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 Seborrheic Dermatitis (STRATUM)
Protocol Number:	ARQ-154-304
Premier Research PCN:	ARCU214357
Document Version:	Final 3.0
Document Date:	22-Sep-2022

APPROVALS

Role	Signatures
Biostatistician Premier Research	Print Name:
	Sign Name:
Peer Reviewer Premier Research	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	19-Apr-2022 – Initial version.
Final V2.0	13-May-2022 – update to multiple endpoint testing strategy for secondary endpoints to reflect amendment 2 of the protocol, clarification that Scalpdex analyses will include only subjects with scalp involvement; minor update to calculation of compliance.
Final V3.0	22-Sep-2022 – Post-database lock and post-unblinding update to SAP as per FDA to amend the primary estimand and multiple imputation procedure based on recommendations from FDA during pre-NDA meeting held 9/14/2022.

TABLE OF CONTENTS

1. OVERVIEW 9
2. STUDY OBJECTIVES AND ENDPOINTS 10
2.1. Study Objectives 10
2.1.1. Primary Objective 10
2.2. Study Endpoints 10
2.2.1. Efficacy Endpoints 10
2.2.2. Safety Endpoints 11
3. OVERALL STUDY DESIGN AND PLAN 12
3.1. Overall Design 12
3.2. Sample Size and Power 12
3.3. Study Population 12
3.4. Treatments Administered 12
3.5. Method of Assigning Subjects to Treatment Groups 13
3.6. Blinding and Unblinding 13
3.7. Schedule of Events 13
4. STATISTICAL ANALYSIS AND REPORTING 16
4.1. Introduction 16
4.2. Interim Analysis and Data Monitoring 16
5. ANALYSIS POPULATIONS 16
6. GENERAL ISSUES FOR STATISTICAL ANALYSIS 17
6.1. Statistical Definitions and Algorithms 17
6.1.1. Baseline 17
6.1.2. Adjustments for Covariates 18
6.1.3. Multiple Comparisons 18
6.1.4. Handling of Dropouts or Missing Data 20
6.1.5. Analysis Visit Windows 23
Table 3 Analysis Windows for Calculation of Average Weekly WI-NRS: 23
6.1.6. Pooling of Sites 24
6.1.7. Derived Variables 24
6.1.8. Data Adjustments/Handling/Conventions 27
6.2. Special Handling for COVID-19 Disruptions 28
7. STUDY PATIENTS/SUBJECTS AND DEMOGRAPHICS 29
7.1. Disposition of Patients/Subjects and Withdrawals 29
7.2. Protocol Deviations 29

7.3. Demographics and Baseline Characteristics 30
7.4. Exposure and Compliance 30
8. EFFICACY ANALYSIS 31
8.1. Primary Efficacy Analysis 31
8.1.1. IGA Success 31
8.1.2. Hypothesis Testing 33
8.1.3. Sensitivity and Supplemental Analyses of the Primary Efficacy Endpoint 33
8.1.4. Impact of Site on the Primary Efficacy Endpoint 34
8.2. Secondary Efficacy Analysis 34
8.2.1. Investigator Global Assessment (IGA) 34
8.2.2. Worst Itch – Numeric Rating Scale (WI-NRS) Pruritus Score 35
8.2.3. Other Efficacy Endpoints 35
8.2.4. Overall Assessment of Erythema 36
8.2.5. Overall Assessment of Scaling 36
8.3. Other Efficacy Analysis 38
8.4. Subgroup Analyses of Efficacy Variables 40
9. SAFETY AND TOLERABILITY ANALYSIS 40
9.1. Adverse Events 40
9.1.1. Adverse Events Leading to Discontinuation 41
9.1.2. Deaths and Serious Adverse Events 41
9.2. Local Tolerability Assessments 41
9.3. Clinical Laboratory Evaluations 42
9.4. Vital Signs 42
9.5. PHQ-8 and Modified PHQ-A 42
9.6. CDI-2 43
9.7. C-SSRS 43
9.8. Physical Examination 43
9.9. Concomitant Medication 44
9.10. Pigmentation Assessment 44
10. CHANGES FROM PLANNED ANALYSIS 44
10.1. Version 2.0 44
10.2. Version 3.0 44
11. OTHER PLANNED ANALYSIS 45
11.1. Pharmacokinetic Analysis 45
12. CHANGES TO ANALYSIS PLANNED IN THE PROTOCOL 45
12.1. Definition of PK Population 45

13. REFERENCES 45
14. TABLES, LISTINGS AND FIGURES 47
14.1. Planned Table Descriptions 47
14.2. Planned Figure descriptions 53
14.3. Planned Listing Descriptions 54
15. TABLES, LISTINGS AND FIGURES SHELLS 56
15.1. Standard Layout for all Tables, Listings and Figures 56

LIST OF ABBREVIATIONS

Abbreviation	Definition
AE	adverse event
ATC	anatomical therapeutic chemical
BLQ	below limit of quantification
BMI	body mass index
CDLQI	children's dermatology life quality index
CI	confidence interval
CMH	Cochran Mantel Haenszel
COVID-19	Novel coronavirus disease
CRF	case report form
CS	clinically significant
CSR	clinical study report
DBL	database lock
DBP	diastolic blood pressure
DLT	dose limiting toxicity
DLQI	dermatology life quality index
DSMB	data safety monitoring board
DSUR	development safety update report
ECG	electrocardiogram
eCRF	electronic case report form
EDC	electronic data capture
EMA	European medicines agency
FDA	food and drug administration
GCP	good clinical practice
HR	heart rate
IB	investigator's brochure
IGA	investigator global assessment
IP	investigational product

Abbreviation	Definition
IRB	institutional review board
IRT	interactive response technology
ITT	intent-to-treat
LLOQ	lower limit of quantification
MedDRA	medical dictionary for regulatory activities
MI	multiple imputation
MTD	maximum tolerated dose
NA	not applicable
NCS	non-clinically significant
PD	protocol deviation
PHQ	patient health questionnaire
PK	Pharmacokinetic
PP	per-protocol
RR	respiratory rate or relative rate
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SE	standard error
SOC	system organ class
SOP	standard operating procedure
TEAE	treatment-emergent adverse event
TESAE	treatment-emergent serious adverse event
WHO-DD	world health organization drug dictionary
WI-NRS	worst itch – numeric rating scale

1. OVERVIEW

This statistical analysis plan (SAP) describes the planned analysis and reporting for Arcutis Biotherapeutics, Inc. protocol number ARQ-154-304 (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 Seborrheic Dermatitis), dated 05-May-2022 Amendment 2. Version 1 of the ARQ-154-SAP was approved on 19-Apr-2022 and was included as background for a Type C meeting submitted to FDA on the same date. FDA's written response to Arcutis questions was received on 3-Jun-2022. In the interim, Protocol Amendment 2 and SAP Version 2.0 (13-May-2022) were created to update the multiple testing strategy for secondary endpoints. The database was locked on 17-May-2022 and unblinded on 19-May-2022. Results have since been presented at a medical conference (Blauvelt A., et al.). On 9-Sep-2022, FDA provided the following advice in preliminary feedback prior to a pre-NDA meeting 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 (ICH 1998). All work planned and reported for this SAP will follow internationally accepted guidelines, published by the American Statistical Association (ASA 2018) and the Royal Statistical Society (RSS 2014), 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 approved before any unblinded inferential or descriptive analysis of data pertaining to Arcutis Biotherapeutics, Inc.'s study ARQ-154-304.

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 ARQ-154 foam 0.3% administered once daily (QD) vs vehicle foam for 8 weeks in subjects with seborrheic dermatitis.

2.2. Study Endpoints

2.2.1. Efficacy Endpoints

2.2.1.1. Primary Efficacy Endpoint

The primary efficacy endpoint of this study is success in Investigator Global Assessment (IGA) of disease severity, defined as IGA score of 'clear' or 'almost clear' plus at least 2-point improvement at Week 8.

