Sponsor:	Arcutis Biotherapeutics, Inc.
Protocol Title:	A Phase 3, 8-Week, Parallel Group, Double Blind, Vehicle-Controlled Study of the Safety and Efficacy of ARQ-154 Foam 0.3% Administered QD in Subjects with Scalp and Body Psoriasis (ARRECTOR)
Protocol Number:	ARQ-154-309
Premier Research PCN:	ARCU214118
Document Version:	Final 3.0
Document Date	08-Nov-2022

APPROVALS

Role	Signatures	Date (dd-Mmm-yyyy)
Biostatistician	Print Name Sign Name:	
Premier Senior Reviewer	Print Name: Sign Name:	
Arcutis Biotherapeutics, Inc. Representative	Print Name: Sign Name:	
Arcutis Biotherapeutics, Inc. Representative	Print Name: Sign Name:	

DOCUMENT HISTORY

Final V1.0	31-May-2022 – Initial version.
Final V2.0	11-Aug-2022 – 1. Synopsis, Sections 2.2.2.2, 6.1.3, 8.2.4, and Table 5: a. Replace the change from baseline in PSD total score at Weeks 2, 4, and 8 in the list of secondary efficacy endpoints within Family 2 with the endpoints listed below to be consistent with Protocol Amendment 3: i. CFB in PSD Items related to Itching, Pain, and Scaling (Questions 1, 9, and 11) aggregate score at Week 8 ii. PSD Item related to Itching (Question 1) = 0 at Week 8 iii. PSD Item related to Scaling (Question 11) = 0 at Week 8 iv. PSD Item related to Pain (Question 9) = 0 at Week 8 b. Add PSD Total score = 0 to the list of endpoints in Family 3 to be consistent with changes in Protocol Amendment 3. c. Moved SI-NRS CFB at Day 1 from Family 1 to Family 3 to be consistent with changes in Protocol Amendment 3 2. Sections 2.2.2.3 and 8.3: a. Remove S-IGA score of Clear or Almost Clear because it would be redundant to the definition of S-IGA Success as subjects enrolled should have baseline S-IGA score of at least 3 b. Re-classify change from baseline in PSD total score at Weeks 2, 4, and 8 as an other efficacy endpoint c. Add B-IGA success at Weeks 2 and 4 as an other efficacy endpoint d. Add additional endpoints related to the new endpoints identified in 1.a.i-iv and 1.b above e. Add the Scalpdex emotion, function, and symptom scales 3. Table 6: Added to clarify the analyses completed for each of the other efficacy endpoints 4. Section 6.1.6: Removed “in each strata combination” as it is not intended nor feasible to have a subject in each strata combination. 5. Sections 6.1.2, 6.1.4.1, and 8: Clarify the intention to use S-IGA and B-IGA strata, not the actual value, used per randomization in the analysis models defined in PROC FREQ and PROC MIXED. 6. Section 6.1.1: Clarified that values collected with the same start date

	and time as the start of study drug application can be used as baseline. 7. Section 6.1.5.2 Table 3: Correct typographical error in Table regarding days to use for weekly average baseline score.
Final V3.0	This was a post-database lock and post-unblinding update to the SAP. The purpose is to clarify that for subjects who discontinued early from the study due to adverse event or lack of efficacy, a subject will be considered as non-responder (for MI and nonresponder imputation analyses) or missing (for observed case analyses) for all analysis visits (refer to Section 5.4) on or after the associated analysis visit of the subject's last dose day (refer to Section 6.1.5). This change to the primary estimand and multiple imputation procedure was based on recommendations from FDA during a pre-NDA meeting for a similar program held on 9/14/2022.

TABLE OF CONTENTS

1. OVERVIEW 11
2. STUDY OBJECTIVES AND ENDPOINTS 12
2.1. Study Objectives 12
2.1.1. Primary Objective 12
2.2. Study Endpoints 12
2.2.1. Safety Endpoints 12
2.2.2. Efficacy Endpoints 12
3. OVERALL STUDY DESIGN AND PLAN 15
3.1. Overall Design 15
3.2. Sample Size and Power 15
3.3. Study Population 15
3.4. Treatments Administered 16
3.5. Method of Assigning Subjects to Treatment Groups 16
3.6. Blinding and Unblinding 16
3.7. Schedule of Events 16
4. STATISTICAL ANALYSIS AND REPORTING 19
4.1. Introduction 19
4.2. Interim Analysis and Data Monitoring 19
5. ANALYSIS POPULATIONS 19
6. GENERAL ISSUES FOR STATISTICAL ANALYSIS 20
6.1. Statistical Definitions and Algorithms 20
6.1.1. Baseline 20
6.1.2. Adjustments for Covariates 21
6.1.3. Covariates for this study include the pooled site group (see Section 6.1.6), baseline S-IGA randomization strata (S-IGA = 3 vs. S-IGA =4) and baseline B-IGA randomization strata (B-IGA = 2 vs. B-IGA ≥ 3). Multiple Comparisons 21
6.1.4. Handling of Dropouts or Missing Data 22
6.1.5. Analysis Visit Windows 26
6.1.6. Pooling of Sites 27
6.1.7. Derived Variables 27
6.1.8. Data Adjustments/Handling/Conventions 31
6.2. Special Handling for COVID-19 Disruptions 32
7. STUDY SUBJECTS AND DEMOGRAPHICS 33
7.1. Disposition of Subjects and Withdrawals 33
7.2. Protocol Deviations 33
7.3. Demographics and Other Baseline Characteristics 33

7.4. Exposure and Compliance 34
8. EFFICACY ANALYSIS 35
8.1. Primary Efficacy Analyses 35
8.1.1. Scalp Investigator Global Assessment (S-IGA) 37
8.1.2. Body Investigator Global Assessment (B-IGA) 38
8.1.3. Hypothesis Testing 38
8.1.4. Sensitivity and Supplemental Analyses 38
8.1.5. Impact of Site 39
8.2. Secondary Efficacy Analyses 39
8.2.1. Scalp Itch – Numeric Rating Scale (SI-NRS) 39
8.2.2. Worst Itch – Numeric Rating Scale (WI-NRS) Pruritus Score 40
8.2.3. Psoriasis Area and Severity Index (PASI) 40
8.2.4. Psoriasis Symptoms Diary (PSD) 41
8.2.5. Psoriasis Scalp Severity Index (PSSI) 41
8.2.6. Scalp Investigator Global Assessment (S-IGA) 42
8.3. Other Efficacy Analyses 44
8.4. Subgroup Analyses of Efficacy Variables 48
9. SAFETY AND TOLERABILITY ANALYSIS 48
9.1. Adverse Events 49
9.1.1. Adverse Events Leading to Withdrawal 49
9.1.2. Deaths and Serious Adverse Events 49
9.2. Local Tolerability Assessments 50
9.3. Clinical Laboratory Evaluations 50
9.4. Vital Signs 50
9.5. PHQ-8 and Modified PHQ-A 50
9.6. C-SSRS 51
9.7. Concomitant Medication 51
9.8. Physical Examination 51
10. OTHER PLANNED ANALYSIS 51
10.1. Pharmacokinetic Analysis 51
11. CHANGES FROM PLANNED ANALYSIS 52
11.1. Version 2.0 52
11.2. Version 3.0 52
12. REFERENCES 52
14. TABLES, LISTINGS AND FIGURES 54
14.1. Planned Table Descriptions 54
14.2. Planned Figure descriptions 61

14.3. Planned Listing Descriptions 62

LIST OF TABLES

Table 1: Schedule of Events 17
Table 2: Analysis Visit Windows (Study Visit) for Non-diary Assessments 26
Table 3: Analysis Window for Calculation of Average Weekly SI-NRS and WI-NRS 27
Table 4: S-IGA and B-IGA 36
Table 5: Summary of Primary and Secondary Efficacy Analyses 42
Table 6: Summary of Analyses of Other Efficacy Endpoints 44

LIST OF ABBREVIATIONS

Abbreviation	Definition
AE	Adverse Event
ANCOVA	Analysis of Covariance
ATC	Anatomical Therapeutic Chemical
B-IGA	Body Investigator Global Assessment of Disease
BSA	Body Surface Area
CDISC	Clinical Data Interchange Standards Consortium
CDLQI	Children's Dermatology Life Quality Index
CFB	Change from baseline
CI	Confidence Interval
CMH	Cochran-Mantel-Haenszel
COVID-19	Novel Coronavirus Disease-19
eCRF	Case Report Form
CSR	Clinical Study Report
C-SSRS	Columbia-Suicide Severity Rating Scale
DLQI	Dermatology Life Quality Index
EMA	European Medicines Agency
EDC	Electronic Data Capture
FDA	U.S. Food and Drug Administration
ICH	International Conference on Harmonisation
IGA	Investigator Global Assessment
IP	Investigational Product
ITT	Intent to Treat
IWRS	Interactive Web Response System
LS-means	Least squares means
MCMC	Markov-Chain Monte-Carlo
MedDRA	Medical Dictionary for Regulatory Activities
NRS	Numeric Rating Score
PASI	Psoriasis Area and Severity Index

Abbreviation	Definition
PASI-75	Psoriasis Area and Severity Index-75; subjects who achieve a 75% reduction in PASI from baseline
PHQ-8	Patient Health Questionnaire
PHQ-A	Modified Patient Health Questionnaire for Adolescents
PK	Pharmacokinetics
PMM	Predictive Mean Matching
PP	Per Protocol
PSD	Psoriasis Symptoms Diary
PSSI	Psoriasis Scalp Severity Index
PSSI-75	Psoriasis Scalp Severity Index-75; subjects who achieve a 75% reduction in PSSI from baseline
PT	Prefered Term
QD	Once Daily ("quaque die")
Q1	First Quartile
Q3	Third Quartile
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Standard Deviation
SE	Standard Errors
S-IGA	Scalp Investigator Global Assessment of Disease
SI-NRS	Scalp Itch - Numeric Rating Scale
SOC	System Organ Class
TEAE	Treatment-emergent Adverse Event
TESAE	Treatment-emergent Serious Adverse Event
TLFs	Tables, Listings, and Figures
WI-NRS	Worst Itch - Numeric Rating Scale
WHO	World Health Organization

1. OVERVIEW

This statistical analysis plan (SAP) describes the planned analysis and reporting for Arcutis Biotherapeutics, Inc. protocol number ARQ-154-309 (A Phase 3, 8-Week, Parallel Group, Double Blind, Vehicle-Controlled Study of the Safety and Efficacy of ARQ-154 Foam 0.3% Administered QD in Subjects with Scalp and Body Psoriasis (ARRECTOR)), dated 29-Jul-2022, Amendment 3. Version 1 of the ARQ-154-309 SAP was based on protocol amendment 2 dated 12-May-2022 and was approved on 31-May-2022. Protocol Amendment 3, dated 29Jul2022, and SAP Version 2.0 (11-Aug-2022) were created to update the multiple testing strategy for secondary endpoints to reflect the change in the PSD secondary endpoints and to move SI-NRS change from baseline at Day 1 endpoint from Family 1 to Family 3 in the multiple testing strategy. The ARQ-154-309 database was locked on 09-Sep-2022 and unblinded on 12-Sep-2022. On 9-Sep-2022, FDA provided the following advice in preliminary feedback prior to a pre-NDA meeting for a similar program held on 14-Sep-2022.

> *For the primary estimand, you proposed a treatment policy strategy to handle all intercurrent events, including treatment discontinuation due to adverse event or lack of efficacy. For intercurrent events of treatment discontinuation due to adverse events or lack of efficacy, we recommend a composite strategy policy where subjects will be defined as non-responders, as we consider this to be the appropriate approach for handling such events. Your proposal to handle intercurrent events of treatment discontinuation using the treatment policy strategy can be used as part of a supportive estimand.*

The purpose of this version of the SAP is to document the change to the primary estimand and related multiple imputation strategies. The revised estimand will be the focus of the Clinical Study Report and NDA. Key results based on the original estimand defined in SAP versions 1 and 2 will be provided as supplemental information for the primary and secondary endpoints covered by the multiple testing strategy.

Reference materials for this statistical plan include the protocol and the accompanying sample data collection documents. Operational aspects related to collection and timing of planned clinical assessments are not repeated in this SAP unless relevant to the planned analysis.

The structure and content of this SAP provides sufficient detail to meet the requirements identified by the Food and Drug Administration (FDA), European Medicines Agency (EMA), and International Conference on Harmonization (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use: Guidance on Statistical Principles in Clinical Trials[1]. All work planned and reported for this SAP will follow internationally accepted guidelines, published by the American Statistical Association[3] and the Royal Statistical Society[1], for statistical practice.

The planned analyses identified in this SAP may be included in clinical study reports (CSRs), regulatory submissions, or future manuscripts. Also, post-hoc exploratory analyses not necessarily identified in this SAP may be performed to further examine study data. Any post-hoc or unplanned, exploratory analysis performed will be clearly identified as such in the final CSR.

The statistical plan described hereafter is an *a priori* plan. It will be submitted to file prior to any unblinded inferential or descriptive analysis of data pertaining to Arcutis Biotherapeutics, Inc.'s study ARQ-154-309.

ARQ-154 foam 0.3% will be described as "roflumilast foam 0.3%" throughout this document and the tables, listings, and figures (TLFs).

2. STUDY OBJECTIVES AND ENDPOINTS

2.1. Study Objectives

2.1.1. Primary Objective

The primary objective of this study is to assess the safety and efficacy of roflumilast foam 0.3% administered once daily (QD) vs. vehicle foam for 8 weeks in adolescent and adult subjects with scalp and body psoriasis.

2.2. Study Endpoints

2.2.1. Safety Endpoints

The safety endpoints of this study include the following:

- Adverse events (AEs)
- Local tolerability assessments
- Clinical laboratory parameters
- Vital signs/weight
- Patient Health Questionnaire depression scale (PHQ-8) and Modified PHQ-Adolescents (PHQ-A)
- Columbia-Suicide Severity Rating Scale (C-SSRS)
- Physical examinations

2.2.2. Efficacy Endpoints

2.2.2.1. Co-Primary Efficacy Endpoints

The co-primary efficacy endpoints of this study are:

- Scalp Investigator Global Assessment (S-IGA) Success at Week 8, defined as achievement of S-IGA score of "Clear" or "Almost Clear" plus at least 2-grade improvement from baseline.
- Body Investigator Global Assessment (B-IGA) Success at Week 8, defined as achievement of B-IGA score of "Clear" or "Almost Clear" plus at least 2-grade improvement from baseline.

