 38.0°C	< 35.5°C
Respiratory rate	> 20 breaths/min	< 8 breaths/min

10. INTERIM ANALYSES

No formal interim analysis is planned for this study. There are 2 formal planned analyses:

- The primary analysis will occur after the primary database lock, when all participants have completed the VC period (ie, up to Week 8). The sponsor will be unblinded after the primary database lock; however, investigators and participants will remain blinded to the individual study treatment assignment.
- 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 6.

Table 6: Statistical Analysis Plan Versions

SAP Version	Date
Original	06 MAY 2025

11.1. Changes to Protocol-Defined Analyses

All endpoints in Section 2.2 have been rephrased from the Protocol in a manner that is consistent with estimand guidance. The rephrasing is a change in presentation only and does not alter the interpretation or intended analyses of these endpoints from the Protocol.

The definition of VCE primary population and VCE escape population in Sections 5.3.5 and 5.3.6 has been modified in a more accurate language.

11.2. Changes to the Statistical Analysis Plan

Not applicable.

12. REFERENCES

Basra MK, Salek MS, Camilleri L, Sturkey R, Finlay AY. Determining the minimal clinically important difference and responsiveness of the Dermatology Life Quality Index (DLQI): further data. Dermatology 2015;230:27-33.

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

Charman CR, Venn AJ, Ravenscroft JC, Williams HC. Translating Patient-Oriented Eczema Measure (POEM) scores into clinical practice by suggesting severity strata derived using anchor-based methods. Br J Dermatol 2013;169:1326-1332.

Charman CR, Venn AJ, Williams HC. The patient-oriented eczema measure: development and initial validation of a new tool for measuring atopic eczema severity from the patients' perspective. Arch Dermatol 2004;140:1513-1519.

Food and Drug Administration. Draft Guidance on Pimecrolimus. October 2022.

Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst 1959;22:719-748.

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

van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. Stat Methods Med Res 2007;16:219-242.

APPENDIX A. PLANNED TABLES, FIGURES, AND LISTINGS

This appendix provides a list of the planned tables, figures, and listing for the clinical study report. Standard tables will follow the conventions in the Standard Safety Tables initial version. Shells are provided for nonstandard tables. In-text tables are identical in structure and content as appendix tables but follow a Rich Text Format.

The list of tables, figures, listings, and the shells are to be used as a guideline. Modifications of the list or shells 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 Demographic and Characteristic			
1.1.1	Analysis Populations	All	X
1.1.2.1	Summary of Participant Disposition in the VC Period	ITT	X
1.1.2.2	Summary of Participant Disposition in the VCE Period	VCE Primary	X
1.1.2.3	Summary of Participant Disposition in the Escape Arm	VCE Escape	X
1.1.3	Summary of Number of Participants Enrolled by Country and Site	ITT	X
1.1.4.1	Summary of Protocol Deviations	ITT	X
1.1.4.2	Summary of Protocol Deviations	VCE Primary	X
1.1.4.3	Summary of Protocol Deviations	VCE Escape	X
1.2.1	Summary of Demographics and Baseline Characteristics	ITT	X
1.2.2	Summary of Demographics and Baseline Characteristics	VCE Primary	X
1.2.3	Summary of Demographics and Baseline Characteristics	VCE Escape	X
1.3.1	Summary of Baseline Disease Characteristics	ITT	X
1.3.2	Summary of Baseline Disease Characteristics	VCE Primary	X
1.3.3	Summary of Baseline Disease Characteristics	VCE Escape	X
1.4.1	Summary of Prior Medications	ITT	X
1.4.2.1	Summary of Prior Medications for Atopic Dermatitis	ITT	X
1.4.2.2	Summary of Prior Medications for Atopic Dermatitis by Therapy Type and Discontinuation Reason	ITT	X
1.4.3.1	Summary of Concomitant Medications During the VC Period	ITT	X
1.4.3.2	Summary of Concomitant Medications During the VCE Period	VCE Primary	X
1.4.3.3	Summary of Concomitant Medications During the Escape Arm	VCE Escape	X
1.4.3.4	Summary of Concomitant Medications for AD During the VC Period	ITT	X
1.4.3.5	Summary of Concomitant Medications for AD During the VCE Period	VCE Primary	X
1.4.3.6	Summary of Concomitant Medications for AD During the Escape Arm	VCE Escape	X
1.5	Summary of General Medical History	ITT	X
Efficacy			
IGA			
2.1.1	Summary and Analysis of Participants Achieving IGA-TS in the DB Treatment Period	ITT	
2.1.2	Summary and Sensitivity Analysis of Participants Achieving IGA-TS at Week 8 by Multiple Imputations	ITT	
2.1.3	Summary and Analysis of Participants Achieving IGA-TS in the VC Period	PP	
2.1.4	Summary and Analysis of Participants Achieving IGA-TS by Baseline EASI Score (< 16 or ≥ 16) in the DB Treatment Period	ITT	
2.1.5	Summary and Analysis of Participants Achieving IGA-TS by Age in the DB Treatment Period	ITT	