2.2.1.2. Secondary Efficacy Endpoint(s)

The secondary efficacy endpoints of this study include the following:

- In subjects with an average weekly Baseline Worst Itch Numeric Rating Scale (WI-NRS) pruritus score of ≥ 4, achievement of at least 4-point from baseline in WI-NRS pruritus score at Week 8 ("WI-NRS Success at Week 8")
- WI-NRS Success at Week 4
- WI-NRS Success at Week 2
- IGA Success at Week 4
- IGA Success at Week 2
- Achievement of an Overall Assessment of Scaling (0-3 scale) score of 0 at Week 8
- Achievement of an Overall Assessment of Erythema (0-3 scale) score of 0 at Week 8

- Achievement of an IGA score of 'clear' at Week 8

2.2.1.3. Other Secondary Efficacy Endpoint(s)

Other secondary efficacy endpoints of this study include the following:

- Change and percent change in Scalpdex total score from baseline at Weeks 2, 4, and 8.
- Achievement of an Overall Assessment of Erythema (0-3 scale) score of 0 at Weeks 2, and 4.
- Achievement of an Overall Assessment of Scaling (0-3 scale) score of 0 at Weeks 2, and 4.
- Change and percent change from baseline in Dermatology Life Quality Index (DLQI), Children's DLQI (CDLQI) at Weeks 2, 4, and 8.
- Change and percent change from baseline in body surface area (BSA) affected at Weeks 2, 4, and 8.
- Change and percent change from baseline in Overall Assessment of Erythema score at Weeks 2, 4, and 8.
- Change and percent change from baseline in Overall Assessment of Scaling score at Weeks 2, 4, and 8.
- Achievement of an IGA score of 'clear' at Week 4.
- Achievement of an IGA score of 'clear' at Week 2.
- Achievement of daily WI-NRS score of 0 or 1.
- Change and percent change from average weekly baseline in WI-NRS

2.2.2. Safety Endpoints

The safety endpoints of this study include the following:

- Adverse events
- Local tolerability assessments
- Clinical laboratory parameters
- Vital signs/weight
- Patient Health Questionnaire (PHQ-8)
- Modified PHQ-Adolescents (PHQ-A)
- Children's Depression Inventory 2 (CDI-2)
- Columbia-Suicide Severity Rating Scale (C-SSRS)
- Physical examinations

- Pigmentation Assessment

3. OVERALL STUDY DESIGN AND PLAN

3.1. Overall Design

This is an 8-week, parallel group, double blind, vehicle-controlled study for the treatment of subjects with seborrheic dermatitis. This study will include both male and female subjects 9 years and older having an IGA score of at least "Moderate" (3). Approximately 450 subjects will be enrolled. Subjects should not have more than 20% body surface area (BSA) of seborrheic dermatitis. After having met all inclusion criteria, and none of the exclusion criteria, subjects will be randomized in a 2:1 ratio to ARQ-154 foam 0.3% QD (Roflumilast foam 0.3%) or Vehicle foam QD on Baseline Day 0 which will be applied to areas of lesions of seborrheic dermatitis. The interval between Screening and Baseline Day 0 could be up to 4 weeks followed by 8 weeks of treatment phase. The total duration of the study is anticipated to be around 12 weeks. Subjects will have to apply the study drug once a day in the evening, except for on Day 0 and Week 2, the study drug is applied at the study site. Subjects have to record the date and time each dose has been applied, any missed doses, and any additional comments.

3.2. Sample Size and Power

A sample size of approximately 450 subjects is planned for the study. Approximately 300 subjects will receive ARQ-154 foam 0.3%; approximately 150 subjects will receive vehicle foam. This will provide at least 90% power to detect a 25% difference between treatment groups on IGA success at α=0.01 assuming the proportion of IGA success at vehicle arm was 40% (odds ratio of 2.79 for Roflumilast vs. vehicle). The number of subjects to be enrolled will also provide sufficient power for the secondary endpoints. Additionally, the larger study size is included in order to provide additional/sufficient number of subjects on ARQ-154 treatment for a safety database.

3.3. Study Population

Study population consists of male and female subjects of age 9 years or older with no more than 20% BSA of seborrheic dermatitis. Subjects should have a minimum IGA of 'Moderate' (3) 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):

- ARQ-154 foam 0.3% QD (Roflumilast foam 0.3%)
- Matching vehicle foam QD

3.5. Method of Assigning Subjects to Treatment Groups

Randomization will take place at the Baseline visit after the Investigator confirms that the subject meets all eligibility criteria listed in the protocol. Subjects will be randomly assigned to apply ARQ-154 foam 0.3% QD or vehicle QD by an interactive response technology system (IRT).

Assignment of treatment arm will be made in a 2:1 ratio stratified by study site and baseline disease severity (IGA = 3 or IGA = 4). Kits containing tubes of investigational product will be assigned to each subject by the IRT system. A subject may receive more than one kit for the treatment period. The kits and tubes will be labeled with a unique number, in a blinded manner.

3.6. Blinding and Unblinding

This is a double-blind study, therefore neither the subjects nor the Investigator, clinical personnel, or sponsor will be aware of which treatment an individual receives. In the event of a medical emergency where breaking the blind is required to provide medical care to the subject, the Investigator may obtain the treatment assignment directly from the IRT system for that subject. Refer to the current version of the IRT plan for details on unblinding. Treatment assignments should, however, remain blinded unless assignment knowledge is necessary to determine subject emergency medical care. The rationale for unblinding must be clearly explained in source documentation and on the CRF, along with the date on which the treatment assignment was obtained. The Investigator is requested to contact the Sponsor promptly in the event of any treatment unblinding.

Blinding of study treatment is critical to the integrity of this clinical trial and therefore, if a subject's treatment assignment is disclosed to the Investigator, the subject will have the study treatment discontinued. All subjects will be followed until study completion unless consent to do so is specifically withdrawn by the subject.

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	Screen	Baseline Day 0	Week 2 Day 14	Week 4 Day 28	Week 8[a] Day 56
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
IGA, Overall Assessment of Erythema, Overall Assessment of Scaling[e]	X	X	X	X	X
BSA[f]	X	X	X	X	X
DLQI/CDLQI[g], Scalpdex	X	X	X	X	X
Daily WI-NRS[h] at home (non-clinic)		Day -7 through Day 56			
Local Tolerability Assessment[i]		X		X	X
Pigmentation Assessment[j]	X	X	X	X	X
C-SSRS, CDI-2/PHQ-8/PHQ-A[k]	X	X		X	X
Medical Photography[l]		X	X	X	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		
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
Prior & 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] To be collected at Screening, Baseline/Day 0, Week 4/Day28, and Week 8/Day 56. For subjects <18 years of age, if the Baseline/Day 0 visit occurs within 3 weeks of Screening, the Screening lab results may be utilized.
[d] Height will be collected at Baseline and Week 8. Weight will be collected at all study visits. Subject to void prior to weight.
[e] IGA will be a 5-point scale ranging from clear (0) to severe (4). **IGA should be completed prior to other physician assessments and by the same Evaluator at each study visit.** Overall Assessment of Erythema (0-3 scale) and Overall Assessment of Scaling (0-3 scale) will be completed.
[f] Total BSA affected by seborrheic dermatitis 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.
[g] The DLQI will be completed by subjects ≥17 years of age. The CDLQI will be completed for subjects 9 to 16 years old, inclusive.
[h] WI-NRS will be completed by subjects daily at home starting 7 days prior to the Baseline/Day 0 visit to Day 56 (Week 8) visit. Daily 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).
[i] Local tolerability assessments to be recorded prior to IP application for the Investigator assessment of skin irritation (Berger and Bowman skin irritation score) and 10-15 minutes post IP application for the subject's '0- 3' burning/stinging assessment. 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 seborrheic dermatitis. At Week 4 and Week 8, subjects will provide a recall assessment of burning/stinging experienced post IP application on the day prior to the clinic visit.
[j] An assessment for hypopigmentation and hyperpigmentation will be performed by the Investigator at all clinic visits.
[k] Adolescents and adults will complete the C-SSRS (12 years of age and older). Adults will complete the PHQ-8. Adolescents (ages 12 to 17 inclusive) will complete the PHQ-A (PHQ-9 modified). Parents/caregivers will complete CDI-2 (parent report) for children 9-11 years of age, inclusive. The CDI-2 will be completed throughout the study for subjects 9-11 years old at the time of assent.
[l] Medical photography will be obtained for target lesions. All efforts will be made to de-identify the subjects. Subjects unwilling to participate in the medical photography will be allowed to opt out of this procedure.
[m] A serum pregnancy test will be administered to all females of child-bearing potential at the Screening visit only.
[n] A urine pregnancy test will be performed at the Baseline, Weeks 2, 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 investigational product.
[o] PK samples (trough) will be collected on Day 28 (Week 4), and Day 56 (Week 8). Ensure investigational product is not applied in the area where PK will be drawn.
[p] Subjects 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 we weighed and recorded at every visit. See IP Handling Plan for details.