2.2.2.2. Secondary Efficacy Endpoints

The secondary efficacy endpoints of this study include the following:

- For the subjects with Baseline Scalp Itch – Numeric Rating Scale (SI-NRS) score ≥ 4, achievement of at least 4-point improvement (reduction) from baseline in SI-NRS at Week 8 (SI-NRS Success at Week 8)
- SI-NRS Success at Week 4
- SI-NRS Success at Week 2
- SI-NRS Change from baseline (CFB) at Week 1
- SI-NRS CFB at 72 hours
- For the subjects with Baseline Worst Itch – Numeric Rating Scale (WI-NRS) score ≥ 4, achievement of at least a 4-point improvement (reduction) from baseline in WI-NRS at Week 8 ("WI-NRS Success at Week 8")
- Psoriasis Area and Severity Index (PASI)-75 at Week 8
- CFB in Psoriasis Symptoms Diary (PSD) Items related to Itching, Pain, and Scaling (Questions 1, 9, and 11) aggregate score at Week 8
- PSD Item related to Scaling (Question 11) = 0 at Week 8
- PSD Item related to Itching (Question 1) = 0 at Week 8
- PSD Item related to Pain (Question 9) = 0 at Week 8
- Psoriasis Scalp Severity Index (PSSI)-75 at Week 8
- S-IGA score of "Clear" at Week 8
- S-IGA Success at Week 4
- CFB in PASI at Week 2
- S-IGA Success at Week 2
- PSD Total Score = 0 at Week 8
- SI-NRS CFB at Day 1

2.2.2.3. Other Efficacy Endpoints

The efficacy endpoints listed below are not explicitly defined in the protocol but are based on other efficacy assessments collected:

S-IGA

- S-IGA score of "Clear" at Weeks 2 and 4
- Change from baseline in S-IGA at Weeks 2, 4, and 8

B-IGA

- B-IGA success at Weeks 2 and 4
- B-IGA score of "Clear" at Weeks 2, 4, and 8
- B-IGA score of "Clear" or "Almost Clear" at Weeks 2, 4, and 8
- Change from baseline in B-IGA at Weeks 2, 4, and 8

SI-NRS

- SI-NRS Success at Weeks 1, 3, 5, 6, and 7

- Change and percent change from baseline in SI-NRS at Weeks 2, 3, 4, 5, 6, 7, and 8

WI-NRS

- WI-NRS Success at Weeks 1, 3, 5, 6, and 7.
- Change and percent change from baseline in WI-NRS at Weeks 1, 2, 3, 4, 5, 6, 7, and 8

PASI

- PASI-75 at Weeks 2 and 4
- PASI-50, PASI-90, and PASI-100 at Weeks 2, 4, and 8
- Change from baseline in PASI at Weeks 4 and 8
- Percent change from baseline in PASI at Weeks 2, 4, and 8

PSD

- Change and percent change from baseline in PSD total score at Weeks 2, 4, and 8
- Change and percent change from baseline in PSD individual item scores at Week 2, 4, and 8
- Change and percent change from baseline in an aggregate score for a subset of 3 PSD individual items at Weeks 2 and 4. Items addressing Itching, Pain and Scaling (Questions 1, 9, and 11) will be aggregated into a single score and results of the aggregate (0-30) score will be analyzed.
- Percent change from baseline in PSD Items related to Itching, Pain, and Scaling (Questions 1, 9, and 11) aggregate score at Week 8
- PSD Item related to Scaling (Question 11) = 0 at Weeks 2 and 4
- PSD Item related to Itching (Question 1) = 0 at Weeks 2 and 4
- PSD Item related to Pain (Question 9) = 0 at Weeks 2 and 4
- Achievement of PSD Total Score = 0 at Weeks 2 and 4

PSSI

- PSSI-75 at Weeks 2 and 4
- PSSI-50, PSSI-90, and PSSI-100 at Weeks 2, 4, and 8
- Change and percent change from baseline in PSSI at Weeks 2, 4, and 8

Scalpdex

- Change and percent change from baseline in Scalpdex total score as well as the emotions, symptoms, and functioning scales at Weeks 2, 4, and 8

% BSA Affected by Psoriasis

- Change and percent change from baseline in total body % BSA affected by psoriasis at Weeks 2, 4, and 8

Extent of Scalp Involvement

- Change and percent change from baseline in percent of scalp affected by psoriasis at Weeks 2, 4, and 8

DLQI/CDLQI

- Change and percent change from baseline in Dermatology Life Quality Index (DLQI)/ Children's DLQI (CDLQI) total scores at Weeks 2, 4, and 8

3. OVERALL STUDY DESIGN AND PLAN

3.1. Overall Design

This is a phase 3, parallel group, double blind, vehicle-controlled study in which roflumilast foam 0.3% or vehicle foam is applied QD x 8 weeks to adolescent and adult subjects with a minimum S-IGA of "Moderate" (3) and a minimum B-IGA of "Mild" (2).

Approximately 420 subjects will be enrolled. Subjects will be randomized in a 2:1 ratio to roflumilast foam 0.3% QD or matching vehicle foam QD, which will be applied to all areas of scalp and body psoriasis (including the face, scalp, trunk, and intertriginous/genital regions; palms and soles may be treated but will not be counted towards any measurements of efficacy). Screening will be for up to 4 weeks followed by an 8-week Treatment Phase. The anticipated maximum duration of subject participation is approximately 12 weeks.

Subjects will have to apply the study drug once a day in the evening, except for Baseline/Day 1 and Week 2 (Visit 3) when the study drug is applied at the study site. Participants have to record the date and time each dose has been applied with details including missed doses and any additional comments.

3.2. Sample Size and Power

A sample size of up to approximately 420 subjects is planned for the study. Approximately 280 subjects will receive roflumilast foam 0.3% QD; approximately 140 subjects will receive vehicle foam QD.

This sample size provides more than 99% power to detect an overall 35% difference between treatment groups on S-IGA success at Week 8 (11% and 46% success rates in roflumilast and vehicle arms corresponding to an odds ratio=6.9). This sample size also provides more than 99% power to detect an overall 25% difference between treatment groups on B-IGA success at Week 8 (7% and 32% success rates in roflumilast and vehicle arms corresponding to an odds ratio of 6.4). Testing will be performed at α=0.025 using a 2-sided stratified (B-IGA at randomization, S-IGA at randomization and study site) Cochran-Mantel-Haenszel (CMH) test. The results from a recent phase 2 study (ARQ-154-204) of roflumilast foam 0.3% compared to vehicle treatment were used as the basis of the power calculations.

The number of subjects to be enrolled will also provide at least 90% power for most of the secondary endpoint analyses. The study is highly powered in order to provide sufficient numbers of subjects on roflumilast foam treatment for a safety database, and to provide robust results due to the single pivotal trial nature of this study.

3.3. Study Population

Study population consists of male and female adolescents (12-17 years old) and adults (≥18 years old). Subjects should have a minimum S-IGA of at least "Moderate" (3) and B-IGA

of at least “Mild” (2) at baseline.

3.4. Treatments Administered

Subjects who meet the eligibility criteria will be randomized to 1 of the 2 following treatment groups in a 2:1 ratio (active: vehicle):

- Roflumilast foam 0.3% QD
- Matching vehicle foam QD

3.5. Method of Assigning Subjects to Treatment Groups

Subjects are randomized and assigned to active treatment or vehicle at the Baseline/Day 1 visit. Assignment of roflumilast foam 0.3% QD or vehicle foam QD is made at a 2:1 ratio according to a computer-generated randomization list. Randomization is stratified by study site, baseline S-IGA (S-IGA = 3 vs. S-IGA = 4), and baseline B-IGA (B-IGA = 2 vs. B-IGA ≥ 3).

3.6. Blinding and Unblinding

This is a double-blinded study, therefore neither the subjects nor the Investigator, clinical personnel, or Sponsor are aware of which treatment an individual subject has received. Emergency unblinding will be done using the study internet-based randomization system (interactive web-response system; IWRS) system in consultation with the medical monitor. Refer to the current IWRS Plan for details on unblinding.

3.7. Schedule of Events

A detailed schedule of events for the study is provided in Table 1.

Table 1: Schedule of Events

Study Procedure	Screening	Baseline Day 1	Week 2 Day 15	Week 4 Day 29	Week 8[a] Day 57
Visit	**1**	**2**	**3**	**4**	**5**
Visit Window	**-4 Weeks**		**± 3 Days**	**± 5 Days**	**± 5 Days**
Informed consent/assent	X				
Medical/surgical history, demography	X				
Physical examination[b]	X	X			X
Fitzpatrick Skin Type	X				
I/E criteria	X	X			
Randomization		X			
Hematology, Chemistry, and Urinalysis[c]	X	X		X	X
Vital signs, weight, height[d]	X	X	X	X	X
S-IGA[e], B-IGA[e], PSSI[e], BSA[f], Extent of Scalp Involvement[g], PASI	X	X	X	X	X
DLQI/CDLQI[h], Scalpdex[h], PSD[h]	X	X	X	X	X
Daily SI-NRS and WI-NRS[i] at home (non-clinic)	Day -7 through Day 57				
Application Site Reaction Assessment/ Local Tolerability[j]		X	X	X	X
C-SSRS, PHQ-8/PHQ-A[k]	X	X		X	X
Medical Photography[l]		X	X	X	X
Follicle Stimulating Hormone (FSH)[m]	X				
Serum pregnancy test[m]	X				
Urine pregnancy test[n]		X		X	X
PK sampling[o]				X	X
IP application and subject/family training[p]		X	X	X	
Dispense investigational product kit[q]		X	X	X	
Dispense / review dosing diary		X	X	X	X
Weigh investigational product[r]		X	X	X	X
Compliance calculation[s]			X	X	X
Adverse event assessment[t]	X	X	X	X	X
Concomitant medications	X	X	X	X	X

***Footnotes from table above*:**

a. Subjects that terminate early should return to the study site for the Week 8 assessments.
b. Limited physical examination: skin (including assessment of Fitzpatrick Skin Type at Screening only), lungs, and heart only.
c. For subjects <18 years of age, if the Baseline/Day 1 visit occurs within 3 weeks of Screening, the Screening lab results may be utilized.
d. Height will be collected at Screening and Week 8. Weight will be collected at all study visits. Subjects should void prior to weight being taken and remove any objects of significant weight (i.e., jackets, outerwear, shoes, cell phones, wallet, key chains, etc.). A 5% unintentional weight loss from Baseline should be reported to the medical monitor.
e. S-IGA will be a 5-point scale ranging from clear (0) to severe (4) and is evaluated for the scalp only. B-IGA will use the same scale as S-IGA but will evaluate the entire body (except the scalp, palms, and soles). PSSI is scored on a 0-72 scale and evaluates the scalp only. PASI assessment should exclude palms and soles. **S-IGA and B-IGA should be completed prior to other Investigator assessments and every effort should be made for the same evaluator to complete at each study visit.**
f. Total BSA affected by scalp and body psoriasis will be determined by the subject's hand method, where the subject's hand (including fingers) surface area is assumed to equal 1% of body surface area. Scalp will be included in total BSA measurement, but palms and soles will be excluded. The total scalp (ie, 100% of the total scalp surface area) can be considered to be approximately 4% BSA.
g. Extent of scalp involvement will be the extent of psoriasis involved, expressed as a percentage of total scalp.
h. The DLQI will be completed by subjects ≥17 years of age. The CDLQI will be completed for subjects 12 to 16 years old, inclusive. Scalpdex will be completed by all subjects. PSD will be completed by subjects ≥ 18 years of age.
i. Subjects/parents/caregivers will complete the SI-NRS and WI-NRS daily at home starting 7 days prior to the Baseline/Day 1 visit to Day 57 (Week 8) visit. Daily SI-NRS and WI-NRS will be completed in the evening prior to IP application (except at Baseline and Week 2 when IP is applied at the clinic). **The SI-NRS should be completed first and then WI-NRS.**
j. Local tolerability assessments include both investigator and subject assessments. The Investigator assessment of skin irritation (Berger and Bowman skin irritation score) is to be recorded prior to IP application.
Note for Investigator tolerability assessments: reactions at the site of IP application, which may occur post-Baseline, should be differentiated from the preexisting inflammation associated with the subject's psoriasis. The subject's "0-3" burning/stinging assessment is to be recorded 10-15 minutes post IP application for the subject's "0-3" burning/stinging assessment. **At Week 8, subjects will provide a recall assessment of burning/stinging experienced post IP application on the day prior to the clinic visit.**
k. All subjects will complete the C-SSRS. Adults will complete the PHQ-8. Adolescents (ages 12 to 17 inclusive) will complete the PHQ-A (PHQ-9 modified). Refer screen failures due to score ≥10 to a mental health professional. For Scores 10-14 at all visits, refer subject promptly to a mental health professional. For scores ≥15 at all visits, refer subject promptly to a mental health professional and consider interruption of IP.
l. Medical photography will be obtained for target lesions by all sites. All efforts will be made to de-identify the subjects.
m. A serum pregnancy test will be administered to all females of child-bearing potential at the Screening visit only. Follicle Stimulating Hormone (FSH) will be performed at Screening (if indicated) to confirm post-menopausal status.
n. A urine pregnancy test will be performed at the Baseline, and Weeks 4 and 8 to all females of child-bearing potential. A negative result is required for continued participation in the study, and results must be available at each visit prior to dispensing of IP.
o. PK samples (trough) will be collected predose on Day 29 (Week 4) and Day 57 (Week 8). Ensure investigational product is not applied in the area where PK will be drawn.
p. Subjects/parents/caregivers to apply assigned IP in the study site at every designated visit.
q. Kits will be dispensed based on % BSA. See IP Handling Plan for details.
r. Each IP canister in the Kit should be weighed and recorded at every visit. See IP Handling Plan for details.
s. Compliance calculation is described Section 6.1.7 and in the IP Handling Plan.

***Footnote from table above*:**

[t.] All AEs should be collected starting after the first application of the IP through the end of the study. All SAEs should be collected starting after the signing of the informed consent through 30 days after the last day of the IP application or the end of the study (whichever is later). Any AEs (whether serious or non-serious) and clinically significant abnormal laboratory test value(s) will be evaluated by the Investigator and treated and/or followed up to 30 days after end of treatment or until symptoms or value(s) return to subject's Baseline, or acceptable level, as judged by the PI.

4. STATISTICAL ANALYSIS AND REPORTING

4.1. Introduction

Data processing, tabulation of descriptive statistics, calculation of inferential statistics, and graphical representations, will be performed primarily using SAS (release 9.4 or higher). If the use of other software is warranted for analyses conducted by Premier Research, the final statistical methodology report will detail what software was used for what purposes.

Continuous (quantitative) variable summaries will include the number of subjects (n) with nonmissing values, mean, standard deviation (SD), median, minimum, maximum, first quartile (Q1), and third quartile (Q3).

Categorical (qualitative) variable summaries will include the frequency and percentage of subjects who are in the particular category or each possible value. In general, the denominator for the percentage calculation will be based upon the total number of subjects in the study population for the treatment group, unless otherwise specified.

Unless otherwise indicated, all statistical tests will be conducted at the familywise α=0.025 significance level using 2-tailed tests, and *P* values will be reported. Corresponding 95% and 97.5% confidence intervals will be presented for statistical tests. Details for splitting the familywise α to control for multiple comparisons will be discussed in Section 6.1.3.

All data collected in the electronic case report form (eCRF) will be provided in subject listings.

4.2. Interim Analysis and Data Monitoring

No interim analyses are planned.

5. ANALYSIS POPULATIONS

The following analysis populations are planned for this study:

- **Safety Population (SAF):** The Safety population will include all subjects who are randomized and received at least 1 confirmed dose of investigational product (IP). This population will be used for all safety analyses.
- **Intent-To-Treat Population (ITT):** The ITT population will include all randomized subjects. This population will be the primary analysis population for the analysis of efficacy endpoints, along with subject disposition.

- **Per Protocol (PP) Population:** The PP population will include the subjects in the ITT population who were at least 80% compliant with study drug application, have an S-IGA and B-IGA assessment at Week 8, and showed no important deviations from the study protocol that would affect the interpretation of efficacy. "Important deviations" will result from a blinded review at the end of the study and the team will determine the important deviations that can impact on efficacy. This population will be used for a supplemental analysis for primary efficacy endpoints.
- **Modified Intent-to-Treat Population (mITT):** The mITT population includes all randomized subjects with the exception of subjects who missed the Week 8 S-IGA or B-IGA assessment specifically due to COVID-19 disruption. This population will be used for sensitivity of the primary endpoints.
- **Scalp Pruritus ITT Population (SPRU4-ITT):** The SPRU4-ITT population is a subset of the ITT population and includes subjects with average weekly SI-NRS score ≥ 4 at baseline. This population will be used for the analysis of achievement of at least 4-point reduction in average weekly SI-NRS score as compared to baseline.
- **Pruritus ITT Population (PRU4-ITT):** The PRU4-ITT population is a subset of the ITT population and includes subjects with average weekly WI-NRS score ≥ 4 at baseline. This population will be used for the analysis of achievement of at least 4-point reduction in average weekly WI-NRS score as compared to baseline.
- **Pharmacokinetic Population (PK):** The PK population will include all subjects who received at least one confirmed dose of investigational product (IP) and provided at least one PK sample. This population will be used for summary of PK concentration results for subjects who applied roflumilast foam 0.3% during this study.

6. GENERAL ISSUES FOR STATISTICAL ANALYSIS

6.1. Statistical Definitions and Algorithms

6.1.1. Baseline

In general, for assessments where time was recorded, the last observation recorded at or before the time of first application of study drug will be used as the baseline observation for all calculations of change from baseline. If the date and time of the assessment is equal to the date and time of the start of the first application of study drug, it will be assumed that the assessment was completed before study drug application and the value will be used as baseline. For assessments where time was not recorded (excluding SI-NRS and WI-NRS), the last observation recorded on or before the day of first application of study drug will be used as the baseline observation for all calculations of change from baseline. If the last non-missing assessment is performed on the same date as the first study treatment administration and time is not available, it is assumed that the assessment took place prior to IP application, per study site training, and the assessment will be considered as baseline, except adverse events (AEs) and medications starting on the first study treatment dose administration date which will be considered postbaseline.