Table No.	Title	Population	Standard
2.1.6	Summary and Analysis of Participants Achieving IGA-TS by Sex in the DB Treatment Period	ITT	
2.1.7	Summary and Analysis of Participants Achieving IGA-TS by Race in the DB Treatment Period	ITT	
2.1.8	Summary and Analysis of Participants Achieving IGA-TS by Region (ROW or Europe) in the DB Treatment Period	ITT	
2.1.9	Summary and Analysis of Participants Achieving IGA-TS by Prior Systemic or Phototherapy in the DB Treatment Period	ITT	
2.1.10	Summary and Analysis of Participants Achieving IGA-TS by Season Enrolled in the DB Treatment Period	ITT	
2.1.11.1	Summary and Analysis of Participants Achieving an IGA Score of 0 or 1 in the DB Treatment Period	ITT	
2.1.12.1	Summary of Participants in Each Category of IGA Score in the DB Treatment Period	ITT	
2.1.13.1	Summary of Participants Achieving IGA-TS in the Escape Arm	VCE Escape (Entered escape arm at Week 8)	
EASI			
2.2.1	Summary and Analysis of Participants Achieving EASI75 in the DB Treatment Period	ITT	
2.2.2	Summary and Sensitivity Analysis of Participant Achieving EASI75 at Week 8 by Multiple Imputations	ITT	
2.2.3	Summary and Analysis of Participants Achieving EASI75 in the VC Period	PP	
2.2.4	Summary and Analysis of Participants Achieving EASI75 by Baseline EASI Score in the DB Treatment Period	ITT	
2.2.5	Summary and Analysis of Participants Achieving EASI75 by Age in the Blinded Treatment Period	ITT	
2.2.6	Summary and Analysis of Participants Achieving EASI75 by Sex in the DB Treatment Period	ITT	
2.2.7	Summary and Analysis of Participants Achieving EASI75 by Race in the DB Treatment Period	ITT	
2.2.8	Summary and Analysis of Participants Achieving EASI75 by Region (ROW or Europe) in the DB Treatment Period	ITT	
2.2.9	Summary and Analysis of Participants Achieving EASI75 by Prior Systemic or Phototherapy in the DB Treatment Period	ITT	
2.2.10	Summary and Analysis of Participants Achieving EASI75 by Season Enrolled in the DB Treatment Period	ITT	
2.2.11	Summary and Analysis of Participants Achieving EASI50 in the DB Treatment Period	ITT	
2.2.12	Summary and Analysis of Participants Achieving EASI90 in the DB Treatment Period	ITT	
2.2.13.1	Summary and Analysis of By-Visit EASI Score in the DB Treatment Period	ITT	
2.2.14.1	Summary of Participants in Each Category of EASI Score in the DB Treatment Period	ITT	
2.2.15	Summary and Analysis of Participants achieving both EASI 75 and IGA-TS in the DB Treatment Period	ITT	
2.2.16	Summary and Analysis of Time to Open-Label Escape Arm	Participants who entered VCE period	
2.2.17.1	Summary and Analysis of Time to concurrently having IGA Score $\geq$ 3, EASI Score $\geq$ 16, Itch NRS Score $\geq$ 4, BSA $\geq$ 10%, and DLQI Score > 10	ITT	