[s] Compliance calculation is described in the IP Handling Plan.
[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).

4. STATISTICAL ANALYSIS AND REPORTING

4.1. Introduction

Data processing, tabulation of descriptive statistics, calculation of inferential statistics, and graphical representations primarily use SAS (release 9.4 or higher). If the use of other software is warranted, 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 non-missing values, mean, standard deviation (SD), median, Q1, Q3, minimum, and maximum.

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 each treatment group unless otherwise specified.

The minimum and maximum will be reported with the same degree of precision (ie, the same number of decimal places) as the observed data. Measures of location (mean and median) will be reported to 1 degree of precision more than the observed data and measures of spread (SD) will be reported to 2 degrees of precision more than the observed data.

Percentages will be presented to 1 decimal place, unless otherwise specified.

Unless otherwise indicated in section 6.1.3, all statistical tests will be conducted at the 0.01 significance level using 2-tailed tests, and p-values will be reported. Corresponding 95% and 99% confidence intervals (CIs) will be presented for statistical tests.

ARQ-154 0.3% foam will be described as “Roflumilast foam 0.3%” throughout the tables, figures, and 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:

- **Intent-To-Treat Population (ITT):** The ITT population includes all randomized subjects. This population will be used as primary analysis population of efficacy endpoints along with disposition.
- **Per-protocol Population (PP):** The PP population will be a subset of the subjects in the ITT population, who are at least 80% compliant with study medication application, have IGA assessment at week 8, and show 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.
- **Modified Intent-to-Treat Population (mITT):** The mITT population includes all randomized subjects with the exception of subjects who missed the week 8 IGA assessment specifically due to COVID-19 disruption. This population will be used for sensitivity of the primary endpoint.
- **Pruritus Population in ITT (PRU4-ITT):** This PRU population is a subset of the ITT population and includes subjects with average weekly WI-NRS pruritus score ≥4 at Baseline. This population will be used for the analysis of achievement of at least 4-point improvement in average weekly WI-NRS pruritus score as compared to Baseline.
- **Safety Population (SAF):** The Safety Population includes all subjects who are randomized and receive at least one confirmed dose of investigational product (IP). This population will be used for all safety analyses.
- **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.

6. GENERAL ISSUES FOR STATISTICAL ANALYSIS

6.1. Statistical Definitions and Algorithms

6.1.1. Baseline

Baseline is defined as Study Day 1. In general, the last observation recorded before the first dose of study drug will be used as the baseline for all calculations of change from baseline. For assessments where time was recorded, the last observation recorded before the 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 for adverse events (AEs) and medications starting on the first study treatment dose

administration date which will be considered post baseline.

Average weekly baseline WI-NRS 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 WI-NRS questionnaires are available. Daily baseline WI-NRS 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 pooled site group (see Section 6.1.6) and IGA severity at baseline (moderate, or severe). Subgroup analyses may be generated for the baseline covariates.

6.1.3. Multiple Comparisons

Secondary endpoints will only be tested if the primary endpoint is considered statistically significant at the α=0.01 level. To control for multiple comparisons among the secondary endpoints, the following multiplicity procedure will be used:

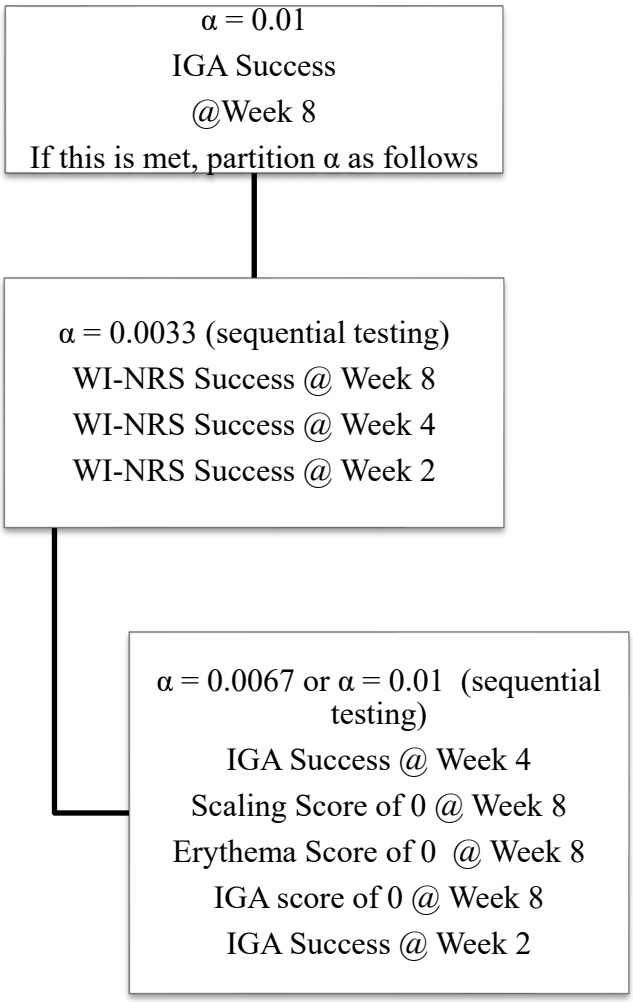
Upon demonstration of statistical significance of the primary endpoint, the alpha level of 0.01 will be unequally split between to families of testing – one family for the WI-NRS success endpoints and another family for the other secondary efficacy endpoints. The testing over the WI-NRS success endpoints and the other secondary efficacy endpoints will be sequential, utilizing the Fallback Method for unused alpha.

For the testing of WI-NRS success endpoints, an alpha level of 0.0033 will be used and will use a hierarchical method to first test WI-NRS success at Week 8, followed by WI-NRS at Week 4, followed by WI-NRS success at Week 2. Should all three of these endpoints be statistically significant at the 0.0033 level, then this unused alpha will be carried to the other secondary endpoint family, which would then be tested at the .0033+.0067=0.01 level. If any of the WI-NRS family do not reach statistical significance, the testing stop at that point in the WI-NRS family, but will continue in the other family, at the 0.0067 level.

In the other secondary endpoint family, an alpha level of 0.0067 (the remainder of the allotted 0.01) or the full 0.01, pending the results of the WI-NRS endpoint testing, will be used to hierarchically test: 1) IGA Success at Week 4, 2) Scaling score of 0 at Week 8, 3) Erythema score of 0 at Week 8, 4) the percentage of subjects who attain IGA of 0 at Week 8, and lastly, 5) IGA Success at Week 2.

This is illustrated in Figure 1.

Figure 1: Primary and Secondary Endpoint Testing


α = 0.01
IGA Success
@Week 8
If this is met, partition α as follows
α = 0.0033 (sequential testing)
WI-NRS Success @ Week 8
WI-NRS Success @ Week 4
WI-NRS Success @ Week 2
α = 0.0067 or α = 0.01 (sequential testing)
IGA Success @ Week 4
Scaling Score of 0 @ Week 8
Erythema Score of 0 @ Week 8
IGA score of 0 @ Week 8
IGA Success @ Week 2

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. To comply with the definition of the primary estimand (Section 8.1.1), data collected on or after the date of last dose for subjects who discontinue due to lack of efficacy or adverse event will be removed from the source data used for multiple imputation process. WI-NRS weekly averages will be removed from the source data used for multiple imputation if the day of last dose falls within the day interval used for the weekly average as defined in Section 6.1.5. This procedure will ensure that the data collected after intercurrent events are not used in the imputation process.

For the primary efficacy endpoint of IGA score, the primary 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 must 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. This MCMC method will use seed of 81054655. The IGA score will be treated as a continuous variable for this step. To avoid values that could not be observed in practice, imputed values will be constrained to be integers in the range of 0 to 4. 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, treatment group, and pooled site using a seed of 13698136. 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. For subjects who discontinue due to lack of efficacy or adverse

event, data from analysis visits that occurred or would have occurred after discontinuation will be replaced with the baseline value. As a result, these subjects will be defined as non-responders for all analyses at these specific analysis visits.