For average weekly SI-NRS and WI-NRS, baseline is defined as the average of all non-missing

scores reported during the last 7 days prior to treatment if at least 4 of 7 evaluable daily questionnaires scores are available. For daily SI-NRS and daily WI-NRS baseline is defined as the last non-missing assessment prior to the first study treatment.

For subject tolerability assessments, baseline is derived as the last non-missing measurement taken on the day of first application of study drug.

6.1.2. Adjustments for Covariates

Covariates for this study include the pooled site group (see Section 6.1.6), baseline S-IGA randomization strata (S-IGA = 3 vs. S-IGA =4) and baseline B-IGA randomization strata (B-IGA = 2 vs. B-IGA ≥ 3).

6.1.3. Multiple Comparisons

Upon demonstration of concurrent statistical significance for S-IGA Success at Week 8 at the 2.5% level, and statistical significance of B-IGA Success at Week 8 at the 2.5% level, the testing scheme described below will be used to test these secondary endpoints. This testing scheme will control the overall type 1 error at the 0.025 level.

The overall α level of 0.025 will be split to test 3 families of secondary endpoints, and the Fallback Method will be applied across the families.

Family 1 (α=0.01 or 0.015):

- SI-NRS Success at Week 8
- SI-NRS Success at Week 4
- SI-NRS Success at Week 2
- SI-NRS change from baseline (CFB) at Week 1
- SI-NRS CFB at 72 hours

Family 2 (α=0.005):

- WI-NRS Success at Week 8
- PASI-75 at Week 8
- CFB in PSD Items related to Itching, Pain, and Scaling (Questions 1, 9, and 11) aggregate score at Week 8
- PSD Item related to Scaling (Question 11) = 0 at Week 8
- PSD Item related to Itching (Question 1) = 0 at Week 8
- PSD Item related to Pain (Question 9) = 0 at Week 8
- PSSI-75 at Week 8

Family 3 (α=0.01 or 0.02 or 0.025):

- S-IGA score of “Clear” at Week 8
- S-IGA Success at Week 4

- CFB in PASI at Week 2
- S-IGA Success at Week 2
- PSD Total Score = 0 at Week 8
- SI-NRS CFB at Day 1

The Fallback Method will be used to pass unused alpha across the Families. The Families will be tested in this order: Family 2, Family 1, Family 3. Testing will be sequential within Family 2 and Family 1, with the order of testing of each endpoint within each family as listed above.

Testing within Family 3 will be implemented with Holm's procedure.

To start, Family 2 will be tested, sequentially within Family 2, each endpoint at the α=0.005 level. Should the testing succeed through all seven endpoints in Family 2, the α=0.005 will remain unused, and will be passed to Family 1.

The testing of the five SI-NRS endpoints in Family 1 will be sequential, either at the α=0.01 level (assumes Family 2 had an unsuccessful test), or the α=0.015 level (assumes all endpoints in Family 2 were successful at the α=0.005 level).

Family 3 contains six endpoints. Family 3 has initially been allocated α=0.01. This creates three possible testing alpha levels for the Family 3 endpoints:

- If all tests in Family 1 (α=0.01) and Family 2 (α=0.005) are successful, then α=0.005+0.01=0.015 is unused, and the α=0.015 will be carried to Family 3, allowing Family 3 endpoints to be tested at the full 0.025 level.
- If the α=0.005 in Family 2 is used, but the subsequent α=0.01 allocated to Family 1 is not used, then Family 3 endpoints will be tested at the α=0.01+0.01=0.02 level.
- If the α=0.01 in Family 1 is used, and the α=0.005 in Family 2 is also used, then Family 3 endpoints will be tested at α=0.01.

6.1.4. Handling of Dropouts or Missing Data

6.1.4.1. Imputation of Missing Data

6.1.4.1.1 Imputation for the primary estimand

Any subject who prematurely withdraws from the study will have their last available data assigned to an analysis window as described in Section 6.1.5. For subjects who discontinue due to lack of efficacy or adverse event, the last dose day will be assigned to an analysis window similarly as described in Section 6.1.5. To comply with the definition of the primary estimand (Section 8.1.1), efficacy data assigned to a pre-specified analysis visit will be removed from the source data used for the multiple imputation process for visits on or after the assigned analysis visit of the last dose day for these subjects. Similarly, SI-NRS and WI-NRS weekly averages will be removed from the source data used for multiple imputation for weeks on or after the assigned analysis week of the last dose day for these subjects . This procedure will ensure that the data collected after intercurrent events are not used in the imputation process.

For the co-primary and secondary efficacy endpoints (i.e., S-IGA, B-IGA, weekly average SI-NRS, weekly average WI-NRS, total PSD score = 0, aggregated PSD score of itching, pain and scaling, PSD score of scaling item of 0, PSD score of the itching item of 0, PSD score of pain item=0, PASI, and PSSI), the analysis will impute missing values using a regression-based multiple imputation model. This is a three-step process.

1. The first step is to understand the pattern of missingness. In order to perform the multiple imputation, a monotone missing pattern has to be achieved. For example, if there exist values for baseline and Week 8 visits, but missing values for the Week 2 or 4 visits, Markov-Chain Monte-Carlo (MCMC) method will be used to impute the small amount of missing data that may be missing at the intermediate visits that is required to make the missing data pattern monotone before applying the multiple imputation algorithm. This method uses a non-informative Jeffreys prior to derive the posterior mode from the expectation-maximization (EM) algorithm as the starting values for the MCMC method. MCMC method will use seed of 66447809. All the endpoints will be considered as continuous for this step. To avoid values that could not be observed in practice for S-IGA, B-IGA, total PSD score, and PSSI, imputed values will be constrained to be integers in the range of the observed endpoint. Since SI-NRS and WI-NRS scores will be calculated as average weekly scores and PASI is a derived score that can have non-integer results (see Section 6.1.7 for more information), this restriction will not apply for those endpoints. Total of 10 imputations are done in the MCMC step.
2. Once the monotone pattern is achieved, the next step is to implement the imputation algorithm. For this, the Predictive Mean Matching method (PMM) will be used. This method is particularly helpful if the normality assumption is violated. For subjects with complete data up to a particular visit, a PMM model will be fit that includes the outcome at that visit as the dependent variable and as independent variables, the outcome at previous visits, pooled site, and treatment group using a seed of 90066927. This process will be repeated 15 times, resulting in a total of 150 complete analysis datasets. The seed may be changed after unblinding in case of any issues with the imputation process, and it will be documented in the CSR if any change is required.
3. For each completed dataset, compute the necessary derived variables. The categorical endpoints will be analyzed using the CMH analysis for each of the complete analysis data sets stratified by pooled site group, baseline S-IGA category, and baseline B-IGA category. The results will be combined into one multiple imputation inference (odds ratio, associated CI, and *P* value) using PROC MIANALYZE as illustrated[4]. The continuous endpoints will be analyzed using descriptive summaries, and an ANCOVA for each of the complete analyses data sets with terms for the baseline outcome, pooled site group, baseline S-IGA category, and baseline B-IGA category as covariates. The descriptive summaries will use the averaged imputed outcomes across all datasets for each subject, and the ANCOVA results will be combined into one multiple imputation inference (LS means, standard errors [SEs], 95% CIs, 97.5% CIs, and *P* values) using PROC MIANALYZE as illustrated[4].

This approach to imputation should be superior to other strategies such as carrying forward the last available observation, which may yield unrealistic imputed values. Also, the use of multiple

imputation avoids the problem of artificially increasing power through data imputation associated with single-imputation methods because it accounts for the uncertainty associated with the imputation. Other missing data will not be imputed, with the exception of incomplete dates as described in Section 6.1.8. For responder analyses, missing data will be treated as a nonresponse with the exception of S-IGA and B-IGA. Only observed data will be summarized using descriptive statistics.

The SAS pseudo code for the multiple imputation process is listed below:

Step 1:

```
proc mi data=example seed=66447809 nimpute=10 round=1 out=example_1;
  mcmc impute=monotone;
  var <baseline score> …… <visit8 score>;
run;
```

Step 2:

```
proc mi data=example_1 seed=90066927 nimpute=15  out=example_2;
  by <MCMC imputation>;
  class <treatment> <pooled site>;
  monotone regpmm(<baseline score> …… <visit8 score>);
  var <treatment> <pooled site> <baseline score> …… <visit8 score>;
run;
```

Step 3:

a) For categorical outcomes, this step involves running a CMH test stratified by pooled site group, baseline S-IGA randomization strata, and baseline B-IGA randomization strata on each completed dataset and combining the results using PROC MIANALYZE.

```
proc freq data=example noprint;
  by <imputationnumber> <visit> ;
  tables <pooled site group> * <BL S-IGA randomization strata> * <BL B-IGA
randomization strata> * <treatment>
         * <outcome> / cmh commonriskdiff(cl=MH) alpha=0.025;
  output out=example_stat cmh;
run;
```

In order to apply PROC MIANALYZE for the categorical outcomes, normalizing transformations have to be applied to odds ratio. *P* values are obtained using Wilson Hilferty transformation as illustrated by Ratitch et.al.[4].

b) For continuous outcomes, this step involves running an ANCOVA with terms for the baseline outcome, pooled site group, baseline S-IGA category, and baseline B-IGA category as covariates for each of the completed dataset and combining the results using PROC MIANALYZE

```
proc mixed data=example;
  by <imputationnumber> <visit> ;
  class <pooled site group> <BL S-IGA randomization strata > <BL B-IGA
randomization strata > <treatment>;
  model <outcome> = <pooled site group> <BL S-IGA randomization strata> <BL
B-IGA randomization strata>
        <treatment> <base> / alpha=0.025;
```

```
  lsmeans <treatment> / alpha=0.025 pdiff cl;
  ods output lsmeans=lsmeans
             diffs=diffs;
run;

*Combines the LS means across the MI datasets;
proc sort data=lsmeans;
  by <treatment> <imputationnumber>;
run:

proc mianalyze data=lsmeans;
  by <treatment>;
  modeleffects estimate;
  stderr stderr;
run;

*Combines the LS mean differences across the MI datasets;
proc sort data=diffs;
  by <treatment> <imputationnumber>;
run:

proc mianalyze data=diffs;
  by <treatment>;
  modeleffects estimate;
  stderr stderr;
run;
```

6.1.4.1.2 Imputation for supplemental estimand

The originally pre-specified primary analysis estimand will remain unchanged, but will be re-labeled as a supplemental estimand and will handle all intercurrent events using a "Treatment Policy Strategy" (Section 8.1.1). Imputation for the supplemental estimand will be similar to the process described above for the primary estimand; however, subjects who discontinue due to lack of efficacy or adverse event will be treated the same as those who did not discontinue for one of those reasons. All observed data for these subjects will be included in the imputation process subject to the analysis visit window strategy described in Section 6.1.5.

6.1.4.2. Tipping Point Analysis

As a sensitivity analysis to the multiple imputation analysis as described in Section 6.1.4.1 for the S-IGA and B-IGA co-primary endpoints, a tipping point analysis will be performed in order to determine the inflection point at which the inference under the MNAR assumption changes substantially. This will be used to check the robustness of the imputation.

The sensitivity analysis will be performed by using a specified sequence of shift parameters, which will adjust the imputed values for observations in both treatment groups. The range of shift parameters to be included in this analysis will be 0 to 2 by 0.5 for the active group and -2 to 0 by 0.5 for the vehicle group. Once the likely point of the shift is determined, the analysis may be rerun using a more focused range around the suspected tipping point. Thus, the value at which the results of the analysis are shifted from significant (i.e., $\alpha \leq 0.025$) to nonsignificant

(i.e., $\alpha > 0.025$) will be determined.

Step 1 and 3 of the analysis will be the same as for the multiple imputation analysis as described in Section 6.1.4.1.1 However, Step 2 of the analysis is where the shift parameters will be applied.

Pseudo-code for Step 2 is as follows:

```
proc mi data=example_1 seed=90066927 nimpute=15  out=example_2;
  by <MCMC imputation>;
  class <treatment> <pooled site group.;
  monotone regpmm(<baseline score> …… <visit8 score>);
  var <treatment> <pooled site group> <baseline score> …… <visit8 score>;
  mnar adjust(<visit2 score> / shift=YY adjustobs=(treatment='Roflumilast
Foam 0.3%'));
  mnar adjust(<visit4 score> / shift=YY adjustobs=(treatment='Roflumilast
Foam 0.3%'));
  mnar adjust(<visit8 score> / shift=YY adjustobs=(treatment='Roflumilast
Foam 0.3%'));
  mnar adjust(<visit2 score> / shift=ZZ adjustobs=(treatment='Vehicle
Foam'));
  mnar adjust(<visit4 score> / shift=ZZ adjustobs=(treatment='Vehicle
Foam'));
  mnar adjust(<visit8 score> / shift=ZZ adjustobs=(treatment='Vehicle
Foam'));
run;

YY and ZZ will encompass all combinations of the range of shift parameters as
pre-specified above.
```

Imputed values for subjects who discontinue due to lack of efficacy or adverse event will be handled as described in Section 6.1.7 to ensure that these subjects are analyzed as non-responders at all visits on or after discontinuation of treatment.

6.1.5. Analysis Visit Windows

6.1.5.1. Non-Diary Assessments

Visits will be analyzed as scheduled (i.e., not subject to visit windowing). Unscheduled, early termination, and/or repeated measurements will only be included if a scheduled measurement is not available and the unscheduled, early termination, or repeated measurement falls within the analysis visit windows as described in Table 2. Visits falling outside of these windows will not be windowed (e.g., an unscheduled visit before first application of study drug). The windows follow the Schedule of Events in Table 1. Unscheduled/repeated measurements will be listed.

Table 2: Analysis Visit Windows (Study Visit) for Non-diary Assessments

Visit Name	Visit Number	Target Start Day	Lower Limit	Upper Limit
Week 2	3	15	2	22
Week 4	4	29	23	42

Visit Name	Visit Number	Target Start Day	Lower Limit	Upper Limit
Week 8	5	57	43	---

6.1.5.2. Diary Assessments (WI-NRS, SI-NRS)

Diary entries will be assigned to a study week for analysis based on the study day relative to the first application of study drug according to Table 3.

Table 3: Analysis Window for Calculation of Average Weekly SI-NRS and WI-NRS

Study Days for Calculation of average weekly WI-NRS	Week (Derived)
(-7, -1) or (-6, 1)*	Baseline
(2, 7)	Week 1
(8, 14)	Week 2
(15, 21)	Week 3
(22, 28)	Week 4
(29, 35)	Week 5
(36, 42)	Week 6
(43, 49)	Week 7
(50, 56)	Week 8

* If the Day 1 score is collected after the first treatment application use data from days (-7, -1), otherwise use data from days (-6, 1). If Day 1 time of assessment is unknown, assume that the assessment was obtained prior to the first treatment application as prescribed by the protocol. If a subject does not have assessment performed on Day 1, then (-7, -1) window will be used in determining the average scores.

6.1.6. Pooling of Sites

Sites will be pooled for statistical analysis as follows. For analysis, sites should have a minimum of 10 randomized subjects. The smallest sites will be grouped sequentially in order of smallest to largest within each country, restricting to those sites that did not meet the minimum enrollment of 10, until each pooled site has a minimum of 10 subjects within each country with at least 1 subject in each treatment group.

6.1.7. Derived Variables

Efficacy variable derivations described in this section will be performed with one exception. The exception is that subjects who discontinued early from study due to an AE or lack of efficacy will be considered as not having response (for MI and non-responder imputation analyses) and will be excluded from observed case analyses for all pre-specified analysis visits (refer to Section

6.1.5) on or after the associated analysis visit of the subject's last dose day). More specifically, for MI and non-responder imputation analyses, these subjects will be counted as non-responders for categorical variables and set to baseline for continuous variables for all visits on or after the associated analysis visit of the subject's last dose day.

- **Study day** is calculated as:
    - observation date – first dose date + 1, if record is collected on or after first dose date
    - observation date – first dose date, if record is collected prior to first dose date.
- **S-IGA Success =** S-IGA of "Clear" or "Almost Clear" plus at least 2-grade improvement from baseline.
- **B-IGA Success =** B-IGA of "Clear" or "Almost Clear" plus at least 2-grade improvement from baseline.
- **Average Weekly SI-NRS Score** = SI-NRS scores will be assigned to a particular study week based on the study week window, as shown in Section 6.1.5.2. Average weekly baseline SI-NRS is derived as the average of all non-missing scores reported during the last 7 days prior to treatment if at least 4 of 7 evaluable daily SI-NRS questionnaires are available. Daily baseline SI-NRS is defined as the last non-missing assessment prior to the first study treatment.