Table No.	Title	Population	Standard
2.2.17.2	Summary of Participants concurrently having IGA Score ≥ 3, EASI Score ≥ 16, Itch NRS Score ≥ 4, BSA ≥ 10%, and DLQI Score > 10 in the DB Treatment Period	ITT	
2.2.18	Summary of Participants Achieving EASI75 in the Escape Arm	VCE Escape (Entered escape arm at Week 8)	
Itch NRS Score			
2.3.1	Summary and Analysis of Participants Achieving ≥ 4-Point Improvement in Itch NRS Score in the VC Period	ITT	
2.3.2	Summary and Analysis of Participants with ITCH4 From Baseline to Day 7 in the VC Period	ITT	
2.3.3	Summary and Analysis of Participants With ITCH4 From Baseline to Day 3 in the VC Period	ITT	
2.3.4	Summary and Analysis of Participants With ITCH4 From Baseline to Day 2 in the VC Period	ITT	
2.3.5.1	Summary and Analysis of By-Visit Itch NRS Score in the VC Period	ITT	
2.3.5.2	Summary of By-Visit Itch NRS Score in the VCE DB Treatment Period	VCE Primary	
2.3.6	Summary and Analysis of Time to ≥ 2-Point Improvement in Itch NRS Score in the VC Period	ITT	
2.3.7	Summary and Analysis of Time to ≥ 4-Point Improvement in Itch NRS Score in the VC Period	ITT	
2.3.8	Summary and Analysis of Daily Itch NRS Score in the VC Period	ITT	X
SCORAD			
2.4.1.1	Summary and Analysis of By-Visit SCORAD Total Score in the DB Treatment Period	ITT	
Current Itch NRS			
2.5.1	Summary and Analysis of By-timepoint Current Itch on Day 1	ITT	
2.5.2	Summary and Analysis of Participants Achieving ≥ 2-Point Improvement in Current Itch NRS Score on Day 1	ITT	
2.5.3	Summary and Analysis of Participants Achieving ≥ 4-Point Improvement in Current Itch NRS Score on Day 1	ITT	
Skin Pain NRS Score			
2.6.1.1	Summary and Analysis of By-Visit Skin Pain NRS Score in the VC Period	ITT	
2.6.1.2	Summary and Analysis of By-Visit Pain NRS Score in the VCE DB Treatment Period	VCE Primary	
2.6.2	Summary and Analysis of Daily Skin Pain NRS Score in the VC Period	ITT	X
DLQI			
2.7.1	Summary and Analysis of By-Visit Total DLQI Score in the DB Treatment Period	ITT	
2.7.3	Summary and Analysis of Participants Achieving ≥ 4-Point Improvement in DLQI From Baseline in the DB Treatment Period	ITT	
Patient-Oriented Eczema Measure			
2.8.1	Summary and Analysis of By-Visit Total POEM Score in the DB Treatment Period	ITT	
2.8.3	Summary of Participants in Each Category of POEM Score in the DB Treatment Period	ITT	
EQ-5D-5L			
2.9.1	Summary and Analysis of By-Visit EQ-5D-5L VAS Score in the DB Treatment Period	ITT	
2.9.3	Summary of Participants in Each Category of EQ-5D-5L Score in the DB Treatment Period	ITT	