3. For each completed dataset, compute the necessary derived variables. The dichotomous success rate (IGA score of 'clear' or 'almost clear' plus at least 2-point improvement) will be derived. The results obtained will be analyzed using the Cochran-Mantel-Haenszel (CMH) analysis for each of the complete analysis data sets stratified by baseline IGA score and pooled site as described in Section 6.1.6. The results will be combined into one multiple imputation inference (odds ratio, associated confidence interval and p-value) using PROC MIANALYZE as illustrated (Ratitch 2013).

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. The multiple imputation is performed for ITT population only. The analysis for the mITT Population will use the MI datasets already created for the ITT population (i.e., no new MI datasets will be generated).

Similar multiple imputation procedure will be implemented for tested secondary endpoints as primary endpoint. Other missing data will not be imputed, with the exception of incomplete dates as described in Section 6.1.8. For responder analysis, only observed data will be used. 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=81054655 nimpute=10 round=1 out=example_1;
   mcmc impute=monotone;
   var <baseline score> …… <visit8 score>;
run;
```

Step 2:

```
proc mi data=example_1 seed=13698136 nimpute=15 out=example_2;
   by <Step 1 imputation number>;
   class <treatment> <site>;
   monotone regpmm(<baseline score> …… <visit8 score>);
   var <treatment> <pooled site> <baseline score> …… <visit8 score>;
run;
```

Step 3: This step involves running CMH test stratified by pooled site and baseline IGA score on each completed dataset and combining the results using proc mianalyze.