  If 2 or more observations are present for a single day, the most severe itch score will be used as a single daily value before calculating the average weekly score. The average of the reported (non-missing) SI-NRS scores assigned to each study week will be calculated. If at least 4 SI-NRS scores are present in this time period, the average weekly SI-NRS score will be calculated; otherwise, the average weekly score will be missing for that week. The derived average score will be rounded to 1 digit. Missing average weekly scores will be imputed as described in Section 6.1.4.1.
- **SI-NRS Success =** achievement of at least 4-point reduction in average weekly SI-NRS score compared to baseline, calculated only for the subjects with SI-NRS ≥ 4 at baseline.

  Day 1 and 72-hour SI-NRS = the Day 1 and 72-hour endpoints are meant to measure the response to the first and first 3 treatment applications, respectively. The relevant assessments are the ones reported prior to IP application on study days 2 and 4 where study day is calculated as described earlier in this section.

- **Average Weekly WI-NRS Pruritus Score** = WI-NRS pruritus scores will be assigned to a particular study week based on the study week window, as shown in Section 6.1.5.2. If 2 or more observations are present for a single day, the most severe itch score will be used for that day. The average of the reported (nonmissing) WI-NRS pruritus scores assigned to each study week will be calculated. If at least 4 WI-NRS pruritus scores are present in this time period, the average weekly WI-NRS pruritus score will be calculated similar to SI-NRS; otherwise, the average weekly score will be missing for that week. Missing average weekly scores will be imputed as described in Section 6.1.4.1.

- **WI-NRS Success** = achievement of at least 4-point reduction in average weekly WI-NRS score compared to baseline, calculated only for the subjects with WI-NRS ≥ 4 at baseline.
- **PASI** = $(0.1\times[E_h + T_h + S_h]\times A_h) + (0.2\times[E_a + T_a + S_a]\times A_a) + (0.3\times[E_t + T_t + S_t]\times A_t) + (0.4\times[E_l + T_l + S_l]\times A_l)$

  where *E*, *T*, and *S* are erythema, thickness, and scaling, respectively, scored on a scale of 0 to 4, *A* is estimated area of skin involved, graded on a scale of 0 to 6, and *h, a, t,* and *l* are head, arms, trunk, and legs, respectively (range for total score 0 to 72). If any of the component scores are missing, the PASI cannot be calculated.
- **PASI-50, PASI-75, PASI-90, PASI-100 =** achievement of a 50%, 75%, 90%, or 100% reduction in PASI from baseline, respectively.
- **PSD Total Score =** sum of the 16 questions (individual questions scored 0 to 10, where higher scores indicate more severe symptoms; range for total score 0 to 160). If 1 or more items are missing, the score is not calculated.
- **PSD total score of 0** = achievement of a PSD total score of 0.
- **PSD scaling score of 0 =** achievement of a PSD scaling score of 0.
- **PSD pain score of 0 =** achievement of a PSD pain score of 0.
- **PSD itch score of 0 =** achievement of a PSD itch score of 0.
- **PSD Itching/Pain/Scaling Aggregate Score =** sum of these 3 questions (individual questions that are scored 0 to 10, where higher scores indicate more severe symptoms; range for this score is 0 to 30. If 1 or more items are missing, the score is not calculated.
- **PSSI =** (erythema score + induration + desquamation) × extent of scalp affected (range 0 to 72). Erythema, induration, and desquamation are scored on a scale of 0 to 4, where 0 = absent and 4 = severest possible; extent of scale affected is scored on a scale of 0 to 6, where 0 = 0% of involved area and 6 = 90-100% of involved area.
- **PSSI-50, PSSI-75, PSSI-90, PSSI-100 =** achievement of a 50%, 75%, 90%, or 100% reduction in PSSI from baseline, respectively.

- **Scalpdex Score Transformation =** Scalpdex is rated on a 1 to 5 scale which will be transformed to 0 to 100 Scale where 1=0; 2=25; 3=50; 4=75; 5=100. This transformed score is used to calculate scale scores.
  - **Emotions Scale =** average of (Q2, Q4, Q5, Q6, Q7, Q9, Q10, Q11, Q12, Q14, Q16, Q17, Q19, Q20, Q22) after transforming to 0 to 100 scale as mentioned above. Q refers to question number. Q19 will be reverse scored i.e., 1=100; 2=75; 3=50; 4=25; 5=0.
  - **Symptoms Scale =** average of (Q1, Q3, Q8) after transforming to 0 to 100 scale as mentioned above. Q refers to question number.
  - **Functioning Scale =** average of (Q13, Q15, Q18, Q21, Q23) after transforming to 0 to 100 scale as mentioned above. Q refers to question number.
- **Scalpdex Total Score =** calculated as mean of all the 23 scalpdex questions using the transformed scale of 0 to 100. Q19 will be reverse scored i.e., 1=100; 2=75; 3=50; 4=25; 5=0 while calculating the mean. Q refers to question number.

- **DLQI Score** (Dermatology Life Quality Index for ages 17+ years) = sum of the 10 questions (individual questions scored as Very much=3, A lot=2, A little=1, Not at all=0, with Not relevant recoded to 0; Question 7: Yes=3, No=0, with Not relevant recoded to 0; range for score 0 to 30). If 1 item is missing, that item is scored as 0. If 2 or more items are missing, the score should not be calculated.
- **CDLQI Score** (Children's Dermatology Life Quality Index for ages 12-16 years) = sum of the 10 questions (individual questions scored as Very much=3, Quite a lot=2, Only a little=1, Not at all=0; Question 7: if the last week was school time, the question was scored as Very much=3, Quite a lot=2, Only a little=1, Not at all=0, with Prevented school recoded to 3, and if the last week was vacation, the standard responses apply; range for score 0 to 30). If 1 item is missing, that item is scored as 0. If 2 or more items are missing, the score should not be calculated.
- **Compliance** = number of applications divided by the expected number of IP applications for each subject. Compliance will be calculated using drug accountability data over the entire treatment period for each subject, up to treatment completion or discontinuation.
  - **Number of Expected IP Applications** = calculated as last treatment/interruption date – first treatment date + 1. If last treatment date ≥ last interruption date, then the last treatment date will be used; otherwise, last interruption date will be used in deriving the expected number of IP applications.
  - **Number of IP Applications =** number of expected IP applications – missed IP applications as collected in the eCRF.
  - **Number of Days on IP** = last treatment date – first treatment date + 1.
- **Body Mass Index (BMI) (kg/m^2)** = (weight in kg)/[(height in cm/100)^2]. For Week 8, Week 8 height will be used to derive BMI since height is not collected at all visits.
- **BMI Categories**
  - Underweight: BMI < 18.5
  - Normal: 18.5 ≤ BMI ≤ 24.9
  - Overweight: 25.0 ≤ BMI ≤ 29.9
  - Obese: BMI ≥ 30.0
- **Age Categories:**
  - 12 – 17 years
  - 18 – 64 years
  - ≥ 65 years
- **Change from Baseline** = value at current time point – value at baseline.
- **Duration (months) since onset of scalp and/or body psoriasis is derived as:**
  - (Randomization date – onset date)/30.4375
  - If day of onset is not available, assume that onset occurred on the 1st of the month.
- **Treatment-emergent AE (TEAE)** = any AE with an onset date/time after the first application of IP.

- **PHQ-8** = sum of the 8 questions (individual questions scored as Not at all=0, Several days=1, More than half the days=2, and Nearly every day=3; range for score 0 to 24). If more than 1 item is missing, the score should not be calculated. If 1 item is missing, the score is calculated as (sum of answered items*8)/number of answered items.
- **Modified PHQ-A** = sum of the 8 questions (individual questions scored as Not at all=0, Several days=1, More than half the days=2, and Nearly every day=3, (range for score 0 to 24). If 1 item is missing, the score should not be calculated. If 1 item is missing, the score is calculated as (sum of answered items*8)/number of answered items.
- **C-SSRS Suicidal Ideation** = A "yes" answer at any time during treatment to any 1 of the 5 suicidal ideation questions (Categories 1-5: Wish to be Dead, Nonspecific Active Suicidal Thoughts, Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act, Active Suicidal Ideation with Some Intent to Act, without Specific Plan, and Active Suicidal Ideation with Specific Plan and Intent).
- **C-SSRS Suicidal Behavior** = A "yes" answer at any time during treatment to any 1 of the 5 suicidal behavior questions (Categories 6-10: Preparatory Acts or Behavior, Aborted Attempt, Interrupted Attempt, Actual Attempt (nonfatal), Completed Suicide).

In cases where variables for questionnaires were derived in the eCRF (e.g., total scores) and a derivation is indicated in Section 6.1.7, the results as derived by statistics (as opposed to the results derived in the eCRF) will be used in analysis.

6.1.8. Data Adjustments/Handling/Conventions

All collected data will be presented in Clinical Data Interchange Standards Consortium (CDISC) datasets. Data not subject to analysis according to this plan will not appear in any tables or graphs but will be included only in the CDISC datasets.

All *P* values will be displayed in 4 decimals and rounded using standard scientific notation (e.g., 0.XXXX). If a *P* value less than 0.0001 occurs it will be shown in tables as < 0.0001; similarly, if a *P* value greater than 0.9999 occurs it will be shown in tables as > 0.9999.

Adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 25.0 18 March 2022.

A treatment-related AE is any AE with a relationship to the study drug of possibly, probably, likely, or missing.

For partial AE and medication start dates:

- If the year is unknown, then do not impute the date but assign a missing value.
- If the year is known, but the month or month and day is unknown, then:
  - If the year matches the year of first dose date and the end date (if present) is after first dose date, then impute as the month and day of the first dose date.
  - Otherwise, assign 01 January.
- If the year and month are known, but the day is unknown, then:
  - If the month and year match the month and year of the first dose date, then impute as the day of the first dose date.

- Otherwise, assign 01.

For partial AE and medication end dates:

- If the year is unknown, then do not impute the date but assign as missing value.
- If the year is known, but the month or month and day is unknown, then:
  - If the year matches the year of the last date of the study (date of last contact if subject lost to follow-up; date of completion or early termination otherwise), then impute as the month and day of the last date of the study.
  - Otherwise, assign 31 December.
- If the year and month are known, but the day is unknown, then:
  - If the month and year match the month and year of the last date of the study, then impute as the day of the last date of the study.
  - Otherwise, assign the last day of the month.

These conventions will be applied to AE and medication onset dates and times with the following precaution: if the missing date and time reflect the date and time of onset of an AE, the modified date and time will be constructed to match the first documented date/time post drug administration while preserving the order in which the AE was reported in the eCRF.

6.2. Special Handling for COVID-19 Disruptions

In the event subjects are unable to complete protocol-specific assessments onsite, study sites may collect data from subjects remotely via telephone and/or by traditional mail or email. The method used for data collection must be clearly documented in the source. Whenever possible, sites should adhere to the protocol visit window for remote data collection. Screening and Baseline/Day 1 visits/assessments must be performed in the clinic and must NOT be completed remotely. If necessary, these visits can be delayed ensuring they are conducted in the clinic and not remotely.

Investigator assessments and subject questionnaires normally completed during on-site visits should be completed on the appropriate paper source documents and entered in the EDC. The following subject assessments/questionnaires are approved to be collected via telemedicine/remotely:

- C-SSRS
- PHQ-8/PHQ-A
- Subject Local Tolerability
- Adverse Events
- Concomitant medication

The following Investigator assessments cannot be completed via telemedicine/remotely:

- S-IGA
- B-IGA
- PASI

- PSSI
- BSA
- Investigator Local Tolerability
- Subject Weight

Subjects who were affected by COVID-19 disruptions by either missing their Week 8 visit or being discontinued before having a Week 8 visit due to COVID-19 related disruptions will be excluded from the mITT population, as described in Section 5.

7. STUDY SUBJECTS AND DEMOGRAPHICS

7.1. Disposition of Subjects and Withdrawals

Disposition will include tabulations of the number of the screen subjects, the number of screen failures, the number of subjects randomized into each treatment group, the number of subjects who received treatment, subjects completing study, tabulated reasons for discontinuation from the study, including whether the subject did not complete the study due to COVID-19 disruption and reasons, and number of subjects in each analysis population. Disposition will be summarized for all subjects who were entered into database by treatment group and overall.

A listing of subject's disposition will be provided. Information on first screening for subjects who were rescreened, will be presented under the first screening subject identifier.

7.2. Protocol Deviations

A data review will be conducted before database lock by the Medical Monitor and the Sponsor to identify protocol deviations as important or non-important.

The number of subjects with important protocol deviations and/or eligibility will be summarized for each deviation category by treatment group and overall for the subjects in the ITT population. In addition, all the protocol deviations (including important and non-important) associated with COVID-19 will be summarized.

A listing of all protocol deviations will also be provided.

7.3. Demographics and Other Baseline Characteristics

Summary statistics for age, age category (12-17, 18-64 and ≥ 65 years old), gender (including child-bearing potential), race, ethnicity, height, weight, BMI, Fitzpatrick Score, and baseline disease characteristics (duration of scalp/body psoriasis at baseline, % BSA affected by disease, S-IGA, B-IGA, PSSI, SI-NRS, WI-NRS, PSD total score, PASI, Scalpdex, DLQI, and CDLQI) will be presented by treatment group and overall, using descriptive statistics (continuous results) or counts and percentages (categorical results) as appropriate.

A summary of treatment history, including history of response, intolerance, or contraindication to topical corticosteroids and/or topical vitamin D derivatives, Apremilast (Otezla), systematic therapy, phototherapy, and psoriasis involvement on knees and/or elbows, face, and genitalia will be provided.

For ordinal variables such as the S-IGA, B-IGA, SI-NRS (average weekly and daily), and WI-NRS (average weekly and daily), summary statistics including the mean, median, Q1, Q3, minimum, and maximum of the ordinal variable will be presented, as well as frequency counts of each level of the ordinal variable.

These analyses will be conducted for the ITT, and Safety populations, except for treatment history summary, which will only be summarized on the Safety population.

A listing of all demographics, analysis population flag, reason not included in the efficacy analysis will be provided.

Medical history for all subjects will be presented in a by-subject listing.

7.4. Exposure and Compliance

The number of expected IP applications will be summarized by treatment group.

The number of IP applications by each subject based on diary data will be summarized using descriptive statistics appropriate for continuous variables.

The amount of IP each subject used based on can weight will be summarized descriptively by treatment group using continuous methods.

The number of days on IP will be summarized descriptively by treatment group.

A subject will be considered compliant with the dosing regimen if the subject applies at least 80% of the expected applications during the treatment period and does not miss more than 3 consecutive doses.

Investigational product application compliance will be calculated based on number of applications divided by the expected number of IP applications. Compliance will be summarized descriptively by treatment group using the following categories:

- $> 100\%$
- $\geq 80\% - \leq 100\%$
- $< 80\%$

A listing of drug exposure, compliance and accountability will also be provided.

8. EFFICACY ANALYSIS

Efficacy analysis will be based on ITT population unless specified otherwise. The randomized treatment arm and strata will be used in the efficacy analysis.

The efficacy analyses that include the stratification factors (pooled sites, baseline S-IGA and baseline B-IGA) will use the strata data as collected in the interactive web response system (IWRS). A table of randomized strata vs. actual strata will be provided if there are any differences.

The order of testing for the co-primary and secondary endpoints is discussed in Section 6.1.3. Other secondary efficacy endpoints are not included in the testing strategy.

Data analyzed using CMH methodology will be based upon the following code:

```
proc freq data=example noprint;
  by <visit> ;
  tables <pooled site group> * <BL S-IGA randomization strata > * <BL B-IGA
randomization strata > * <treatment>
         * <outcome> / cmh commonriskdiff(cl=MH) alpha=<0.025 or 0.05>;
  output out=example_stat cmh;
run;
```

If a non-estimable odds ratio and associated p-value occur as part of the multiple testing strategy, the p-value to be used in the testing strategy will be based on a test of a difference in proportions. In cases where the odds ratio and difference in proportion are non-estimable, the stratification factors will be removed. Statistical significance will be concluded at the 2.5% significance level (2-sided) or less, as discussed in Section 6.1.3. Odds ratios, 95% CIs, and 97.5% CIs for the odds ratios and differences in percent will be provided, whenever possible.