Table No.	Title	Population	Standard
WPAI-AD			
2.10.1	Summary and Analysis of By-Visit WPAI-AD Score in the DB Treatment Period	ITT	
PROMIS			
2.11.1	Summary and Analysis of By-Visit PROMIS Short Form- Sleep-Related Impairment 8a Score in the DB Treatment Period	ITT	
2.11.3	Summary and Analysis of By-Visit PROMIS Short Form- Sleep Disturbance 8b Score in the DB Treatment Period	ITT	
HADS			
2.12.1	Summary and Analysis of By-Visit HADS Score in the DB Treatment Period	ITT	
BSA			
2.13.1	Summary and Analysis of Total %BSA Affected by AD in the DB Treatment Period	ITT	
2.13.3	Summary of By-Visit Change From Baseline in the Number of Intertriginous Regions With Active AD Involvement in the DB Treatment Period	ITT	
Desensitization and Rebound in the VCE DB Period			
2.14.1	Summary and Analysis of Participants Experiencing a Relapse After Study Treatment Discontinuation	Study completers who achieved EASI75 at Week 24 and entered safety follow-up	
2.14.2	Summary and Analysis of Participants' By-Visit Time Off Study Treatment Due to Lesion Clearance During VCE Period	VCE Primary	
2.14.3	Summary and Analysis of Participants' By-Visit Time on Study Treatment During the VCE Period	VCE Primary	
2.14.4	Summary and Analysis of Time to First Re-Treatment During the VCE Period	VCE Re-Treatment evaluable	
Safety			
Exposure			
3.1.1.X	Summary of Exposure During the VC Period	Safety	X
3.1.2	Summary of Exposure (Baseline to Week 24)	Safety	X
3.1.3	Summary of Study Drug Compliance During the VC Period	Safety	X
Adverse Events			
3.2.1.1.X	Overall Summary of Treatment-Emergent Adverse Events in the VC Period	Safety	X
3.2.1.2	Overall Summary of Treatment-Emergent Adverse Events for Participants Who Applied Ruxolitinib Cream Throughout the Study (Baseline to Week 24)	Safety	X
3.2.1.3	Overall Summary of Exposure-Adjusted Treatment-Emergent Adverse Events (Baseline to Week 24)	Safety	X
3.2.2.1.X	Summary of Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term in the VC Period	Safety	X
3.2.2.2	Summary of Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term for Participants Who Applied Ruxolitinib Cream Throughout the Study (Baseline to Week 24)	Safety	X
3.2.3.1.X	Summary of Treatment-Emergent Adverse Events by MedDRA Preferred Term in Decreasing Order of Frequency in the VC Period	Safety	X
3.2.3.2	Summary of Treatment-Emergent Adverse Events by MedDRA Preferred Term in Decreasing Order of Frequency for Participants Who Applied Ruxolitinib Cream Throughout the Study (Baseline to Week 24)	Safety	X

Table No.	Title	Population	Standard
3.2.3.3	Summary of Exposure-Adjusted Incidence Rates of Treatment-Emergent Adverse Events by MedDRA Preferred Term in Decreasing Order of Frequency (Baseline to Week 24)	Safety	X
3.2.3.4.X	Summary of Application Site Reactions by MedDRA Preferred Term in Decreasing Order of Frequency in the VC Period	Safety	X
3.2.3.5	Summary of Application Site Reactions by MedDRA Preferred Term in Decreasing Order of Frequency for Participants Who Applied Ruxolitinib Cream Throughout the Study (Baseline to Week 24)	Safety	X
3.2.3.6	Summary of Exposure-Adjusted Incidence Rates of Application Site Reactions by MedDRA Preferred Term in Decreasing Order of Frequency (Baseline to Week 24)	Safety	X
3.2.4.1.X	Summary of Grade 3 or Higher Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term in the VC Period	Safety	X
3.2.4.2	Summary of Grade 3 or Higher Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term for Participants Who Applied Ruxolitinib Cream Throughout the Study (Baseline to Week 24)	Safety	X
3.2.5.1.X	Summary of Grade 3 or Higher Treatment-Emergent Adverse Events by MedDRA Preferred Term in Decreasing Order of Frequency in the VC Period	Safety	X
3.2.5.2	Summary of Grade 3 or Higher Treatment-Emergent Adverse Events by MedDRA Preferred Term in Decreasing Order of Frequency for Participants Who Applied Ruxolitinib Cream Throughout the Study (Baseline to Week 24)	Safety	X
3.2.6.1.X	Summary of Serious Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term	Safety	X
3.2.6.2	Summary of Serious Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term for Participants Who Applied Ruxolitinib Cream Throughout the Study (Baseline to Week 24)	Safety	X
3.2.7.1.X	Summary of Serious Treatment-Emergent Adverse Events by MedDRA Preferred Term in Decreasing Order of Frequency in the VC Period	Safety	X
3.2.7.2	Summary of Serious Treatment-Emergent Adverse Events by MedDRA Preferred Term in Decreasing Order of Frequency for Participants Who Applied Ruxolitinib Cream Throughout the Study (Baseline to Week 24)	Safety	X
3.2.8.1.X	Summary of Treatment-Related Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term in the VC Period	Safety	X
3.2.8.2	Summary of Treatment-Related Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term for Participants Who Applied Ruxolitinib Cream Throughout the Study (Baseline to Week 24)	Safety	X
3.2.9.1.X	Summary of Treatment-Related Treatment-Emergent Adverse Events by MedDRA Preferred Term in Decreasing Order of Frequency in the VC Period	Safety	X
3.2.9.2	Summary of Treatment-Related Treatment-Emergent Adverse Events by MedDRA Preferred Term in Decreasing Order of Frequency for Participants Who Applied Ruxolitinib Cream Throughout the Study (Baseline to Week 24)	Safety	X
3.2.10.1.X	Summary of Grade 3 or Higher Treatment-Related Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term in the VC Period	Safety	X
3.2.10.2	Summary of Grade 3 or Higher Treatment-Related Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term for Participants Who Applied Ruxolitinib Cream Throughout the Study (Baseline to Week 24)	Safety	X
3.2.11.1.X	Summary of Treatment-Related Serious Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term in the VC Period	Safety	X