```
Proc freq data=example noprint;
   by <imputationnumber> <visit> ;
   tables <pooled site>*<BL IGA>*<treatment>*<outcome>/ cmh alpha=0.01
COMMONRISKDIFF(TEST=MH cl=mh);
   output out=example_stat cmh;
run;
```

In order to apply PROC MIANALYZE, normalizing transformations must be applied to odds ratio. P-Values are obtained using Wilson Hilferty transformation as illustrated (Ratitch 2013).

6.1.4.1.2 Imputation for supplemental estimand

Based on the feedback from FDA, 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

The sensitivity analysis for the primary endpoint will be performed by using a specified sequence of shift parameters. The range of shift parameters to be included in this analysis are 0 to 2 by 0.5 for active and -2 to 0 by 0.5 for Vehicle. Once the approximate point of the shift is determined, the analysis will be rerun using an expanded range around the approximate tipping point, to determine the tipping point with greater precision (i.e., to 2 decimal places, by 0.01). Thus, the value at which the results of the primary analysis are shifted from significant (i.e., $\alpha \leq 0.01$) to non-significant (i.e., $\alpha > 0.01$) will be determined.

Steps 1 and 3 of the analysis will be the same as for the multiple imputation analysis as described in Section 6.1.4.1.16.1.4.1. 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=13698136 nimpute=15  out=example_2;
    by <Step 1 imputation number>;
    class <treatment> <site>;
    monotone regpmm(<baseline score> …… <visit8 score>);
    var <treatment> <site> <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 Step 2 of Section 6.1.4.1.1 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

Visits will be analyzed as scheduled. Unscheduled, early termination visits, and/or repeated measurements will only be included if a scheduled measurement is not available, and the early termination or unscheduled/repeated measurement falls within the analysis visit windows as described in Table 2. Unscheduled/repeated measurements will be listed.

Table 2: Analysis Visit Windows

Visit Name	Visit Number	Target Start Day	Lower Limit	Upper Limit
Week 2	3	14	2	22
Week 4	4	28	23	42
Week 8	5	56	43	

Table 3 Analysis Windows for Calculation of Average Weekly WI-NRS:

Study Days for Calculation of average weekly WI-NRS	Week (Derived)
(-7, 1)*	Baseline
(2, 7)	Week 1
(8, 14)	Week 2
(15, 21)	Week 3

Study Days for Calculation of average weekly WI-NRS	Week (Derived)
(22, 28)	Week 4
(29, 35)	Week 5
(36, 42)	Week 6
(43, 49)	Week 7
(50, 56)	Week 8

* Day 1 WI-NRS score will be used to calculate the "Baseline" only when it is collected before the application of the first study drug.

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 (ITT population), with at least one subject in each treatment group. The smallest sites within each country will be grouped sequentially in order of smallest to largest, restricting to those sites that did not meet the minimum enrollment of 10, until each pooled site has a minimum of 10 subjects with at least 1 subject in each treatment group in each strata combination.

As sensitivity analysis of different pooling strategy, sites that have randomized less than 50% of the number of randomized subjects at the site with the largest number of randomized subjects, those sites will be pooled within each country. The sensitivity analysis of pooling strategy will be applied for the primary endpoint only.

6.1.7. Derived Variables

- **Average weekly baseline WI-NRS** 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 WI-NRS questionnaires are available. Daily baseline WI-NRS is defined as the last non-missing assessment prior to the first study treatment.
- **IGA success** = IGA of "Clear" or "Almost Clear" plus at least 2-point improvement from baseline.
- **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.

- **Compliance** = **number of applications** divided by the **expected number of IP applications** for each subject. Number of expected study treatment applications will be calculated as the sum of number of investigational product applications and number of doses missed. 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** = last treatment or missed dose date – first treatment date + 1.
  - **Number of IP applications =** number of expected IP applications – number of days with missed IP applications as collected in the CRF.
  - **Number of Days on IP =** last treatment date – first treatment date + 1.
- **Total Weight of IP Applied (g)** = dispensed can weight – returned can weight.
- **Average Daily Weight of IP Applied (g/day)** = total weight of IP applied/number of days on IP.
- **BMI (kg/m^2)** = (weight in kg)/[(height in cm/100)^2]. For Week 2, and 4, baseline height will be used to derive BMI. 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:**
  - 9 – 17 years
  - 18 – 64 years
  - ≥ 65 years
- **Duration (months) since diagnosis of seborrheic dermatitis** is derived as:
  - (Randomization date – diagnosis date)/30.4375
  - If day of diagnosis is not available, assume that diagnosis occurred on the 1st of the month.
- **WI-NRS 4-point Reduction** = achievement of at least 4-point reduction in average weekly WI-NRS pruritus score at Weeks 2, 4, and 8 compared to baseline, calculated only for subjects with a pruritus score of ≥ 4 at baseline (PRU4-ITT population).

- **Average Weekly WI-NRS Pruritus Score** will be calculated in a 7-day period, see table 3 for the analysis window. A minimum of 4 days of observations are needed to calculate average weekly WI-NRS.
- **DLQI Score** = sum of the 10 questions (individual questions scored as Very much=3, A lot=2, A little=1, Not at all=0, Not relevant=0, Question 7: Yes=3, No=0, with Not relevant recoded to 0; range for score 0 to 30). If 1 item is missing, it is scored as 0 for that item. If 2 or more items are missing, the score should not be calculated.
- **CDLQI Score** = 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.
- **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 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.
- **CDI-2** = total score is calculated as a sum of the 17 questions (individual questions scored as much or most of the time=3, often=2, some of the time=1, Not at all=0; range for score 0 to 51).
  - CDI-2 emotional problem scale is a sum of 9 questions (Q1, Q3-6, Q8, Q10 -12)
  - CDI-2 function problem scale is a sum of 8 questions (Q2, Q7, Q9, Q13-17)
- **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.
- **Change from Baseline** = value at current time point – value at baseline.
- **TEAE** = any AE with onset date/time after the first application of IP.
- **C-SSRS Suicidal Ideation** = A "yes" answer at any time during treatment to any one of the five suicidal ideation questions (Categories 1-5: Wish to be Dead, Non-specific 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 Active Suicidal Ideation with Specific Plan and Intent).
- **C-SSRS Suicidal Behavior** = A "yes" answer at any time during treatment to any one of the five suicidal behavior questions (Categories 6-10: Preparatory Acts or Behavior, Aborted Attempt, Interrupted Attempt, Actual Attempt (non-fatal), Completed Suicide).
- **Trunk Involvement =** "Yes" if a subject has coverage area (trunk) greater than 0% as captured in the BSA page.

6.1.8. Data Adjustments/Handling/Conventions

All collected data will be presented in listings. Data not subject to analysis according to this plan will not appear in any tables or graphs but will be included only in the data listings.

All p-values will be displayed in four decimals and rounded using standard scientific notation (eg, 0.XXXX). If a p-value is less than 0.0001 occurs, it will be shown in tables as < 0.0001. Similarly, if a p-value is greater than 0.9999, occurs it will be shown in tables as > 0.9999.

Adverse events will be coded using the MedDRA version 24.1 or above.

Medications will be coded using the World Health Organization Drug Dictionary Enhanced (WHO-DDE) Global B3 version September 2021.

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.

6.2. Special Handling for COVID-19 Disruptions

In some cases, study visits will have to be delayed/not performed as a result of COVID-19 disruptions (e.g., sites were closed, or subjects were under stay-at-home orders). Where possible, study sites may collect post-baseline data from subjects remotely by telemedicine, traditional mail, and/or email; this will be clearly documented in the source. If possible, sites should adhere to the protocol visit window for remote data collection.

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

- WI-NRS
- DLQI/CDLQI
- Scalpdex
- C-SSRS
- PHQ-8/PHQ-A/CDI-2
- Subject Local Tolerability

The following assessments cannot be completed via telemedicine/remotely:

- IGA
- BSA
- Investigator Local Tolerability
- Overall Assessment of Erythema
- Overall Assessment of Scaling
- Pigmentation Assessment
- Subject Weight

Study visits and procedures must be followed per protocol whenever possible. Any specific changes in study conduct that deviate from the protocol should be communicated to the institutional review board and Sponsor. All protocol deviations which occurred as a result of COVID-19 disruptions (e.g., visits out of window, missed assessments, etc.) will be differentiated from other protocol deviations.

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 PATIENTS/SUBJECTS AND DEMOGRAPHICS

7.1. Disposition of Patients/Subjects and Withdrawals

Disposition will include tabulations of the number of the screened subjects, the number of the subjects who failed screen, the number of subjects randomized into each treatment group, the number of subjects who received treatment, subjects completing the study, the reasons for discontinuation from the study overall and due to COVID-19 disruption, and the 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. The reasons for screen failures will be listed.

7.2. Protocol Deviations

Important deviations will be identified based on a blinded review at the end of the study performed by the study team. The number of subjects with important protocol deviations will be summarized in categories by treatment group and overall, for subjects in the ITT population.

A listing of all protocol deviations will also be provided.

In addition, all the protocol deviations associated with COVID-19 will be summarized by deviation category and listed as described above.

7.3. Demographics and Baseline Characteristics

Summary statistics for age, age groups (9-17 years, 18-64 years and ≥ 65 years old), gender (including child-bearing potential), race, ethnicity, height, weight, BMI, baseline disease characteristics (IGA, Overall Assessment of Erythema, Overall Assessment of Scaling, Scalpdex, average weekly WI-NRS, daily WI-NRS, DLQI, CDLQI), Fitzpatrick skin type, percent BSA affected by disease, and body parts involved will be presented by treatment group and overall.