Data analyzed using ANCOVA will be based upon the following code:

```
proc mixed data=example;
  by <visit>;
  class <treatment> <pooled site group> <BL S-IGA randomization strata > <BL
B-IGA randomization strata >;
  model <outcome> = <treatment> <pooled site group> <BL S-IGA randomization
strata > <BL B-IGA randomization strata > <baseline value> / alpha=<0.025 or
0.05>;
  lsmeans trtn / alpha=<0.025 or 0.05> pdiff cl;
run;
quit;
```

8.1. Primary Efficacy Analyses

The two co-primary efficacy endpoints of S-IGA Success and B-IGA Success at Week 8 will be analyzed using ITT population.

The S-IGA and B-IGA are ordinal scales with 5 severity grades which are reported only in integers. Table 4 illustrates the description of each severity grade.

Table 4: S-IGA and B-IGA

Score	Grade	Description
0	Clear	Plaque thickening = no elevation or thickening over normal skin Scaling = no evidence of scaling Erythema = none (no residual red coloration but post-inflammatory hyperpigmentation may be present)
1	Almost Clear	Plaque thickening = none or possible thickening but difficult to ascertain if there is a slight elevation above normal skin level Scaling = none or residual surface drying and scaling Erythema = light pink coloration
2	Mild	Plaque thickening = slight but definite elevation Scaling = fine scales partially or mostly covering the lesions Erythema = light red coloration
3	Moderate	Plaque thickening = moderate elevation with rounded or sloped edges Scaling = most lesions at least partially covered Erythema = definite red coloration
4	Severe	Plaque thickening = marked or very marked elevation typically with hard or sharp edges Scaling = non-tenacious or thick tenacious scale, covering most or all of lesions Erythema = very bright red coloration; extreme red coloration; deep red coloration

The co-primary estimands can be described by the following attributes:

Population: Subjects with Scalp and Body Psoriasis

Endpoints: S-IGA Success and B-IGA Success

Treatments: Roflumilast foam 0.3% vs. matching vehicle foam

Intercurrent events: Over the 8-week randomized treatment period, subjects may be exposed to possible known or unknown inter-current events that could impact the estimates of the estimand, such as treatment discontinuation due to a specific adverse effect or a lack of effect. A composite strategy will be implemented that handles subjects who discontinue due to lack of efficacy or adverse event as missing not at random, differently than all other subjects. Subjects who discontinue due to lack of efficacy or adverse event will be considered as non-responders (for MI and nonresponder imputation analyses) or missing (for observed case analyses) for all analysis visits (refer to Section 5.4) on or after the associated analysis visit of the subject's last dose day (refer to Section 6.1.5)

The original primary estimand has been left unchanged but re-labeled as a supplemental estimand and will handle all intercurrent events using a "Treatment Policy Strategy", including discontinuation due to lack of efficacy or adverse event. Multiple imputation procedures are described in Section 6.1.4.1.2.

Population-level summary: For this study, the primary population summary will be the ratio of

the odds of achieving S-IGA/B-IGA success and also the ratio of the odds for B-IGA success at Week 8 week using ARQ-154 (roflumilast foam 0.3%), relative to the odds of success at Week 8 using the matching vehicle foam. An additional estimand will be the between group difference in the proportion of subjects achieving S-IGA success and the proportion of subjects achieving B-IGA success after receiving 8 weeks of ARQ-154 (roflumilast foam 0.3%) or a matching vehicle foam.

Missing data will be imputed as described in Section 6.1.4. For each imputation dataset, the odds ratio will be obtained from CMH test stratified by the randomization factors (pooled site group, baseline S-IGA category (3 vs. 4) and baseline B-IGA category (2 vs. ≥ 3)). The proportion differences will also be obtained based on stratification factors of baseline S-IGA, baseline B-IGA and pooled site group. SAS® PROC MIANALYZE will be used to combine the estimates from the imputation datasets (Section 6.1.4). These approaches produce estimates which are the combined odds ratio and the proportion difference resulting from adjusting for the possible confounding effects of the three classification factors – pooled site group, baseline S-IGA category (3 vs. 4) and B-IGA category (2 vs. ≥ 3). Randomized stratification factors will be used in the efficacy analysis. S-IGA and BIGA will be tested separately, but each endpoint must be statistically significant at the 2.5% level for the study to be considered a success.

8.1.1. Scalp Investigator Global Assessment (S-IGA)

Categorical Data Analysis

The first co-primary efficacy endpoint is success in S-IGA of disease severity, defined as an S-IGA of "Clear" or "Almost Clear" plus at least a 2-point improvement from baseline at Week 8.

The primary endpoint will be analyzed using a CMH test stratified by pooled site group, baseline S-IGA stratum, and baseline B-IGA stratum. Statistical significance will be concluded at the 2.5% significance level (2-sided) or less, as discussed in Section 6.1.3. Odds ratios, 95% CIs, and 97.5% CIs for the odds ratios will be provided. Additionally, the 95% and 97.5% CIs for proportion of successes in each treatment group will be presented using an extension of Wilson's method proposed by Lott and Reider[5]. Finally, the proportion difference (using Mantel-Haenzsel stratum weights and the Sato variance estimator) and 95% and 97.5% CIs of the risk difference between roflumilast foam 0.3% and vehicle will be presented.

For the primary analysis, missing S-IGA scores will be imputed using multiple imputation as described in Section 6.1.4.1. The CMH analyses will be performed separately for each of the complete analysis data sets, and the results will be combined into one multiple imputation inference (estimated treatment effect and associated CI and *P* value).

A tipping point analysis will be performed as described in Section 6.1.4.2. For the tipping point analysis, a table will be provided that displays the *P* values in the analysis against the range of shift parameters; included in this table will be the shift parameters, odds ratios, 97.5% CIs, and *P* values for the range of shift parameters to provide information on where the analysis tips from significant to non-significant.

8.1.2. Body Investigator Global Assessment (B-IGA)

The scale, grade, and description for the B-IGA, is identical to the S-IGA, as discussed in Section 8.1. This endpoint will be analyzed in the same manner as S-IGA (Section 8.1.1).

8.1.3. Hypothesis Testing

Primary hypothesis testing will be performed independently on the two co-primary endpoints. The null hypothesis is that S-IGA/B-IGA Success at Week 8 does not differ between roflumilast foam 0.3% and matching vehicle foam. The alternative hypothesis is that the S-IGA/B-IGA Success at Week 8 does differ between roflumilast foam 0.3% and matching vehicle foam.

Null Hypothesis (H_0): $P_R\ Q_V\ /\ P_V\ Q_R = 1.0$,
Alternative Hypothesis (H_A): $P_R\ Q_V\ /\ P_V\ Q_R \neq 1.0$, where

P_R = the proportion of S-IGA/B-IGA Success in roflumilast foam 0.3%
P_V = the proportion of S-IGA/B-IGA Success in matching vehicle foam
$Q_R = 1 - P_R$
$Q_V = 1 - P_V$.

Each co-primary endpoint will be tested with using 2-sided alpha value of 0.025. The null hypotheses of no difference in S-IGA/B-IGA Success at Week 8 between roflumilast foam 0.3% and matching vehicle foam will be rejected if the two individual tests are significant.

8.1.4. Sensitivity and Supplemental Analyses

The following sensitivity and supplemental analyses will be performed for S-IGA and B-IGA, based on the CMH test as described in Sections 8.1.1 and 8.1.2. Analyses of observed data will utilize the same CMH method but will not require the combination of multiple estimates required for the analyses with multiple imputation.:

- The two co-primary efficacy endpoints analyses repeated on subjects in the PP population using observed data.
- The two co-primary efficacy endpoint analyses repeated on subjects in the mITT population using MI data.
- The two co-primary efficacy endpoints analyses repeated on subjects in the ITT population using observed data.
- The two co-primary efficacy endpoints analyses repeated on subjects in the ITT population using a non-responder imputation in which the post baseline missing will be imputed as non-responder, meaning no S-IGA or B-IGA success.
- The tipping point analysis as described in Section 6.1.4.2
- The co-primary efficacy endpoints analyses repeated on subjects in the ITT population using MI imputed data, using the CMH model but excluding site as a stratification factor.

8.1.5. Impact of Site

The following supportive analysis will be performed to investigate the impact of site on the co-primary endpoints.

- To access the impact of pooled site on the primary analysis endpoint, the proportion of subjects achieving S-IGA success within each site will be plotted by treatment groups.
- An additional analysis to examine the impact of pooled study site will examine the changes in p-values that occur after removal of a pooled site from the overall analysis. First, remove one pooled site from the MI datasets, recalculate the combined CMH p-value as described for the primary analysis, and repeat removing a different site for each iteration. The p-values from the iterations will be plotted by the pooled site removed.

8.2. Secondary Efficacy Analyses

All secondary efficacy analyses will be performed using the ITT population, unless otherwise specified, on both observed and multiple imputation data. The non-responder imputation analysis will also be performed for secondary endpoints.

8.2.1. Scalp Itch – Numeric Rating Scale (SI-NRS)

The SI-NRS scale ranges from 0 to 10 with 0 being "no scalp itch" and 10 equaling "worst scalp itch imaginable." This will be determined by asking the subject's assessment of worst scalp itch over the past 24 hours.

Secondary efficacy endpoints relating to SI-NRS are all based on average weekly SI-NRS unless specified otherwise. The SPRU4-ITT population will be used for analyses of SI-NRS success. For other SI-NRS related endpoints, the ITT population will be used.

- For the subjects with baseline SI-NRS ≥ 4, achievement of SI-NRS Success at Week 8
- For the subjects with baseline SI-NRS ≥ 4, achievement of SI-NRS Success at Week 4
- For the subjects with baseline SI-NRS ≥ 4, achievement of SI-NRS Success at Week 2
- SI-NRS CFB at Week 1
- SI-NRS CFB at 72 hours (daily score)
- SI-NRS CFB at Day 1(daily score)

SI-NRS Success Analysis

The estimand attributes (Section 8.1) can be described similarly to those described for the co-primary endpoints with the exception that the population of interest is changed to subjects with a baseline SI-NRS score ≥4. This endpoint will be analyzed in the same manner as S-IGA (Section 8.1.1).

Change in SI-NRS Score Analysis

Change from baseline is the primary population level summary for this estimand, and all other attributes are identical to those specified for the co-primary endpoints. Percent change from

baseline is included as a supplemental summary measure. As specified in Section 2.2.2.2, timepoints of primary interest include Day 1 in daily SI-NRS score, 72 hours in daily SI-NRS score, and Week 1 in average weekly SI-NRS score. The analysis will include descriptive summaries of observed data as well as analyses based on multiple imputation datasets using the ITT population. ANCOVA with terms for pooled site group, baseline S-IGA category, baseline B-IGA category, and baseline SI-NRS score as covariates. The LS means, SEs, 95% CIs, 97.5% CIs, and *P* values will be presented. Refer to Section 6.1.4 for additional details.

Plots of daily SI-NRS (mean and SE of observed value and percent change from baseline) will be provided by treatment groups.

Other efficacy endpoints relating to daily SI-NRS

- Analysis of change and percent from baseline of SI-NRS daily scalp itch assessment will be performed using descriptive summaries. This will be summarized by study day.
- Similar ANCOVA model will be performed for change and percent change from baseline daily SI-NRS every day in the first week, at week 2, 4 and 8

8.2.2. Worst Itch – Numeric Rating Scale (WI-NRS) Pruritus Score

The WI-NRS scale ranges from 0 to 10 with 0 being "no scalp itch" and 10 equaling "worst imaginable itch." This will be determined by asking the subject's assessment of worst scalp itch over the past 24 hours.

The secondary efficacy endpoint relating to WI-NRS pruritus score, which is based on average weekly WI-NRS unless specified otherwise, is:

- For the subjects with baseline WI-NRS score ≥ 4, achievement of at least 4-point improvement from baseline in WI-NRS at Week 8 ("WI-NRS Success at Week 8")

WI-NRS Success Analysis

The estimand attributes (Section 8.1) can be described similarly to those described for the co-primary endpoints with the exception that the population of interest is the PRU4-ITT population. This endpoint will be analyzed in the same manner as S-IGA (Section 8.1.1).

8.2.3. Psoriasis Area and Severity Index (PASI)

Psoriasis Area and Severity Index (PASI) is used for the measurement of severity of psoriasis.

PASI combines the assessment of the severity of lesions and the area affected into a single score in the range 0 (no disease) to 72 (maximal disease).

The secondary efficacy endpoints relating to PASI score are:

- PASI-75 at Week 8
- CFB in PASI at Week 2

PASI-75 Analysis

The estimand attributes (Section 8.1) can be described similarly to those described for the co-primary endpoints. This endpoint will be analyzed in the same manner as S-IGA (Section 8.1.1).

Change in PASI Score Analysis

Change in the PASI total score from baseline to Week 2 will be analyzed in the same way as change from baseline in SI-NRS (Section 8.2.1).

8.2.4. Psoriasis Symptoms Diary (PSD)

The PSD assesses the severity of symptoms or how bothersome are the symptoms of psoriasis. The total score combines the scores of the 16 individual questions into a single score in the range of 0 to 72 (individual questions scored 0 to 10, where higher scores indicate more severe symptoms).

The secondary efficacy endpoints relating to PSD score are:

- CFB in PSD Items related to Itching, Pain, and Scaling (Questions 1, 9, and 11) aggregate score at Week 8
- PSD Item related to Scaling (Question 11) = 0 at Week 8
- PSD Item related to Itching (Question 1) = 0 at Week 8
- PSD Item related to Pain (Question 9) = 0 at Week 8
- PSD Total Score = 0 at Week 8

Change in Total PSD Score Analysis

Change in the PSD score from baseline to Week 2, 4, and 8 will be analyzed in the same way as change from baseline in SI-NRS (Section 8.2.1).

The occurrence of PSD itching, scaling, pain, and total scores of 0 will be analyzed with methods previously described for the primary efficacy endpoints (Section 8.1.1).

8.2.5. Psoriasis Scalp Severity Index (PSSI)

The PSSI is used to measure the severity of psoriasis, combining the assessment of the severity of scalp lesions and the area of scalp affected into a single score ranging from 0 (no disease) to 72 (maximal disease).

The secondary efficacy endpoint relating to PSSI score is:

- PSSI-75 at Week 8

PSSI-75 Analysis

The estimand attributes (Section 8.1) can be described similarly to those described for the co-primary endpoints. This endpoint will be analyzed in the same manner as S-IGA (Section 8.1.1).

8.2.6. Scalp Investigator Global Assessment (S-IGA)

The secondary efficacy endpoints relating to S-IGA score are:

- S-IGA score of "Clear" at Week 8
- S-IGA Success at Week 4
- S-IGA Success at Week 2

S-IGA Categorical Analysis

The estimand attributes (Section 8.1) can be described similarly to those described for the co-primary endpoints. This endpoint will be analyzed in the same manner as S-IGA (Section 8.1.1).

A summary of primary and secondary efficacy analyses is shown in Table 5.