Table No.	Title	Population	Standard
3.2.11.2	Summary of Treatment-Related Serious Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term for Participants Who Applied Ruxolitinib Cream Throughout the Study (Baseline to Week 24)	Safety	X
3.2.12.1.X	Summary of Treatment-Emergent Adverse Events With a Fatal Outcome by MedDRA System Organ Class and Preferred Term in the VC Period	Safety	X
3.2.12.2	Summary of Treatment-Emergent Adverse Events With a Fatal Outcome by MedDRA System Organ Class and Preferred Term for Participants Who Applied Ruxolitinib Cream Throughout the Study (Baseline to Week 24)	Safety	X
3.2.13.1.X	Summary of Treatment-Emergent Adverse Events Leading to Dose Interruption by MedDRA System Organ Class and Preferred Term in the VC Period	Safety	X
3.2.13.2	Summary of Treatment-Emergent Adverse Events Leading to Dose Interruption by MedDRA System Organ Class and Preferred Term for Participants who Applied Ruxolitinib Cream Throughout the Study (Baseline to Week 24)	Safety	X
3.2.14.1.X	Summary of Treatment-Emergent Adverse Events Leading to Discontinuation of Study Drug by MedDRA System Organ Class and Preferred Term in the VC Period	Safety	X
3.2.14.2	Summary of Treatment-Emergent Adverse Events Leading to Discontinuation of Study Drug by MedDRA System Organ Class and Preferred Term for Participants Who Applied Ruxolitinib Cream Throughout the Study (Baseline to Week 24)	Safety	X
3.2.15.X	Summary of Treatment-Emergent Adverse Events of Interest by MedDRA System Organ Class and Preferred Term in the VC period	Safety	X
3.2.16.X	Summary of Treatment-Emergent Adverse Events of Interest by MedDRA Preferred Term in Decreasing Order of Frequency in the VC period	Safety	X
Laboratory			
3.3.1.X	Summary of Laboratory Values – Hematology in the VC Period	Safety	X
3.3.2.X	Summary of Laboratory Values – Chemistry in the VC Period	Safety	X
3.3.3.X	Shift Summary of Hematology Laboratory Values in CTCAE Grade – To the Worst Abnormal Value in the VC Period	Safety	X
3.3.4.X	Shift Summary of Chemistry Laboratory Values in CTCAE Grade – To the Worst Abnormal Value in the VC Period	Safety	X
3.3.5.X	Shift Summary of Hematology Values – To the Worst Abnormal Value in the VC Period	Safety	X
3.3.6.X	Shift Summary of Chemistry Values – To the Worst Abnormal Value in the VC Period	Safety	X
Vital Signs			
3.5.1.X	Summary of Systolic Blood Pressure in the VC Period	Safety	X
3.5.2.X	Summary of Diastolic Blood Pressure in the VC Period	Safety	X
3.5.3.X	Summary of Pulse in the VC Period	Safety	X
3.5.4.X	Summary of Respiratory Rate in the VC Period	Safety	X
3.5.5.X	Summary of Body Temperature in the VC Period	Safety	X