A summary of seborrheic treatment history, including history of response, intolerance, or contraindication to topical corticosteroids and/or topical antifungals, and off-label use of topical calcineurin inhibitors will be provided.

For the continuous variables, the number of non-missing values and the mean, standard deviation, median, Q1, Q3, minimum, and maximum will be tabulated.

For the categorical variables, the counts and proportions of each value will be tabulated.

For ordinal variables such as the IGA, Overall Assessment of Erythema, Overall Assessment of Scaling, and daily WI-NRS, summary statistics including the mean, median, Q1, Q3, and range of the ordinal variable will be presented. Frequency counts of each level of the categorized variables will also be presented.

These analyses will be conducted for the ITT and Safety populations separately (if these two populations include different sets of subjects).

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

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 by treatment group using descriptive statistics appropriate for continuous variables.

The amount of IP used by each subject 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 IP application 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%
- ≥ 80% - ≤100%
- < 80%

A listing of drug exposure 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, IGA strata ("3" (moderate) and "4" (severe)) will use the data as collected in the interactive web response system (IWRS). A table of randomized strata vs. actual strata will be provided if there is any mis-randomization.

8.1. Primary Efficacy Analysis

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

In general, for efficacy analysis, the CMH test is performed on both imputed and observed data. Numerical summaries, categorical frequencies of each severity group are summarized using observed data only. Analysis of imputation results are summarized for ITT population only unless specified otherwise. In rare cases, the odds ratio is not estimable due to zero event in one of the categories, the odds ratio and associated p-value will not be reported. 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.

8.1.1. IGA Success

The IGA is an ordinal scale with five severity grades which is reported only in integers. Table 3 illustrates the description of each severity grade.

Table 4: IGA

Score	Description
0	**Clear:** No erythema, no scaling (hypo-hyperpigmentation can be present)
1	**Almost Clear:** Slight erythema and/or trace (barely perceptible) amounts of scaling
2	**Mild:** Pink to red color and/or slight scaling
3	**Moderate:** Distinct erythema (redness) and/or clearly visible scaling

4	**Severe:** Severe erythema (intense, fiery red) and/or severe scaling (coarse, thick scales with flaking onto clothes or skin)

The primary efficacy endpoint is success in IGA of disease severity, defined as IGA score of "Clear" or "Almost Clear" plus at least 2-point improvement from baseline at Week 8. The definition is the same as IGA score of "Clear" or "Almost Clear" given all subjects are enrolled in this study with moderate or severe IGA.

The primary statistical comparison will be as follows:

- Roflumilast foam 0.3% versus Vehicle foam

The primary estimand is based on feedback from FDA received after database lock and unblinding and is described by the following attributes:

Population: Seborrheic Dermatitis

Intercurrent events: In the course of the 8-week randomized treatment period, subjects may be exposed to possible known or unknown intercurrent events that could possibly impact the estimates of the estimand, such as treatment discontinuation due to a specific adverse effect or perhaps 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. The "Treatment Policy Strategy" has been adopted for handling all other known or unknown intercurrent events in this study other than discontinuation due to lack of efficacy or adverse event. Subjects who discontinue due to lack of efficacy or adverse event will be treated as non-responders at all analysis visits that occurred or would have occurred on or after the date of last dose of treatment application.

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: ratio of the odds of achieving IGA success after 8 weeks of using ARQ-151 (roflumilast foam 0.3%), relative to the odds of success after 8 weeks using a matching vehicle foam.

The supportive population-level summary: the proportion difference between ARQ-151 foam 0.3% and vehicle groups will be provided for the patients who achieve IGA success at week 8.

Rationale for the estimands: Odds ratio of achieving IGA success for ARQ-151 (roflumilast foam 0.3%) relative to vehicle after 8 weeks and the proportion difference of achieving IGA success will be evaluated regardless of the occurrence of any such intercurrent event. The odds ratio will be obtained from CMH test stratified by the randomization factors (by baseline IGA

score and by pooled study site). The strata-adjusted proportion difference will be obtained from CMH method based on stratification factors of baseline IGA and study sites.

Statistical significance will be concluded as described in section 6.1.3. 99% CIs for the odds ratios and for the stratified proportion difference between treatment arm will be provided. Additionally, the 99% CI for proportion of subjects who achieved IGA successes in each treatment group will be presented. 95% CIs for the odds ratios and the stratified proportion difference in each treatment arm are provided as well. Both 99% CI & 95% CI will be presented for all the efficacy endpoints. For the primary analysis, missing IGA scores will be imputed using multiple imputation as described in Section 6.1.4. These imputations will result in a total of 150 complete analysis datasets.

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 (stratified proportion differences and associated CI, odds ratio and associated CI and p-value). The proportion of subjects who achieved IGA success and associated CI, using an extension of Wilson's method proposed by Lott and Reiter, will be reported using multiple imputation as well.

IGA numeric data and change from baseline will be summarized using descriptive summaries (mean, median, inter quartile range) by treatment group and study visit using observed data.

8.1.2. Hypothesis Testing

Primary hypothesis testing on the odds ratio: The null hypothesis is that the IGA Success does not differ between roflumilast foam 0.3% and matching vehicle foam. The alternative hypothesis is that the IGA Success 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 IGA Success in roflumilast foam 0.3%
P_V = the proportion of IGA Success in matching vehicle foam
$Q_R = 1 - P_R$
$Q_V = 1 - P_V$.

8.1.3. Sensitivity and Supplemental Analyses of the Primary Efficacy Endpoint

The following sensitivity and supplemental analyses will be performed for primary efficacy endpoint, based on the CMH test as described in Section 8.1.1:

- The primary efficacy endpoint analysis repeated on subjects in the PP population using observed data.
- The primary efficacy endpoint analysis repeated on subjects in the mITT population on

MI imputed data.

- The primary efficacy endpoint analysis repeated on subjects in the ITT population using observed data.
- The primary efficacy endpoint analysis repeated on subjects in the ITT population using non-responder imputation in which the post baseline missing will be considered as non-responder (no IGA success).
- The primary efficacy endpoint analysis repeated on subjects in the ITT population on MI imputed data using CMH model without site stratification.
- A tipping point analysis will be performed as described in Section 6.1.4.2

8.1.4. Impact of Site on the Primary Efficacy Endpoint

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

- To assess the impact of site on the primary analysis endpoint, the proportion of subjects achieving IGA success and 99% CI within each site will be plotted by treatment groups. No p-values will be provided.
- An additional analysis to examine the impact of study site will examine the changes in p-values that occur after removal of subjects from a site. First, remove subjects from one 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 site removed.

8.2. Secondary Efficacy Analysis

All secondary efficacy analyses will be performed using the ITT unless specified otherwise. The sensitivity analysis will be performed in a similar fashion as those in the primary efficacy endpoint as described in Section 8.1.3. See the list of sensitivity analysis for secondary endpoints in table 6).

The statistical comparison will be as follows:

- Roflumilast foam 0.3% versus Vehicle foam

8.2.1. Investigator Global Assessment (IGA)

Secondary efficacy endpoints in the multiple testing scheme related to the IGA include the following:

- IGA Success at Week 4
- IGA Success at Week 2

- Achievement of an IGA score of 'clear' at Week 8.

The above analysis will be performed on subjects in the ITT, for imputed and observed data, unless otherwise specified using the similar CMH test and the stratified proportion difference as described in Sections 8.1.1. and 8.1.3. The sensitivity analysis for IGA endpoints (IGA success at Week 2, Week 4 and IGA score of 0 at Week 8) will be provided in a similar fashion as those in the primary endpoint as summarized in Table 6.

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

Secondary efficacy endpoints related to WI-NRS pruritus score, which are all based on average weekly WI-NRS unless specified otherwise, include the following:

- In subjects with baseline WI-NRS pruritus score of ≥ 4, achievement of at least 4-point improvement from baseline in WI-NRS pruritus score at Week 8 ("WI-NRS Success at Week 8").
- WI-NRS Success at Week 4.
- WI-NRS Success at Week 2.

The following analyses are performed for WI-NRS:

- In subjects with a baseline WI-NRS score ≥ 4, WI-NRS success endpoint (at Week 2, at Week 4 and at Week 8) will be analyzed in a similar fashion as those for the primary endpoint.
- The primary efficacy endpoint analysis repeated on subjects in the PRU4-ITT population using non-responder imputation in which the post baseline missing will be considered as non-responder (no WI-NRS success).

8.2.3. Other Efficacy Endpoints

Other efficacy endpoints related to average weekly WI-NRS pruritus score:

- Change and percent change from baseline in average weekly WI-NRS pruritus score at Weeks 2, 4, and 8 will be performed using descriptive summaries (mean, median, inter quartile range) by treatment group and study visit using observed values.