Table 5: Summary of Primary and Secondary Efficacy Analyses

Efficacy Endpoint	Primary Analysis	Sensitivity or Supplemental Analysis
Primary (α = 0.025)		
S-IGA Success (IGA score of '0' or '1' plus at least 2-point improvement from baseline) at week 8	ITT, multiple imputation (CMH)	#1 PP, observed data (CMH) #2 mITT, multiple imputation (CMH) #3 ITT, observed data (CMH) #4 ITT, non-responder imputation (CMH) #5 ITT, Tipping point #6 ITT based on different site pooling strategy #7 ITT without site stratification
B-IGA Success (IGA score of '0' or '1' plus at least 2-point improvement from baseline) at week 8	ITT, multiple imputation (CMH)	#1 PP, observed data (CMH) #2 mITT, multiple imputation (CMH) #3 ITT, observed data (CMH) #4 ITT, non-responder imputation (CMH) #5 ITT, Tipping point #6 ITT based on different site pooling strategy #7 ITT without site stratification
Secondary Family 2 (α=0.005)		

Family 1 (α=0.01 or 0.015): Average weekly SI-NRS Success (achievement of at least a 4-point improvement) at Weeks 2, 4, 8	SPRU4-ITT population, multiple imputation (CMH, odds ratio)	#1 SPRU4-ITT, observed data (CMH) #2 SPRU4-ITT, non-responder imputation (CMH)
Family 1 (α=0.01 or 0.015): • SI-NRS change from baseline at Week 1 • SI-NRS change from baseline at 72 hours •	ITT population, multiple imputation, ANCOVA	#1 ITT, observed data
Family 2 (α=0.005): Average weekly WI-NRS Success (achievement of at least a 4-point improvement) at Week 8	PRU4-ITT, multiple imputation (CMH) , odds ratio)	#1 ITT, observed data (CMH) #2 ITT, non-responder imputation (CMH)
Family 2 (α=0.005): PASI-75 at Week 8	ITT, multiple imputation (CMH)	#1 ITT, observed data (CMH) #2 ITT, non-responder imputation (CMH)
Family 2 (α=0.005): CFB in PSD Items related to Itching, Pain, and Scaling (Questions 1, 9, and 11) aggregate score at Week 8	ITT, multiple imputation, ANCOVA	#1 ITT, observed data
Family 2 (α=0.005): • PSD Item related to Scaling (Question 11) = 0 at Week 8 • PSD Item related to Itching (Question 1) = 0	ITT, multiple imputation (CMH)	#1 ITT, observed data (CMH) #2 ITT, non-responder imputation (CMH)

at Week 8 • PSD Item related to Pain (Question 9) = 0 at Week 8		
Family 2 (α=0.005): PSSI-75 at week 8	ITT, multiple imputation (CMH)	#1 ITT, observed data (CMH) #2 ITT, non-responder imputation (CMH)
Family 3 (α=0.01 or 0.02 or 0.025): S-IGA score of 'Clear' at Week 8	ITT, multiple imputation (CMH)	#1 ITT, observed data (CMH) #2 ITT, non-responder imputation (CMH)
Family 3 (α=0.01 or 0.02 or 0.025): S-IGA success at week 4, 2	ITT, multiple imputation (CMH)	#1 ITT, observed data (CMH) #2 ITT, non-responder imputation (CMH)
Family 3 (α=0.01 or 0.02 or 0.025): Change from baseline in PASI at week 2	ITT, multiple imputation, ANCOVA	#1 ITT, observed data
Family 3 (α=0.01 or 0.02 or 0.025): PSD Total Score = 0 at Week 8	ITT, multiple imputation (CMH)	#1 ITT, observed data (CMH) #2 ITT, non-responder imputation (CMH)
Family 3 (α=0.01 or 0.02 or 0.025): SI-NRS change from baseline at Day 1	ITT population, multiple imputation, ANCOVA	#1 ITT, observed data

8.3. Other Efficacy Analyses

Analyses of other secondary efficacy endpoints will be performed for the ITT population. Table 6 provides a summary of analyses of other efficacy endpoints that were not included in the multiple testing strategy.

Table 6: Summary of Analyses of Other Efficacy Endpoints

Endpoint	**Missing Data Approach**	**Analysis Method**
S-IGA		

Endpoint	Missing Data Approach	Analysis Method
S-IGA score of "Clear" at Weeks 2 and 4	Multiple Imputation, Observed Data, Non-responder Imputation	CMH
Change from baseline in S-IGA at Weeks 2, 4, and 8 (Observed Data)	Observed Data	Summary Statistics
B-IGA		
B-IGA success at Weeks 2 and 4	Multiple Imputation, Observed Data, Non-responder Imputation	CMH
B-IGA score of "Clear" at Weeks 2, 4, and 8	Multiple Imputation, Observed Data, Non-responder Imputation	CMH
B-IGA score of "Clear" or "Almost Clear" at Weeks 2, 4, and 8	Multiple Imputation, Observed Data, Non-responder Imputation	CMH
Change from baseline in B-IGA at Weeks 2, 4, and 8 (Observed Data)	Observed Data	Summary Statistics
SI-NRS		
SI-NRS Success at Weeks 1, 3, 5, 6, and 7	Multiple Imputation, Observed Data, Non-responder Imputation	CMH
Change and percent change from baseline in SI-NRS at Weeks 2, 3, 4, 5, 6, 7, and 8	Multiple Imputation and Observed Data	ANCOVA
WI-NRS		
WI-NRS Success at Weeks 1, 3, 5, 6, and 7	Multiple Imputation, Observed Data, Non-responder Imputation	CMH
Change and percent change from baseline in WI-NRS at Weeks 1, 2, 3, 4, 5, 6, 7, and 8	Observed Data	ANCOVA

Endpoint	Missing Data Approach	Analysis Method
PASI		
PASI-75, at Weeks 2 and 4 and PASI-50, PASI-90, and PASI-100 at Weeks 2, 4, and 8	Multiple Imputation, Observed Data, Non-responder Imputation	CMH
Change from baseline in PASI at Weeks 4 and 8	Multiple Imputation and Observed Data	ANCOVA
Percent change from baseline in PASI at Weeks 2, 4, and 8	Multiple Imputation and Observed Data	ANCOVA
PSSI		
PSSI-75 at Weeks 2 and 4	Multiple Imputation, Observed Data, Non-responder Imputation	CMH
PSSI-50, PSSI-90, and PSSI-100 at Weeks 2, 4, and 8	Multiple Imputation, Observed Data, Non-responder Imputation	CMH
Change and percent change from baseline in PSSI at Weeks 2, 4, and 8	Multiple Imputation and Observed Data	ANCOVA
PSD		
Change and percent change from baseline in PSD total score at Weeks 2, 4, and 8	Multiple Imputation and Observed Data	ANCOVA
Change and percent change from baseline in an aggregate score 3 PSD individual items dealing with Itching, Pain, and Scaling at Weeks 2 and 4.	Multiple Imputation and Observed Data	ANCOVA
Percent change from baseline in an aggregate score 3 PSD individual items dealing with Itching, Pain, and Scaling at Week 8	Multiple Imputation and Observed Data	ANCOVA
PSD Item related to Itching (Question 1) = 0 at Weeks 2 and 4 PSD Item related to Scaling (Question	Multiple Imputation, Observed Data, Non-responder Imputation	CMH

Endpoint	Missing Data Approach	Analysis Method
11) = 0 at Weeks 2 and 4 PSD Item related to Pain (Question 9) = 0 at Weeks 2 and 4		
Achievement of PSD Total Score=0 at Weeks 2 and 4	Multiple Imputation, Observed Data, Non-responder Imputation	CMH
Change and percent change from baseline in PSD individual item scores at Week 2, 4, and 8	Observed Data	ANCOVA
Scalpdex		
Change and percent change from baseline in Scalpdex total score as well as the emotions, symptoms, and functioning scales at Weeks 2, 4, and 8	Observed Data	ANCOVA
DLQI/CDLQI		
Change and percent change from baseline in Dermatology Life Quality Index (DLQI)/ Children's DLQI (CDLQI) total scores at Weeks 2, 4, and 8	Observed Data	ANCOVA
% BSA Affected by Psoriasis		
Change and percent change from baseline in total body % BSA affected by psoriasis at Weeks 2, 4, and 8	Observed Data	ANCOVA
Extent of Scalp Involvement		
Change and percent change from baseline in percent of scalp affected by psoriasis at Weeks 2, 4, and 8	Observed Data	Summary Statistics

The other secondary efficacy endpoints are listed in Section 2.2.2.3. Categorical endpoints will be analyzed with methods similar to those described for the sensitivity and supplemental analyses of S-IGA and B-IGA (Section 8.1. Methods for analyses of change and percent change from baseline are described in Section 8.2.3. Table 6 identifies the methods used to handle missing data and the analysis method. The same methods used for the primary and sensitivity analyses of the primary endpoints will be used. For these other efficacy endpoints, p-values will

be presented outside of the multiple testing paradigm, and therefore should be interpreted with caution.

8.4. Subgroup Analyses of Efficacy Variables

For the co-primary endpoints, S-IGA success at Week 8 and B-IGA success at Week 8, subgroup analysis will be performed on the following using the primary estimand:

- Topical Corticosteroids – Inadequate Response, intolerance, or contraindication (Y/N)
- Topical Vitamin D derivatives –- Inadequate Response, intolerance, or contraindication (Y/N)
- Age categories (12-17, 18-64, ≥ 65)
- B-IGA at baseline (2, ≥3 based on data entered in EDC)
- S-IGA at baseline (3 – Moderate, 4 – Severe based on data entered in EDC)
- Gender (male, female)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- Race (White, Black or African American, Others)
- Fitzpatrick score of I-III vs IV-VI
- Subjects with facial involvement
- Subjects with genitalia involvement
- Subjects with knee involvement
- Subjects with elbow involvement
- Subjects with knee or elbow involvement

The above subgroup analyses are generated with imputed data only. Methods will be similar to those used for the primary analysis of the endpoint. Baseline B-IGA/S-IGA will be removed from the CMH model for analyses by the Baseline B-IGA/S-IGA subgroups, respectively. Forest plots will be produced for the odds ratio point estimates and associated 95% confidence intervals, as well as for the and difference in proportions point estimate and associated 95% confidence intervals. The forest plots will include the overall results (point estimate and 95% confidence interval) along with vertical reference lines at 1 (odds ratio) or 0 (difference in proportions) as well as another vertical reference line at the point estimate for the overall ITT population.

9. SAFETY AND TOLERABILITY ANALYSIS

Safety will be evaluated from reported AEs, physical examinations, local tolerance assessments, changes in clinical laboratory values, changes in vital signs/weight, C-SSRS, PHQ-8, and modified PHQ-A results. No inferential statistical tests will be performed.

All safety analyses will be performed on the Safety population.

9.1. Adverse Events

An overall summary of TEAEs will be provided; this will present number and percent of subjects who reported at least 1: TEAE (including all TEAEs, TEAEs by maximum severity, and TEAEs by greatest relationship), treatment emergent serious adverse events (TESAEs), TEAE leading to discontinuation of the study drug, TEAE leading to study discontinuation, and TEAE resulting in death.

Tables including the number and percentage of subjects will be provided for the following categories:

1. TEAEs by SOC and PT
2. TEAEs by SOC, PT and Maximum Severity
3. TEAEs by SOC, PT, and Strongest Relationship to Study Drug (i.e., related vs. unrelated)
4. TEAEs by PT in descending order
5. TESAE by SOC and PT
6. TEAE leading to discontinuation of the study drug by SOC and PT
7. TEAE leading to study discontinuation by SOC and PT

In the summaries showing severity and relationship to study medication the event with the maximum severity (mild < moderate < severe) or strongest relationship (not related < related) will be reported. If a particular event is missing the severity and/or relationship, then the strongest possible severity or relationship will be assumed for analysis (severity = severe, relationship = likely).

In the AE data listings, all AEs will be displayed. AEs that are treatment emergent will be flagged and AE occurred in the application site will be flagged.

Most frequent TEAE (≥ 1%) of PT will be plotted by treatment groups along with overall incidence of TEAEs and overall incidence of TESAEs.

9.1.1. Adverse Events Leading to Withdrawal

A summary of incidence rates (frequencies and percentages) of TEAEs leading to withdrawal of study drug by treatment group, SOC, and preferred term will be prepared for the Safety population. Additionally, this table will be repeated for TEAEs leading to withdrawal from the study. No inferential statistical tests will be performed.

9.1.2. Deaths and Serious Adverse Events

Any deaths that occur during the study will be flagged in the AE listing.

9.2. Local Tolerability Assessments

The investigator's assessment of the application site reaction will be summarized by study visit using both categorical methods (number and percentage of subjects with each score) as well as continuous methods (e.g., mean, median). Categorical summaries will be provided for dermal response as well as other effects. No inferential statistical tests will be performed.

The subject's assessment of the application site reaction will be summarized similarly.

9.3. Clinical Laboratory Evaluations

Laboratory test results will be summarized descriptively by treatment and study visit as both observed values, change and percent change from baseline values for continuous hematology, chemistry, and urinalysis results. Categorical urinalysis results will be summarized using frequencies by study visit and treatment.

The number of subjects with clinical laboratory values below, within, or above the normal range by study visit and in relation to baseline will be tabulated for each clinical laboratory analyte by treatment group (shift table).

Abnormal results will be flagged in the listings.

9.4. Vital Signs

Descriptive summaries of observed values, changes from baseline, and percent changes from baseline will be calculated for systolic blood pressure, diastolic blood pressure, heart rate, height, weight, BMI, and oral body temperature by treatment group and visit.

The number of subjects who gain or lose ≥ 5% of their baseline body weight during the course of the study, as well as subjects who gain or lose ≥10% of their baseline body weight over the course of the study, will be summarized by treatment group. This summary will be repeated by intentionality of weight loss (intentional or unintentional, as captured on the eCRF).

Shift tables for subjects who shift from their baseline BMI category as defined in Section 6.1.7 (underweight, normal, overweight, obese) to a different BMI category throughout the course of the study will be provided by treatment group and visit. This summary will be repeated by intentionality of weight loss (intentional or unintentional, as captured on the eCRF).

9.5. PHQ-8 and Modified PHQ-A

Data for PHQ-8 and modified PHQ-A will be classified using each subject's total score at a time point into a category based on the following scoring system:

- None – Minimal depression (0 to 4)
- Mild depression (5 to 9)
- Moderate depression (10 to 14)
- Moderately severe depression (15 to 19)
- Severe depression (20 to 24)

Shift tables showing the category of severity at each visit by treatment group will be presented.

9.6. C-SSRS

The C-SSRS is a questionnaire that prospectively assesses Suicidal Ideation and Suicidal Behavior. At the Screening study visit, "Baseline/Screening" version of the C-SSRS will be used. This version assesses Suicidal Ideation and Suicidal Behavior during the subject's lifetime and during the past 6 months. For the Screening visit, "lifetime" experience of the subject with Suicidal Ideation and Suicidal Behavior will be summarized. For baseline visit, the "Since Last Visit" version will be used.

Tables will include results from the Suicidal Ideation and Suicidal Behavior sections of the C-SSRS. Frequencies and percentages of subjects with a response of "Yes" at any point on the Suicidal Ideation and Suicidal Behavior items will be summarized by study visit and treatment group.

9.7. Concomitant Medication

Prior and concomitant medications will be summarized descriptively by treatment, Anatomical Therapeutic Chemical (ATC) Class Level 4, and Preferred Term using counts and percentages.

Prior medications will be presented separately from concomitant medications. Medications that started before the first application of IP will be considered prior medications whether or not they were stopped prior to the first application of study drug. Any medications continuing or starting after the first application of study drug will be considered to be concomitant. If a medication starts before the first application of study drug and continues after the first application of study drug it will be considered both prior and concomitant.

Medications will be coded using World Health Organization (WHO) Drug Global B3, version March 2022.

9.8. Physical Examination

The number and percentage of subjects with normal and abnormal findings in the physical examination will be displayed at each study visit and treatment group.

10. OTHER PLANNED ANALYSIS

10.1. Pharmacokinetic Analysis

All PK collection information from the eCRF, along with PK concentrations, will be presented in a listing. Concentration data will be summarized by study visit and treatment group using summary statistics. Sample values reported below the limit of quantification (BLQ) will be set equal to zero for plasma descriptive statistics. The lower limit of quantification (LLOQ) is 0.100 ng/mL for roflumilast and the N-oxide metabolite. For the analysis of the pre-dose plasma concentrations only, to allow for the calculation of geometric mean, 0.0001 will be added to all concentration values to remove all zero values. Addition of 0.0001 will only be necessary if there are 1 or more values BLQ.

11. CHANGES FROM PLANNED ANALYSIS

11.1. Version 2.0

All the exploratory endpoints have been added in the SAP, but they were not specified in the protocol.

The following analysis populations have been added:

- The SPRU4-ITT population has been added to facilitate interpretation of the SI-NRS analysis in addition to the ITT population; and
- The PRU4-ITT population has been added to facilitate interpretation of the WI-NRS analysis in addition to the ITT population.

These populations are not discussed in the protocol.

In this SAP, PK population is defined as subjects who received at least one confirmed dose of investigational product (IP) and provided at least one PK sample. This population will be used for summary of PK concentration results. The definition is different from that in protocol where PK population is defined as subjects receiving active drug with sufficient plasma concentrations of roflumilast to define a profile, as determined by the pharmacokineticist, which will be used for analysis of PK parameters.

Version 2.0 of the SAP is based on the final version of Amendment 3 for the ARQ-154-309 protocol dated 29-Jul-2022. At the time of signing this SAP version 2.0, the protocol is approved and in the process of being published.