Note: Safety table that ends with X will be repeated for VCE primary population in the VCE DB period, and VCE escape population in the escape arm.

Figures

Figure No.	Title	Population
IGA		
4.1.1	Proportion of Participants Achieving IGA-TS	ITT
4.1.1.1	Proportion of Participants Achieving IGA-TS by Baseline EASI Score	ITT
4.1.1.2	Proportion of Participants Achieving IGA-TS by Age	ITT
4.1.1.3	Proportion of Participants Achieving IGA-TS by Sex	ITT
4.1.1.4	Proportion of Participants Achieving IGA-TS by Race	ITT
4.1.1.5	Proportion of Participants Achieving IGA-TS by Region	ITT
4.1.1.6	Proportion of Participants Achieving IGA-TS by Prior Systemic Therapy	ITT
4.1.1.7	Proportion of Participants Achieving IGA-TS by Season Enrolled	ITT
4.1.2	Forest Plot of Response Rate Difference in Achieving IGA-TS at Week 8	ITT
EASI		
4.2.1	Proportion of Participants Achieving EASI75	ITT
4.2.1.1	Proportion of Participants Achieving EASI75 by Baseline EASI Score	ITT
4.2.1.2	Proportion of Participants Achieving EASI75 by Age	ITT
4.2.1.3	Proportion of Participants Achieving EASI75 by Sex	ITT
4.2.1.4	Proportion of Participants Achieving EASI75 by Race	ITT
4.2.1.5	Proportion of Participants Achieving EASI75 by Region	ITT
4.2.1.6	Proportion of Participants Achieving EASI75 by Prior Systemic Therapy	ITT
4.2.1.7	Proportion of Participants Achieving EASI75 by Season Enrolled	ITT
4.2.2	Proportion of Participants Achieving EASI50	ITT
4.2.3	Proportion of Participants Achieving EASI90	ITT
4.2.4	Mean and Standard Error Plot of By-Visit EASI Score	ITT
4.2.5	Mean and Standard Error Plot of By-Visit Change From Baseline in EASI Score	ITT
4.2.6	Proportion of Participants Achieving Both EASI75 and IGA-TS	ITT
4.2.7	Kaplan-Meier plot of time to open-label escape arm	ITT
4.2.8	Forest Plot of Response Rate Difference in Achieving EASI75 at Week 8	ITT
ITCH NRS Score		
4.3.1	Proportion of Participants Achieving ≥ 4-Point Improvement in Itch NRS Score in the VC Period	ITT
4.3.2	Proportion of Participants with ITCH4 From Day 1 to Day 7 in the VC Period	ITT
4.3.3	Mean and Standard Error Plot of By-Visit Itch NRS Score in the VC period	ITT
4.3.4	Mean and Standard Error Plot of By-Visit Change From Baseline in Itch NRS Score in the VC period	ITT
4.3.5	Mean and Standard Error Plot of By-Visit Itch NRS Score in the VCE period	VCE Primary
4.3.5	Kaplan-Meier Plot of Time to ≥ 2-Point Improvement in Itch NRS Score in the VC Period	ITT
4.3.6	Kaplan-Meier Plot of Time to ≥ 4-Point Improvement in Itch NRS Score in the VC Period	ITT
4.3.7	Mean and Standard Error Plot of By-Visit Daily Itch NRS Score in the VC Period	ITT
4.3.8	Mean and Standard Error Plot of By-Visit Change from Baseline in Daily Itch NRS Score in the VC Period	ITT
SCORAD		
4.4.1	Mean and Standard Error Plot of By-Visit SCORAD Total Score	ITT
4.4.2	Mean and Standard Error Plot of By-Visit Change From Baseline in SCORAD Total Score	ITT
Current Itch NRS		
4.5.1	Mean and Standard Error Plot of By-Timepoint Current Itch on Day 1	ITT
4.5.2	Mean and Standard Error Plot of By-Timepoint Change From Baseline in Current Itch on Day 1	ITT