- Change and percent change from baseline in average weekly WI-NRS at Weeks 2, 4, and 8 will be analyzed using an analysis of covariance (ANCOVA) model with terms for treatment, pooled study site, and baseline IGA as covariate. The observed value will be used. Statistical comparisons between the treatment groups will be obtained using contrasts. The LS Means, standard errors, 95% and 99% confidence intervals, and p-values will be presented at each visit. These analyses will be performed on the ITT populations. The p-values are nominal and there is no formal inferential testing on other efficacy endpoints.

Other efficacy endpoints related to daily WI-NRS pruritus score:

- Change and percent change from baseline in daily WI-NRS score will be provided by treatment groups along with plots of daily WI-NRS (mean and SD of observed value and percent change from baseline).
- Similar ANCOVA model will be performed for change and percent change from baseline daily WI-NRS every day in the first week, at Weeks 2, 4 and 8 using observed data.

8.2.4. Overall Assessment of Erythema

The assessment is performed at Screening, Baseline, and Weeks 2, 4, and 8. The score is reported in ordinal scale with severity grades ranging from 0-3. Table 4 illustrates description of erythema grades.

Table 3: Erythema Grades

Score	Description
0	**None:** No evidence of erythema
1	**Mild:** Barely perceptible erythema which is faint or patchy
2	**Moderate:** Distinct erythema,
3	**Severe:** Intense (fiery red) erythema

Secondary endpoint that is included in the multiple testing scheme is

- Achievement of an Overall Assessment of Erythema (0-3 scale) score of 0 at Week 8

The analysis of Erythema of 0 is performed similar to the primary endpoint including the sensitivity analysis on both imputed and observed data.

Analysis of efficacy endpoints related to erythema score will be performed using ITT population using observed value:

- The number and percentage of subjects in each category will be summarized by treatment and study visit descriptively
- An ANCOVA model similar to what is defined in Section 8.2.3 will be used for analysis of the change and percent change from Erythema baseline.

8.2.5. Overall Assessment of Scaling

This assessment is performed at Screening, Baseline, and Weeks 2, 4, and 8. The score is

reported in ordinal scale with severity grades ranging from 0-3.

Table 5 illustrates description of scaling grades.

Table 4: Scaling Grades

Score	Description
0	**None:** No scaling evident on lesions
1	**Mild:** Barely detectable, scattered, small flaking scales
2	**Moderate:** Scales clearly visible and prominent
3	**Severe:** Coarse, thick scales, with flaking into clothes or skin

The analysis of Scaling will be performed similar to those of erythema score as indicated in Section 8.2.3.

Separate listings of IGA, average weekly WI-NRS, daily WI-NRS score, Overall assessment of Erythema, and overall assessment of scale score will be provided along with IGA success, and WI-NRS success flag.

Table 6: Summary of Primary and Secondary Efficacy Analyses

Efficacy Endpoint	Primary Analysis	Sensitivity or Supplemental Analysis
Primary		
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 1 (α=0.0033): Average weekly WI-NRS Success (achievement of at least a 4-point improvement) at Weeks 2, 4, 8	PRU4-ITT population, multiple imputation (CMH)	#1 PRU4-ITT, observed data (CMH) #2 PRU4-ITT, non-responder imputation (CMH)
Family 2 (α=0.0067): IGA Success at weeks 4	ITT, multiple imputation (CMH, odds ratio)	#1 ITT, observed data (CMH) #2 ITT, non-responder imputation (CMH)
Family 2 (α=0.0067): Scaling score of '0' at weeks 8	ITT, multiple imputation (CMH, odds ratio)	#1 ITT, observed data (CMH) #2 ITT, non-responder imputation (CMH)
Family 2 (α=0.0067): Erythema Score of '0' at Week 8	ITT, multiple imputation (CMH, odds ratio)	#1 ITT, observed data (CMH) #2 ITT, non-responder imputation (CMH)
Family 2 (α=0.0067): IGA score of '0' at Week 8	ITT, multiple imputation (CMH)	#1 ITT, observed data (CMH) #2 ITT, non-responder imputation (CMH)
Family 2 (α=0.0067): IGA Success at Week 2	ITT, multiple imputation (CMH)	#1 ITT, observed data (CMH) #2 ITT, non-responder imputation (CMH)

8.3. Other Efficacy Analysis

Other secondary efficacy analysis will be performed for subjects belonging to ITT population on the observed data.

The other secondary efficacy endpoints include:

- Change and percent change in Scalpdex total score from baseline at Weeks 2, 4, and 8.
- Achievement of an Overall Assessment of Erythema (0-3 scale) score of 0 at Weeks 2, and 4.
- Achievement of an Overall Assessment of Scaling (0-3 scale) score of 0 at Weeks 2, and 4.

- Change and percent change from baseline in DLQI/CDLQI at Weeks 2, 4, and 8.
- Change and percent change from baseline in BSA affected at Weeks 2, 4, and 8.
- Change and percent change from baseline in Overall Assessment of Erythema at Weeks 2, 4, and 8.
- Change and percent change from baseline in Overall Assessment of Scaling at Weeks 2, 4, and 8.
- Achievement of an IGA score of 'clear' at Week 4.
- Achievement of an IGA score of 'clear' at Week 2.
- Achievement of daily WI-NRS score of 0 or 1 at Weeks 2, 4, 8.
- Change and percent change in WI-NRS from average weekly baseline score.

All the categorical analyses involving achievement of Erythema, Scaling, IGA, WI-NRS are summarized using CMH and proportion difference stratified by pooled site and baseline IGA and presented by study visit.

Continuous endpoints including change and percent change from baseline in average weekly WI-NRS, % BSA, Scalpdex, DLQI/CDLQI, Erythema, Scaling at Week 2, Week 4 and Week 8 will be analyzed using an analysis of covariance (ANCOVA) with the factors of treatment, pooled study site, baseline IGA score, and baseline of the variable under analysis. Similar analysis will be performed for daily WI-NRS every day in the first week and at Week 2, Week 4 and Week 8. Statistical comparisons between the treatment groups will be obtained using contrasts. The LS Means, standard errors, 95% and 99% confidence intervals, and p-values will be presented at each visit. These analyses will be performed on the ITT populations.

Scalpdex questionnaire consists of 23 questions which are categorized into emotions, symptoms and functioning scales. The scale scores and total score are calculated as described in the Section 6.1.7. Analyses of Scalpdex will be limited to subjects in the ITT population who have Seborrheic Dermatitis involvement on the scalp. Change and percent change from baseline for Scalpdex scale scores along with total score are summarized descriptively. In addition, number and percentages for each Scalpdex question is summarized by study visit and treatment group. An ANCOVA model similar to what is defined in Section 8.2.2 will be used for analysis of the change and percent change from Scalpdex (total score, emotions, symptoms and functioning scales) baseline.

Change and percent change from baseline for Overall Assessment of Erythema, Scaling, DLQI, CDLQI, and BSA will be summarized descriptively.

8.4. Subgroup Analyses of Efficacy Variables

For the primary endpoint of IGA success (IGA of 'Clear' or 'Almost Clear' plus at least 2-point improvement at Week 8) and tested secondary endpoints, subgroup analysis will be performed using the primary estimand on the following:

- Topical Corticosteroids – Inadequate Response, intolerance or contraindication (Y/N)
- Topical Antifungals – Inadequate Response, intolerance or contraindication (Y/N)
- By age categories (9-17, 18-64 and ≥ 65 years)
- By IGA at baseline
- Gender (male, female)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- Race (White, Black, Others)
- Baseline BSA% (< 5% vs. ≥ 5%)
- Baseline BSA% based on tertiles
- Fitzpatrick score of I-III vs IV-VI
- Subjects with facial involvement only vs. subjects with scalp involvement only vs. subjects with facial and scalp involvement

The above subgroup analyses are generated for imputed data only. If odds ratio is not estimable due to zero counts, associated C.I. will not be provided as well. Specifically, if a result is non-estimable, then the CMH test will not include the strata for study site and will not include the strata for baseline disease severity. A footnote in the summary table will indicate this change. For subgroup analysis, no p-value will be provided. Forest plots of subgroup analysis for primary efficacy endpoint and secondary endpoints will be provided.

9. SAFETY AND TOLERABILITY ANALYSIS

All safety analyses will be performed on the Safety population. No inferential statistical tests will be performed.

9.1. Adverse Events

All the AEs will be coded using the MedDRA dictionary version 24.1.

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 Discontinuation

A summary of incidence rates (frequencies and percentages) of TEAEs leading to discontinuation of the study drug, by treatment group, SOC, and preferred term will be provided for the Safety Population. Additionally, this table will be repeated for TEAEs leading to study discontinuation.

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 subject with each score) as well as continuous methods (e.g., mean, median, etc.). Categorical summaries will be provided for dermal response as well as other effects.

Investigator tolerability assessment is performed before the subject is dosed on Baseline Day 0. Subject tolerability assessment is performed 10-15 minutes after the application of IP. Subjects not meeting this condition are excluded from table summaries for tolerability assessments.

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 and 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 laboratory values will be listed.

9.4. Vital Signs

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

Changes in weight by treatment group will summarize 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 by treatment group and visit.

BMI is derived as specified in Section 6.1.7. Shift tables (overall, subjects who indicated weight loss was intentional and subjects weight loss was not intentional at each visit) by treatment group for subjects who shift from their baseline BMI category (underweight, normal, overweight, obese) to a different BMI category throughout the course of the study will be provided by treatment group and visit.

9.5. PHQ-8 and Modified PHQ-A

PHQ-8 will be completed by adult subjects and modified PHQ-8 is filled by adolescent subjects of age 12-17 years inclusive. 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 time point by treatment group will be presented.

9.6. CDI-2

Parents/caregivers will complete the CDI-2 for subjects belonging to 9-11 years of age inclusive. For Children's Depression Inventory 2(CDI-2), descriptive summaries of observed values and changes from baseline will be calculated for the total score and the 2 scales emotional problems and functional problems by treatment group and visit.