11.2. Version 3.0

Section 1 Overview provides a description of the rationale for SAP Version 3. The purpose of the SAP is to document the change to the primary estimand and related multiple imputation strategies. The revised estimand will be the focus of the Clinical Study Report and NDA. Key results based on the original estimand defined in SAP versions 1 and 2 will be provided as supplemental information for the primary and secondary endpoints covered by the multiple testing strategy. This change to the primary estimand and multiple imputation procedure was based on recommendations from FDA during a pre-NDA meeting for a similar program held on 14Sep2022.

12. REFERENCES

1. US Federal Register. (1998) International Conference on Harmonization; Guidance on Statistical Principles for Clinical Trials. Department of Health and Human Services: Food and Drug Administration [Docket No. 97D-0174]. Federal Register Volume 63, Number 179, pages 49583-49598. September 16, 1998.
2. ASA. (2016) Ethical Guidelines for Statistical Practice. Prepared by the Committee on Professional Ethics, April 2016. http://www.amstat.org/about/ethicalguidelines.cfm
3. RSS. (2014) The Royal Statistical Society: Code of Conduct, 2014. http://www.rss.org.uk/Images/PDF/join-us/RSS-Code-of-Conduct-2014.pdf.

4. Ratitch, B., Lipkovich, I., & O'Kelly, M. (2013). *Combining Analysis Results from Multiply Imputed Categorical Data*. PharmaSUG. https://www.pharmasug.org/proceedings/2013/SP/PharmaSUG-2013-SP03.pdf

5. Lott, A and Reiter, J.P. (2020), "Wilson Confidence Intervals for Binomial Proportions With Multiple Imputation for Missing Data," The American Statistician, 74:2, 109-115

14. TABLES, LISTINGS AND FIGURES

All listings, tables, and graphs will have a header showing the sponsor company name and protocol and a footer showing the version of SAS, the file name and path, and the source of the data (listing number).

14.1. Planned Table Descriptions

The following are planned summary tables for protocol number ARQ-154-309. The table numbers are place holders only and may be changed when the tables are produced.

Table Number	Population	Table Title/Summary
14.1	**Disposition, Baseline Characteristics and Exposure Summary**	
14.1.1	All Subjects	Subject Disposition
14.1.2.1	ITT	Demographics and Baseline Characteristics
14.1.2.1 (cont.)	ITT	Demographics and Baseline Characteristics
14.1.2.1 (cont.)	ITT	Demographics and Baseline Characteristics
14.1.2.1 (cont.)	ITT	Demographics and Baseline Characteristics
14.1.2.1 (cont.)	ITT	Demographics and Baseline Characteristics
14.1.2.1 (cont.)	ITT	Demographics and Baseline Characteristics
14.1.2.2	Safety	Demographics and Baseline Characteristics
14.1.3	Safety	Previous Treatment History of Scalp and Body Psoriasis
14.1.4	ITT	Protocol Deviations
14.1.5	Safety	Prior Medications by Anatomic Therapeutic Chemical (ATC) Class Level 4 and Preferred Term
14.1.6	Safety	Summary of Study Drug Exposure
14.2	**Efficacy Summary**	
14.2.1	ITT, SPRU4-ITT, or PRU4-ITT	Hypothesis Testing for Primary and Secondary Efficacy Endpoints - Multiple Imputation
14.2.1.1	ITT	Scalp Investigator Global Assessment (S-IGA) Grades and Success by Study Visit – Multiple Imputation (Primary Analysis) Categorical Results
14.2.1.1.s	ITT	Scalp Investigator Global Assessment (S-IGA) Grades and Success by Study Visit – Multiple Imputation (Primary Analysis, Supplemental Estimand) Categorical Results
14.2.1.2	mITT	Scalp Investigator Global Assessment (S-IGA) Grades and Success by Study Visit – Multiple Imputation (Sensitivity Analysis) Categorical Results
14.2.1.3	ITT	Scalp Investigator Global Assessment (S-IGA) Grades and Success by Study Visit – Observed Data (Sensitivity Analysis) Categorical Results
14.2.1.4	Per-protocol	Scalp Investigator Global Assessment (S-IGA) Grades and Success by Study Visit – Observed Data (Supplemental Analysis) Categorical Results
14.2.1.5	ITT	Scalp Investigator Global Assessment (S-IGA) Grades and Success by Study Visit – Non-responder Imputation (Sensitivity Analysis) Categorical Results
14.2.1.6	ITT	Scalp Investigator Global Assessment (S-IGA) Grades and Success at Week 8 – Tipping Point Analysis (Sensitivity Analysis) Categorical Results

14.2.1.7	ITT	Scalp Investigator Global Assessment (S-IGA) Grades and Success by Study Visit – Multiple Imputation, without Site Stratification (Sensitivity Analysis) Categorical Results
14.2.1.8.1	ITT	Scalp Investigator Global Assessment (S-IGA) Success at Week 8 – Multiple Imputation, Impact of Site (Supporting Analysis)
14.2.1.8.2	ITT	Scalp Investigator Global Assessment (S-IGA) Success at Week 8 by Site – Observed Data (Supporting Analysis)
14.2.1.9.1	ITT	Scalp Investigator Global Assessment (S-IGA) Grades and Success by Study Visit and Response to Topical Corticosteroids – Multiple Imputation (Subgroup Analysis)
14.2.1.9.2	ITT	Scalp Investigator Global Assessment (S-IGA) Grades and Success by Study Visit and Age Group – Multiple Imputation (Subgroup Analysis)
14.2.1.9.3	ITT	Scalp Investigator Global Assessment (S-IGA) Grades and Success by Study Visit and Baseline S-IGA Disease Severity – Multiple Imputation (Subgroup Analysis)
14.2.1.9.4	ITT	Scalp Investigator Global Assessment (S-IGA) Grades and Success by Study Visit by Study Visit and Baseline B-IGA Disease Severity – Multiple Imputation (Subgroup Analysis)
14.2.1.9.5	ITT	Scalp Investigator Global Assessment (S-IGA) Grades and Success by Study Visit and Gender – Multiple Imputation (Subgroup Analysis)
14.2.1.9.6	ITT	Scalp Investigator Global Assessment (S-IGA) Grades and Success by Study Visit and Ethnicity – Multiple Imputation (Subgroup Analysis)
14.2.1.9.7	ITT	Scalp Investigator Global Assessment (S-IGA) Grades and Success by Study Visit and Race – Multiple Imputation (Subgroup Analysis)
14.2.1.9.8	ITT	Scalp Investigator Global Assessment (S-IGA) Grades and Success by Study Visit and Fitzpatrick Score – Multiple Imputation (Subgroup Analysis)
14.2.1.9.9	ITT	Scalp Investigator Global Assessment (S-IGA) Grades and Success by Study Visit and Facial Involvement – Multiple Imputation (Subgroup Analysis)
14.2.1.9.10	ITT	Scalp Investigator Global Assessment (S-IGA) Grades and Success by Study Visit and Genitalia Involvement – Multiple Imputation (Subgroup Analysis)
14.2.1.9.11	ITT	Scalp Investigator Global Assessment (S-IGA) Grades and Success by Study Visit and Knee Involvement – Multiple Imputation (Subgroup Analysis)
14.2.1.9.12	ITT	Scalp Investigator Global Assessment (S-IGA) Grades and Success by Study Visit and Elbow Involvement – Multiple Imputation (Subgroup Analysis)
14.2.1.9.13	ITT	Scalp Investigator Global Assessment (S-IGA) Grades and Success by Study Visit and Knee or Elbow Involvement – Multiple Imputation (Subgroup Analysis)
14.2.1.9.14	ITT	Scalp Investigator Global Assessment (S-IGA) Grades and Success by Study Visit and Response to Topical Vitamin D Derivatives – Multiple Imputation (Subgroup Analysis)
14.2.1.10	ITT	Summary, Change, and Percent Change from Baseline in Scalp Investigator Global Assessment (S-IGA) by Study Visit – Observed Data
14.2.2.1	ITT	Body Investigator Global Assessment (B-IGA) Grades and Success by Study Visit – Multiple Imputation (Primary Analysis) Categorical Results
14.2.2.1.s	ITT	Body Investigator Global Assessment (B-IGA) Grades and Success by Study Visit – Multiple Imputation (Primary Analysis, Supplemental Estimand) Categorical Results

14.2.2.2	mITT	Body Investigator Global Assessment (B-IGA) Grades and Success by Study Visit – Multiple Imputation (Sensitivity Analysis) Categorical Results
14.2.2.3	ITT	Body Investigator Global Assessment (B-IGA) Grades and Success by Study Visit– Observed Data(Sensitivity Analysis) Categorical Results
14.2.2.4	Per-Protocol	Body Investigator Global Assessment (B-IGA) Grades and Success by Study Visit – Observed Data (Supplemental Analysis) Categorical Results
14.2.2.5	ITT	Body Investigator Global Assessment (B-IGA) Grades and Success by Study Visit – Non-responder Imputation (Sensitivity Analysis) Categorical Results
14.2.2.6	ITT	Body Investigator Global Assessment (B-IGA) Grades and Success at Week 8 – Tipping Point Analysis (Sensitivity Analysis) Categorical Results
14.2.2.7	ITT	Body Investigator Global Assessment (B-IGA) Grades and Success by Study Visit – Multiple Imputation, without Site Stratification (Sensitivity Analysis) Categorical Results
14.2.2.8.1	ITT	Body Investigator Global Assessment (B-IGA) Success at Week 8 – Multiple Imputation, Impact of Site (Supporting Analysis)
14.2.2.8.2	ITT	Body Investigator Global Assessment (B-IGA) Success at Week 8 by Site – Observed Data (Supporting Analysis)
14.2.2.9.1	ITT	Body Investigator Global Assessment (B-IGA) Grades and Success by Study Visit and Response to Topical Corticosteroids – Multiple Imputation (Subgroup Analysis)
14.2.2.9.2	ITT	Body Investigator Global Assessment (B-IGA) Grades and Success by Study Visit by Study Visit and Age Group – Multiple Imputation (Subgroup Analysis)
14.2.2.9.3	ITT	Body Investigator Global Assessment (B-IGA) Grades and Success by Study Visit and Baseline S-IGA Disease Severity – Multiple Imputation (Subgroup Analysis)
14.2.2.9.4	ITT	Body Investigator Global Assessment (B-IGA) Grades and Success by Study Visit and Baseline B-IGA Disease Severity – Multiple Imputation (Subgroup Analysis)
14.2.2.9.5	ITT	Body Investigator Global Assessment (B-IGA) Grades and Success by Study Visit and Gender – Multiple Imputation (Subgroup Analysis)
14.2.2.9.6	ITT	Body Investigator Global Assessment (B-IGA) Grades and Success by Study Visit and Ethnicity – Multiple Imputation (Subgroup Analysis)
14.2.2.9.7	ITT	Body Investigator Global Assessment (B-IGA) Grades and Success by Study Visit and Race – Multiple Imputation (Subgroup Analysis)
14.2.2.9.8	ITT	Body Investigator Global Assessment (B-IGA) Grades and Success by Study Visit and Fitzpatrick Score – Multiple Imputation (Subgroup Analysis)
14.2.2.9.9	ITT	Body Investigator Global Assessment (B-IGA) Grades and Success by Study Visit and Facial Involvement – Multiple Imputation (Subgroup Analysis)
14.2.2.9.10	ITT	Body Investigator Global Assessment (B-IGA) Grades and Success by Study Visit and Genitalia Involvement – Multiple Imputation (Subgroup Analysis)
14.2.2.9.11	ITT	Body Investigator Global Assessment (B-IGA) Grades and Success by Study Visit and Knee Involvement – Multiple Imputation (Subgroup Analysis)
14.2.2.9.12	ITT	Body Investigator Global Assessment (B-IGA) Grades and Success by Study Visit and Elbow Involvement – Multiple Imputation (Subgroup Analysis)

14.2.2.9.13	ITT	Body Investigator Global Assessment (B-IGA) Grades and Success by Study Visit and Knee or Elbow Involvement – Multiple Imputation (Subgroup Analysis)
14.2.2.9.14	ITT	Body Investigator Global Assessment (B-IGA) Grades and Success by Study Visit and Response to Topical Vitamin D Derivatives – Multiple Imputation (Subgroup Analysis)
14.2.2.10	ITT	Summary, Change, and Percent Change from Baseline in Body Investigator Global Assessment (B-IGA) by Study Visit – Observed Data
14.2.3.1	SPRU4-ITT	Average Weekly Scalp Itch–Numeric Rating Scale (SI-NRS) Success by Study Visit – Multiple Imputation (Primary Analysis)
14.2.3.1.s	SPRU4-ITT	Average Weekly Scalp Itch–Numeric Rating Scale (SI-NRS) Success by Study Visit – Multiple Imputation (Primary Analysis, Supplemental Estimand)
14.2.3.2	SPRU4-ITT	Average Weekly Scalp Itch–Numeric Rating Scale (SI-NRS) Success by Study Visit – Observed Data (Sensitivity Analysis)
14.2.3.3	SPRU4-ITT	Average Weekly Scalp Itch–Numeric Rating Scale (SI-NRS) Success by Study Visit - Non-responder Imputation (Sensitivity Analysis)
14.2.3.4	ITT	Average Weekly Scalp Itch-Numeric Rating Scale (SI-NRS) by Study Visit – Multiple Imputation (ANCOVA)
14.2.3.5	ITT	Summary, Change, and Percent Change of Average Weekly Scalp Itch-Numeric Rating Scale (SI-NRS) by Study Visit – Multiple Imputation
14.2.3.6	ITT	Average Weekly Scalp Itch-Numeric Rating Scale (SI-NRS) by Study Visit – Observed Data (ANCOVA)
14.2.3.7	ITT	Summary, Change, and Percent Change of Average Weekly Scalp Itch-Numeric Rating Scale (SI-NRS) by Study Visit – Observed Data
14.2.3.8	ITT	Daily Scalp Itch-Numeric Rating Scale (SI-NRS) for Day 1 (24-hour) and 72-hour Endpoints – Multiple Imputation (ANCOVA)
14.2.3.9	ITT	Summary, Change, and Percent Change of Daily Scalp Itch-Numeric Rating Scale (SI-NRS) by Study Day – Multiple Imputation
14.2.4.1	PRU4-ITT	Average Weekly Worst Itch–Numeric Rating Scale (WI-NRS) Success by Study Visit – Multiple Imputation (Primary Analysis)
14.2.4.1.s	PRU4-ITT	Average Weekly Worst Itch–Numeric Rating Scale (WI-NRS) Success by Study Visit – Multiple Imputation (Primary Analysis, Supplemental Estimand)
14.2.4.2	PRU4-ITT	Average Weekly Worst Itch–Numeric Rating Scale (WI-NRS) Success by Study Visit – Observed Data (Sensitivity Analysis)
14.2.4.3	PRU4-ITT	Average Weekly Worst Itch–Numeric Rating Scale (WI-NRS) Success by Study Visit - Non-responder Imputation (Sensitivity Analysis)
14.2.4.4	ITT	Summary, Change, and Percent Change of Average Weekly Worst Itch-Numeric Rating Scale (WI-NRS) by Study Visit – Observed Data
14.2.4.5	ITT	Summary, Change, and Percent Change of Daily Worst Itch-Numeric Rating Scale (WI-NRS) by Study Day – Observed Data
14.2.4.6	ITT	Average Weekly Worst Itch-Numeric Rating Scale (WI-NRS) by Study Visit– Observed Data (ANCOVA)
14.2.4.7	ITT	Daily Worst Itch-Numeric Rating Scale (WI-NRS) by Study Day – Observed Data (ANCOVA)
14.2.5.1	ITT	Achievement of Psoriasis Area and Severity Index (PASI-75, PASI-90, PASI-50, PASI-100) by Study Visit – Visit–Multiple Imputation
14.2.5.2	ITT	Achievement of Psoriasis Area and Severity Index (PASI-75, PASI-90, PASI-50) by Study Visit – Observed Data