Figure No.	Title	Population
4.5.3	Proportion of Participants achieving ≥ 2-Point Improvement in Current Itch NRS Score	ITT
4.5.4	Proportion of Participants achieving ≥ 4-Point Improvement in Current Itch NRS Score	ITT
Skin Pain NRS Score		
4.6.1	Mean and Standard Error Plot of By-Visit Skin Pain NRS Score	ITT
4.6.2	Mean and Standard Error Plot of By-Visit Change From Baseline in Skin Pain NRS Score	ITT
4.6.3	Mean and Standard Error Plot of By-Visit Daily Skin Pain NRS Score in the VC Period	ITT
4.6.4	Mean and Standard Error Plot of By-Visit Change From Baseline in Daily Skin Pain NRS Score in the VC Period	ITT
DLQI		
4.7.1	Mean and Standard Error Plot of By-Visit Total DLQI Score	ITT
4.7.2	Mean and Standard Error Plot of By-Visit Change From Baseline in Total DLQI Score	ITT
4.7.3	Proportion of Participants Achieving ≥ 4-Point Improvement in DLQI From Baseline	ITT
Patient-Oriented Eczema Measure		
4.8.1	Mean and Standard Error Plot of By-Visit POEM Score	ITT
4.8.2	Mean and Standard Error Plot of By-Visit Change From Baseline in POEM Score	ITT
EQ-5D-5L		
4.9.1	Mean and Standard Error Plot of By-Visit EQ-5D-5L VAS Score	ITT
4.9.2	Mean and Standard Error Plot of By-Visit Change From Baseline in EQ-5D-5L VAS Score	ITT
WPAI-AD		
4.10.1	Mean and Standard Error Plot of By-Visit WPAI-AD Score	ITT
4.10.2	Mean and Standard Error Plot of By-Visit Change From Baseline in WPAI-AD Score	ITT
PROMIS		
4.11.1	Mean and Standard Error Plot of By-Visit PROMIS Short Form- Sleep-Related Impairment 8a Score	ITT
4.11.2	Mean and Standard Error Plot of By-Visit Change From Baseline in PROMIS Short Form- Sleep-Related Impairment 8a Score	ITT
4.11.3	Mean and Standard Error Plot of By-Visit PROMIS Short Form Sleep Disturbance 8b Score	ITT
4.11.4	Mean and Standard Error Plot of By-Visit Change From Baseline in PROMIS Short Form Sleep Disturbance 8b Score	ITT
Hospital Anxiety and Depression Scale		
4.12.1	Mean and Standard Error Plot of By-Visit HADS Score	ITT
4.12.2	Mean and Standard Error Plot of By-Visit Change From Baseline in HADS Score	ITT
BSA		
4.13.1	Mean and Standard Error Plot of By-Visit Total %BSA Affected by AD	ITT
4.13.2	Mean and Standard Error Plot of By-Visit Change From Baseline in Total %BSA Affected by AD	ITT
4.13.3	Mean and Standard Error Plot of By-Visit Change From Baseline in the Number of Intertriginous Regions With Active AD Involvement	ITT
Laboratory		
4.14.1.X	Box Plot of Selected Hematology Results by Visit in the VC Period	Safety
4.14.2.X	Box Plot of change from baseline in Selected Hematology Results by Visit in the VC Period	Safety
4.14.3.X	Box Plot of Percentage Change from Baseline in Selected Hematology Results by Visit in the VC Period	Safety
Vital Signs		
4.15.1.X	Box Plot of Systolic Blood Pressure in the VC Period	Safety
4.15.2.X	Box Plot of Diastolic Blood Pressure in the VC Period	Safety
4.15.3.X	Box Plot of Pulse in the VC Period	Safety

Figure No.	Title	Population
4.15.4.X	Box Plot of Respiratory Rate in the VC Period	Safety
4.15.5.X	Box Plot of Body Temperature in the VC Period	Safety

Listings

Listing No.	Title
Baseline Demographic and Characteristic	
2.1.1.1	Participant Enrollment and Disposition Status in the VC Period
2.1.1.2	Participant Enrollment and Disposition Status in the VCE Primary
2.1.1.3	Participant Enrollment and Disposition Status in the VCE Escape
2.1.2	Participant Inclusion and Exclusion Criteria Violations
2.2	Protocol Deviations
2.3	Analysis Populations
2.4.1	Demographic Characteristics
2.4.2	Baseline Disease Characteristics
2.4.3	Prior and Concomitant Medications
2.4.4	Prior Therapies for Atopic Dermatitis
2.4.5	Medical History
2.5.1	Study Drug Exposure and Compliance in the VC Period
2.5.2	Study Drug Exposure in the VCE Primary
2.5.3	Study Drug Exposure in the VCE Escape
2.5.4	Study Drug Exposure from Baseline to Week 24
Efficacy	
2.6.1	IGA Score
2.6.2	EASI
2.6.3.1	Itch NRS Daily Score
2.6.3.2	Itch NRS By-Visit Score
2.6.3.3	Time to Itch Response
2.6.4	SCORAD scores
2.6.5	Current Itch NRS
2.6.6.1	Skin Pain NRS Daily Score
2.6.6.2	Skin Pain NRS By-Visit Score
2.6.7	DLQI Score
2.6.8	POEM Score
2.6.9	EQ-5D-5L
2.6.10	WPAI-AD score
2.6.11	PROMIS Short Form– Sleep Disturbance 8b Score
2.6.12	PROMIS Short Form– Sleep-Related Impairment 8a Score
2.6.13	HADS Scores
2.6.14	Atopic Dermatitis Affected Percent BSA
2.6.15	Time to Open-Label Escape Arm
2.6.16	Time to First Re-Treatment in the VCE Period
2.6.17	Time to Concurrently Meeting All of the Following Criteria: IGA Score ≥ 3, EASI Score ≥ 16, Itch NRS Score ≥ 4, BSA ≥ 10%, and DLQI Score > 10
2.6.18	Relapse After Study Treatment Discontinuation
2.6.19	Proportion of Study On and Off Study Treatment During VCE Period
2.6.20	Intertriginous Regions With Active AD Involvement
Adverse Events	
2.7.1	Adverse Events
2.7.2	Adverse Events Leading to Study Drug Discontinuation

Listing No.	Title
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
2.7.7	Grade 3 or Higher Adverse Events
2.7.8	Application Site Reactions
2.7.9	Adverse Events of Interest
Laboratory Data	
2.8.1	Clinical Laboratory Values – Hematology
2.8.2	Clinical Laboratory Values – Chemistry
2.8.3	Abnormal Clinical Laboratory Values – Hematology
2.8.4	Abnormal Clinical Laboratory Values – Chemistry
Vital Signs	
2.10.1	Vital Signs
2.10.2	Abnormal Vital Sign Values
2.10.3	Alert Vital Sign Values

APPENDIX B. SAS CODE FOR MULTIPLE IMPUTATION

The following sample SAS code will be used for the MI of IGA:

```
proc mi data=mi_wide seed=184243261 nimpute=40 out=impute_IGA;
  class trt01p strat1 strat2;
       var trt01p strat1 Strat2 avalBASELINE avalWEEK2 avalWEEK4 avalWEEK8;
       fcs REGPMM;

run;
```

The following sample SAS code will be used for the MI of EASI:

```
proc mi data=mi_wide seed=184243262 nimpute=40 out=impute_EASI;
  class trt01p strat1 strat2;
       var trt01p strat1 Strat2 avalBASELINE avalWEEK2 avalWEEK4 avalWEEK8;
       fcs;
run;
```