14.2.5.3	ITT	Achievement of Psoriasis Area and Severity Index (PASI-75, PASI-90, PASI-50) by Study Visit – Non-responder Imputation (Sensitivity Analysis) Categorical Results
14.2.5.4	ITT	Summary of Psoriasis Area and Severity Index (PASI) by Study Visit – Multiple Imputation (ANCOVA)
14.2.5.5	ITT	Summary, Change and Percent Change from Baseline in Psoriasis Area and Severity Index (PASI) by Study Visit - Multiple Imputation
14.2.5.6	ITT	Summary of Psoriasis Area and Severity Index (PASI) by Study Visit – Observed Data (ANCOVA)
14.2.5.7	ITT	Summary, Change and Percent Change from Baseline in Psoriasis Area and Severity Index (PASI) by Study Visit – Observed Data
14.2.6.1	ITT	Achievement of Psoriasis Scalp Severity Index (PSSI-75, PSSI-50, PSSI-90, PSSI-100) by Study Visit – Multiple Imputation
14.2.6.1.s	ITT	Achievement of Psoriasis Scalp Severity Index (PSSI-75, PSSI-50, PSSI-90, PSSI-100) by Study Visit – Multiple Imputation (Secondary Analysis, Supplemental Estimand)
14.2.6.2	ITT	Achievement of Psoriasis Scalp Severity Index (PSSI-75, PSSI-50, PSSI-90, PSSI-100) by Study Visit – Observed Data
14.2.6.3	ITT	Achievement of Psoriasis Scalp Severity Index (PSSI-75, PSSI-90, PSSI-50, PSSI-100) by Study Visit – Non-responder Imputation (Sensitivity Analysis) Categorical Results
14.2.6.4	ITT	Summary of Psoriasis Scalp Severity Index (PSSI) by Study Visit – Multiple Imputation (ANCOVA)
14.2.6.5	ITT	Summary, Change and Percent Change from Baseline in Psoriasis Scalp Severity Index (PSSI) by Study Visit - Multiple Imputation
14.2.6.6	ITT	Summary of Psoriasis Scalp Severity Index (PSSI) by Study Visit – Observed Data (ANCOVA)
14.2.6.7	ITT	Summary, Change and Percent Change from Baseline in Psoriasis Scalp Severity Index (PSSI) by Study Visit – Observed Data
14.2.7.1.1	ITT - Patients with Age ≥18 years	Summary of Psoriasis Symptoms Diary (PSD) Aggregate Score of Itching/Pain/Scaling by Study Visit – Multiple Imputation (ANCOVA)
14.2.7.1.1.s	ITT - Patients with Age ≥18 years	Summary of Psoriasis Symptoms Diary (PSD) Aggregate Score of Itching/Pain/Scaling by Study Visit – Multiple Imputation (ANCOVA) (Secondary Analysis, Supplemental Estimand)
14.2.7.1.2	ITT - Patients with Age ≥18 years	Summary, Change, and Percent Change from Baseline in Psoriasis Symptoms Diary (PSD) Aggregate Score of Itching/Pain/Scaling by Study Visit – Multiple Imputation
14.2.7.1.3	ITT - Patients with Age ≥18 years	Summary of Psoriasis Symptoms Diary (PSD) Aggregate Score of Itching/Pain/Scaling by Study Visit – Observed Data (ANCOVA)
14.2.7.1.4	ITT - Patients with Age ≥18 years	Summary, Change, and Percent Change from Baseline in Psoriasis Symptoms Diary (PSD) Aggregate Score of Itching/Pain/Scaling by Study Visit – Observed Data
14.2.7.2.1	ITT - Patients with Age ≥18 years	Achievement of Score of 0 for Psoriasis Symptoms Diary (PSD) Item Related to Scaling by Study Visit – Multiple Imputation
14.2.7.2.1.s	ITT - Patients with Age ≥18 years	Achievement of Score of 0 for Psoriasis Symptoms Diary (PSD) Item Related to Scaling by Study Visit – Multiple Imputation (Secondary Analysis, Supplemental Estimand)
14.2.7.2.2	ITT - Patients with Age ≥18 years	Achievement of Score of 0 for Psoriasis Symptoms Diary (PSD) Item Related to Scaling by Study Visit – Observed Data (Sensitivity Analysis)

14.2.7.2.3	ITT - Patients with Age ≥18 years	Achievement of Score of 0 for Psoriasis Symptoms Diary (PSD) Item Related to Scaling by Study Visit – Non-responder Imputation (Sensitivity Analysis)
14.2.7.3.1	ITT - Patients with Age ≥18 years	Achievement of Score of 0 for Psoriasis Symptoms Diary (PSD) Item Related to Itching by Study Visit – Multiple Imputation
14.2.7.3.1.s	ITT - Patients with Age ≥18 years	Achievement of Score of 0 for Psoriasis Symptoms Diary (PSD) Item Related to Itching by Study Visit – Multiple Imputation (Secondary Analysis, Supplemental Estimand)
14.2.7.3.2	ITT - Patients with Age ≥18 years	Achievement of Score of 0 for Psoriasis Symptoms Diary (PSD) Item Related to Itching by Study Visit – Observed Data (Sensitivity Analysis)
14.2.7.3.3	ITT - Patients with Age ≥18 years	Achievement of Score of 0 for Psoriasis Symptoms Diary (PSD) Item Related to Itching by Study Visit – Non-responder Imputation (Sensitivity Analysis)
14.2.7.4.1	ITT - Patients with Age ≥18 years	Achievement of Score of 0 for Psoriasis Symptoms Diary (PSD) Item Related to Pain by Study Visit – Multiple Imputation
14.2.7.4.1.s	ITT - Patients with Age ≥18 years	Achievement of Score of 0 for Psoriasis Symptoms Diary (PSD) Item Related to Pain by Study Visit – Multiple Imputation (Secondary Analysis, Supplemental Estimand)
14.2.7.4.2	ITT - Patients with Age ≥18 years	Achievement of Score of 0 for Psoriasis Symptoms Diary (PSD) Item Related to Pain by Study Visit – Observed Data (Sensitivity Analysis)
14.2.7.4.3	ITT - Patients with Age ≥18 years	Achievement of Score of 0 for Psoriasis Symptoms Diary (PSD) Item Related to Pain by Study Visit – Non-responder Imputation (Sensitivity Analysis)
14.2.7.5.1	ITT - Patients with Age ≥18 years	Achievement of Score of 0 for Psoriasis Symptoms Diary (PSD) Total Score by Study Visit – Multiple Imputation
14.2.7.5.1.s	ITT - Patients with Age ≥18 years	Achievement of Score of 0 for Psoriasis Symptoms Diary (PSD) Total Score by Study Visit – Multiple Imputation (Secondary Analysis, Supplemental Estimand)
14.2.7.5.2	ITT - Patients with Age ≥18 years	Achievement of Score of 0 for Psoriasis Symptoms Diary (PSD) Total Score by Study Visit – Observed Data (Sensitivity Analysis)
14.2.7.5.3	ITT - Patients with Age ≥18 years	Achievement of Score of 0 for Psoriasis Symptoms Diary (PSD) Total Score by Study Visit – Non-responder Imputation (Sensitivity Analysis)
14.2.7.6.1	ITT - Patients with Age ≥18 years	Summary of Psoriasis Symptoms Diary (PSD) by Study Visit – Multiple Imputation (ANCOVA)
14.2.7.6.2	ITT - Patients with Age ≥18 years	Summary, Change and Percent Change from Baseline in Psoriasis Symptoms Diary (PSD) by Study Visit – Multiple Imputation
14.2.7.6.3	ITT - Patients with Age ≥18 years	Summary of Psoriasis Symptoms Diary (PSD) by Study Visit – Observed Data (ANCOVA)
14.2.7.6.4	ITT - Patients with Age ≥18 years	Summary, Change and Percent Change from Baseline in Psoriasis Symptoms Diary (PSD) by Study Visit – Observed Data
14.2.7.7.1	ITT - Patients with Age ≥18 years	Summary, Change, and Percent Change from Baseline in Psoriasis Symptoms Diary (PSD) Individual Item Scores by Study Visit – Observed Data

14.2.7.7.2	ITT - Patients with Age ≥18 years	Summary of Psoriasis Symptoms Diary (PSD) Individual Item Scores by Study Visit – Observed Data (ANCOVA)
14.2.8.1	ITT	Summary, Change and Percent Change from Baseline in Scalpdex Total Score by Study Visit - Observed Data
14.2.8.2	ITT	Summary of Scalpdex Total Score by Study Visit - Observed Data (ANCOVA)
14.2.8.3	ITT	Scalpdex Individual Questionnaire Responses by Study Visit (Categorical)
14.2.8.4	ITT	Scalpdex Individual Questionnaire Responses by Study Visit (Continuous)
14.2.8.5.1	ITT	Summary, Change and Percent Change from Baseline in Scalpdex Score - Emotions Scale by Study Visit - Observed Data
14.2.8.5.2	ITT	Summary of Scalpdex Score - Emotions Scale by Study Visit - Observed Data (ANCOVA)
14.2.8.6.1	ITT	Summary, Change, and Percent Change from Baseline in Scalpdex Score - Symptoms Scale by Study Visit - Observed Data
14.2.8.6.2	ITT	Summary of Scalpdex Score - Symptoms Scale by Study Visit - Observed Data (ANCOVA)
14.2.8.7.1	ITT	Summary, Change, and Percent Change from Baseline in Scalpdex Score - Functioning Scale by Study Visit -Observed Data
14.2.8.7.2	ITT	Summary of Scalpdex Score - Functioning Scale by Study Visit - Observed Data (ANCOVA)
14.2.9.1	ITT - Patients with Age ≥17 years	Summary, Change and Percent Change from Baseline in Dermatology Life Quality Index (DLQI) by Study Visit - Observed Data
14.2.9.2	ITT - Patients with Age ≥17 years	Summary of Dermatology Life Quality Index (DLQI) by Study Visit - Observed Data (ANCOVA)
14.2.9.3	ITT - Patients with Age 12-16 years	Summary, Change and Percent Change from Baseline in Children's Dermatology Life Quality Index (CDLQI) by Study Visit - Observed Data
14.2.10.1	ITT	Summary, Change, and Percent Change from Baseline in Body Surface Area (%) by Study Visit - Observed Data
14.2.10.2	ITT	Summary of Body Surface Area (%) by Study Visit - Observed Data (ANCOVA)
14.2.11.1	ITT	Summary, Change, and Percent Change from Baseline in Extent of Scalp Involvement by Study Visit - Observed Data
14.3	**Safety Summary**	
14.3.1.1	Safety	Overall Summary of Treatment Emergent Adverse Events
14.3.1.2	Safety	Treatment Emergent Adverse Events by System Organ Class and Preferred Term
14.3.1.3	Safety	Treatment Emergent Adverse Events by System Organ Class, Preferred Term, and Maximum Severity
14.3.1.4	Safety	Treatment Emergent Adverse Events by System Organ Class, Preferred Term, and Relationship to Study Drug
14.3.1.5	Safety	Treatment Emergent Adverse Events by Preferred Term
14.3.1.6	Safety	Treatment Emergent Serious Adverse Events by System Organ Class and Preferred Term
14.3.1.7	Safety	Treatment Emergent Adverse Events Leading to Study Drug Discontinuation by System Organ Class and Preferred Term
14.3.1.8	Safety	Treatment Emergent Adverse Events Leading to Study Discontinuation by System Organ Class and Preferred Term

14.3.2.1	Safety	Quantitative Summary of Investigator Local Tolerability Assessment (Dermal Response) by Study Visit
14.3.2.2	Safety	Categorical Summary of Investigator Local Tolerability Assessment (Dermal Response) by Study Visit
14.3.2.3	Safety	Categorical Summary of Investigator Local Tolerability Assessment (Other Effects) by Study Visit
14.3.2.4	Safety	Quantitative Summary of Subject Local Tolerability Assessment by Study Visit
14.3.2.5	Safety	Categorical Summary of Subject Local Tolerability Assessment by Study Visit Categorical Results
14.3.3.1	Safety	Summary of Clinical Chemistry Laboratory Results (Standard Units) by Study Visit
14.3.3.2	Safety	Shift from Baseline in Clinical Chemistry Laboratory Results (Standard Units) by Study Visit
14.3.3.3	Safety	Summary of Hematology Laboratory Results (Standard Units) by Study Visit
14.3.3.4	Safety	Shift from Baseline in Hematology Laboratory Results (Standard Units) by Study Visit
14.3.3.5	Safety	Summary of Quantitative Urinalysis Laboratory Results (Standard Units) by Study Visit
14.3.3.6	Safety	Shift from Baseline in Quantitative Urinalysis Laboratory Results (Standard Units) by Study Visit
14.3.3.7	Safety	Summary of Qualitative Urinalysis Laboratory Results by Study Visit
14.3.4.1	Safety	Summary of Vital Signs by Study Visit
14.3.4.2	Safety	Change in Weight from Baseline by Study Visit
14.3.4.3	Safety	Change in Weight from Baseline by Study Visit and Weight Loss Intentional/Non-Intentional Categories
14.3.4.4	Safety	Shift from Baseline in BMI by Study Visit
14.3.4.5	Safety	Shift from Baseline in BMI by Study Visit and Weight Loss Intentional/Non-Intentional Categories
14.3.5	Safety	Shift from Baseline in Patient Health Questionnaire Depression Scale (PHQ-8)/Modified PHQ-Adolescents by Study Visit
14.3.6	Safety	Summary of Columbia-Suicide Severity Rating Scale (C-SSRS) Suicidal Ideation and Suicidal Behavior by Study Visit
14.3.7	Safety	Summary of Physical Examination by Study Visit
14.3.8	Safety	Summary of Concomitant Medications by Anatomic Therapeutic Chemical (ATC) Class Level 4 and Preferred Term
14.4	**Pharmacokinetic Data**	
14.4.1	PK	Summary of Pharmacokinetic Concentration Results by Study Visit

14.2. Planned Figure descriptions

The following are planned summary figures for protocol ARQ-154-309. The figure numbers are place holders only and may be changed when figures are produced.

Figure Number	Population	Figure Title/Summary
14.2.1.11	ITT	Forest Plot of S-IGA Success at Week 8 - Multiple Imputation (Subgroup Analysis)

14.2.1.12	ITT	Plot of Proportion (97.5 CI%) of Scalp Investigator Global Assessment (S-IGA) Success at Week 8 by Site - Observed Data (Supporting Analysis) ITT Population
14.2.1.13	ITT	Plot of Proportion (97.5 CI%) of Scalp Investigator Global Assessment (S-IGA) Success at Week 8 - Multiple Imputation, Impact of Site (Supporting Analysis)
14.2.2.11	ITT	Forest Plot of B-IGA Success at Week 8 - Multiple Imputation (Subgroup Analysis) ITT Population
14.2.2.12	ITT	Plot of Proportion (97.5 CI%) of Body Investigator Global Assessment (B-IGA) Success at Week 8 by Site - Observed Data (Supporting Analysis) ITT Population
14.2.2.13	ITT	Plot of Proportion (97.5 CI%) of Body Investigator Global Assessment (B-IGA) Success at Week 8 - Multiple Imputation, Impact of Site (Supporting Analysis)
14.2.3.6	ITT	Plot of Mean (+/-SE) Change, and Percent Change from Baseline of Daily Scalp Itch-Numeric Rating Scale (SI-NRS) by Treatment Group over Time ITT Population
14.3.1.9	Safety	Plot of Most Frequent (≥1.0%) Treatment Emergent Adverse Events by Preferred Term and Treatment Group

14.3. Planned Listing Descriptions

The following are planned data and patient/subject data listings for protocol number ARQ-154-309.

In general, all listings will be sorted by site, and subject number. Screen failures will only be presented in listing 16.1.1.

In all listings, a blank line will be placed between each subject. Within a data listing, if an item appears line after line (e.g., repetition of subject number), then only the first occurrence will be displayed.

In data listings, the information for one subject will be kept on one page if at all possible, rather than splitting a subject's information across pages.

Listing Number	**Population**	**Listing Title/Summary**
16.2.1	Screened Subjects	Subject Disposition
16.2.2.1	ITT	Protocol Deviations
16.1.7	ITT	Randomization
16.2.4.1	ITT	Demographics and Analysis Population
16.2.4.2	ITT	Subject Baseline Characteristics
16.2.4.3	Safety	Medical History
16.2.5.1	Safety	Study Drug Application at the Study Site
16.2.5.2	PK	Pharmacokinetic Blood Collection Dates and Times and Drug Concentrations
16.2.6.1	ITT	Efficacy Endpoints
16.2.6.2	ITT	Daily Scalp Itch and Worst Itch Numerical Rating Scale (SI-NRS and WI-NRS)
16.2.7	Safety	Adverse Events
16.2.8.1	Safety	Abnormal Clinical Laboratory Data: Clinical Chemistry

16.2.8.2	Safety	Abnormal Clinical Laboratory Data: Hematology
16.2.8.3	Safety	Abnormal Clinical Laboratory Data: Urinalysis
16.2.9.1	Safety	Abnormal Vital Signs