9.7. 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. From 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.

9.8. 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.

9.9. Concomitant Medication

Prior and concomitant medications will be summarized descriptively by treatment group, ATC 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 WHODrug Global B3, version September 2021.

9.10. Pigmentation Assessment

Hypopigmentation and hyperpigmentation are scored individually on a 0-3 scale where '0' indicates none, '1' for mild, '2' for moderate and '3' for severe. A shift table of hypopigmentation at baseline and at each study visit will be provided by treatment group as well as the shift table of hyperpigmentation. Proportion of subjects who achieves no hypopigmentation and hyperpigmentation at each study visit will be summarized for those subjects who presents of Hypopigmentation or/and hyperpigmentation at baseline. Categorical summary of pigmentation assessment will be provided by study visit and treatment group stratified by race groups (white, others, overall).

10. CHANGES FROM PLANNED ANALYSIS

10.1. Version 2.0

Version 2.0 of the SAP is based on the final signed version of Amendment 2 for the ARQ-154-304 protocol. At the time of signing this SAP version 2.0, the protocol is in the process of being published and is planned to undergo IRB review and approval after database lock.

The intent of this SAP is to document, prior to study database lock and unblinding, the changes specified in protocol amendment 2 regarding the testing strategy used for the secondary endpoints and also other minor updates to analyses not detailed in the protocol. Any changes to the final protocol or additional analyses created after database lock and unblinding will be clearly labeled as such in the CSR.

10.2. 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.

11. OTHER PLANNED ANALYSIS

11.1. Pharmacokinetic Analysis

Concentration data will be summarized by timepoint and treatment group using descriptive statistics, reporting n, mean, standard deviation, median, Q1, Q3, minimum, and maximum, and geometric statistics including geometric mean and coefficient of variation. For computation of mean plasma concentrations, data that are below the limit of quantification (BLQ) will be set to zero. The PK population will be used for these analyses.

All PK collection information from the eCRF will be presented in a listing. PK concentration results will be listed as well.

12. CHANGES TO ANALYSIS PLANNED IN THE PROTOCOL

12.1. Definition of PK Population

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.

13. REFERENCES

Blauvelt A., et al. Efficacy and safety of roflumilast foam 0.3% in patients with seborrheic dermatitis in a phase 3 trial. European Academy of Dermatology and Venereology (EADV) Congress, Sept 7-11, Milano, Italy. 2022. Abstract 3531 and presentation.

ASA. (2018) Ethical Guidelines for Statistical Practice. Prepared by the Committee on Professional Ethics, April 2018. http://www.amstat.org/about/ethicalguidelines.cfm

ICH (1998). ICH Harmonised Tripartite Guideline. Statistical Principles for Clinical Trials E9; 1998. https://database.ich.org/sites/default/files/E9_Guideline.pdf

RSS. (2014) The Royal Statistical Society: Code of Conduct, 2014. https://rss.org.uk/about/policy-and-guidelines/code-of-conduct/.

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

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-154-304. 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
14.1.2.2	Safety	Demographics and Baseline Characteristics
14.1.3	Safety	Previous Treatment History for Seborrheic Dermatitis
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	Exposure to Investigational Product
14.2	**Efficacy Summary**	
14.2.1	ITT	Hypothesis Testing for Primary and Secondary Efficacy Endpoints – Multiple Imputation
14.2.1.1	ITT	Investigator Global Assessment (IGA) Grades and Success by Study Visit – Multiple Imputation (Primary Analysis) Categorical Results
14.2.1.1.s	ITT	Investigator Global Assessment (IGA) Grades and Success by Study Visit – Multiple Imputation (Primary Analysis, Supplemental Estimand) Categorical Results
14.2.1.2.1	PP	Investigator Global Assessment (IGA) Grades and Success by Study Visit – Observed Data (Supplemental Analysis) Categorical Results
14.2.1.2.2	mITT	Investigator Global Assessment (IGA) Grades and Success by Study Visit – Multiple Imputation (Sensitivity Analysis) Categorical Results
14.2.1.2.3	ITT	Investigator Global Assessment (IGA) Grades and Success by Study Visit – Observed Data (Sensitivity Analysis) Categorical Results
14.2.1.2.4	ITT	Investigator Global Assessment (IGA) Grades and Success by Study Visit – Non-responder Imputation (Sensitivity Analysis) Categorical Results
14.2.1.2.5	ITT	Investigator Global Assessment (IGA) Grades and Success at Week 8 – Tipping Point Analysis (Sensitivity Analysis)
14.2.1.2.6.1	ITT	Investigator Global Assessment (IGA) Success at Week 8 by Site – Observed data (Supporting Analysis)

Table Number	Population	Table Title/Summary
14.2.1.2.6.2	ITT	Investigator Global Assessment (IGA) Success at Week 8 – Multiple Imputation, Impact of Site (Supporting Analysis)
14.2.1.2.7	ITT	Investigator Global Assessment (IGA) Grades and Success at Week 8 – Multiple Imputation, without Site Stratification (Sensitivity Analysis) Categorical Results
14.2.1.3.1	ITT	Investigator Global Assessment (IGA) Grades and Success by Study Visit and Response to Topical Corticosteroids – Multiple Imputation (Subgroup Analysis)
14.2.1.3.2	ITT	Investigator Global Assessment (IGA) Grades and Success by Study Visit and Response to Topical Antifungals – Multiple Imputation (Subgroup Analysis)
14.2.1.3.3	ITT	Investigator Global Assessment (IGA) Grades and Success by Study Visit by Study Visit and Age Group – Multiple Imputation (Subgroup Analysis)
14.2.1.3.4	ITT	Investigator Global Assessment (IGA) Grades and Success by Study Visit by Study Visit and Baseline IGA Disease Severity – Multiple Imputation (Subgroup Analysis)
14.2.1.3.5	ITT	Investigator Global Assessment (IGA) Grades and Success by Study Visit and Gender – Multiple Imputation (Subgroup Analysis)
14.2.1.3.6	ITT	Investigator Global Assessment (IGA) Grades and Success by Study Visit and Ethnicity – Multiple Imputation (Subgroup Analysis)
14.2.1.3.7	ITT	Investigator Global Assessment (IGA) Grades and Success by Study Visit and Race – Multiple Imputation (Subgroup Analysis)
14.2.1.3.8	ITT	Investigator Global Assessment (IGA) Grades and Success by Study Visit and Baseline BSA% – Multiple Imputation (Subgroup Analysis)
14.2.1.3.9	ITT	Investigator Global Assessment (IGA) Grades and Success by Study Visit and Baseline BSA% Tertiles – Multiple Imputation (Subgroup Analysis)
14.2.1.3.10	ITT	Investigator Global Assessment (IGA) Grades and Success by Study Visit and Fitzpatrick Score – Multiple Imputation (Subgroup Analysis)
14.2.1.3.11	ITT	Investigator Global Assessment (IGA) Grades and Success by Study Visit and Body Part Involvement – Multiple Imputation (Subgroup Analysis)
14.2.1.4	ITT	Summary of Investigator Global Assessment (IGA) by Study Visit – Observed Data
14.2.2.1	PRU4-ITT	Worst Itch-Numeric Rating Scale (WI-NRS) Success by Study Visit – Multiple Imputation (Primary Analysis)
14.2.2.1.s	PRU4-ITT	Worst Itch-Numeric Rating Scale (WI-NRS) Success by Study Visit – Multiple Imputation (Primary Analysis, Supplemental Estimand)
14.2.2.2.1	PRU4-ITT	Worst Itch-Numeric Rating Scale (WI-NRS) Success by Study Visit – Observed Data (Sensitivity Analysis)

Table Number	Population	Table Title/Summary
14.2.2.2.2	PRU4-ITT	Worst Itch-Numeric Rating Scale (WI-NRS) Success by Study Visit - Non-responder Imputation (Sensitivity Analysis)
14.2.2.3.1	PRU4-ITT	Worst Itch-Numeric Rating Scale (WI-NRS) Success by Study Visit and Response to Topical Corticosteroids - Multiple Imputation (Subgroup Analysis)
14.2.2.3.2	PRU4-ITT	Worst Itch-Numeric Rating Scale (WI-NRS) Success by Study Visit and Response to Topical Antifungals - Multiple Imputation (Subgroup Analysis)
14.2.2.3.3	PRU4-ITT	Worst Itch-Numeric Rating Scale (WI-NRS) Success by Study Visit and Age Group - Multiple Imputation (Subgroup Analysis)
14.2.2.3.4	PRU4-ITT	Worst Itch-Numeric Rating Scale (WI-NRS) Success by Study Visit and Baseline IGA Disease Severity - Multiple Imputation (Subgroup Analysis)
14.2.2.3.5	PRU4-ITT	Worst Itch-Numeric Rating Scale (WI-NRS) Success by Study Visit and Gender - Multiple Imputation (Subgroup Analysis)
14.2.2.3.6	PRU4-ITT	Worst Itch-Numeric Rating Scale (WI-NRS) Success by Study Visit and Ethnicity - Multiple Imputation (Subgroup Analysis)
14.2.2.3.7	PRU4-ITT	Worst Itch-Numeric Rating Scale (WI-NRS) by Study Visit and Race - Multiple Imputation (Subgroup Analysis)
14.2.2.3.8	PRU4-ITT	Worst Itch-Numeric Rating Scale (WI-NRS) Success by Study Visit and Baseline BSA% - Multiple Imputation (Subgroup Analysis)
14.2.2.3.9	PRU4-ITT	Worst Itch-Numeric Rating Scale (WI-NRS) Success by Study Visit and Baseline BSA% Tertiles - Multiple Imputation (Subgroup Analysis)
14.2.2.3.10	PRU4-ITT	Worst Itch-Numeric Rating Scale (WI-NRS) Success by Study Visit and Fitzpatrick Score - Multiple Imputation (Subgroup Analysis)
14.2.2.3.11	PRU4-ITT	Worst Itch-Numeric Rating Scale (WI-NRS) Success by Study Visit and Body Part Involvement - Multiple Imputation (Subgroup Analysis)
14.2.2.4	ITT	Achievement of Daily Worst Itch-Numeric Rating Scale (WI-NRS) 0 or 1 by Study Visit - Observed Data
14.2.2.5	ITT	Summary of Average Weekly Worst Itch-Numeric Rating Scale (WI-NRS) by Study Visit - Observed Data
14.2.2.6	ITT	Change from Baseline in Average Weekly Worst Itch-Numeric Rating Scale (WI-NRS) by Study Week – Observed Data (ANCOVA)
14.2.2.7	ITT	Worst Itch-Numeric Rating Score (WI-NRS) by Study Visit Categorical Results
14.2.2.8	ITT	Change from Baseline in Daily Worst Itch-Numeric Rating Scale (WI-NRS) by Study Week - Observed Data (ANCOVA)
14.2.3.1	ITT	Overall Assessment of Erythema Severity Grades by Study Visit - Multiple Imputation (Primary Analysis) Categorical Results

Table Number	Population	Table Title/Summary
14.2.3.1.s	ITT	Overall Assessment of Erythema Severity Grades by Study Visit - Multiple Imputation (Primary Analysis, Supplemental Estimand) Categorical Results
14.2.3.2.1	ITT	Overall Assessment of Erythema Severity Grades by Study Visit – Observed Data (Sensitivity Analysis) Categorical Results
14.2.3.2.2	ITT	Overall Assessment of Erythema Score of 0 at Week 8 - Non-responder Imputation – (Sensitivity Analysis)
14.2.3.3.1	ITT	Overall Assessment of Erythema Score of 0 at Week 8 by Response to Topical Corticosteroids - Multiple Imputation (Subgroup Analysis)
14.2.3.3.2	ITT	Overall Assessment of Erythema Score of 0 at Week 8 by Response to Topical Antifungals - Multiple Imputation (Subgroup Analysis)
14.2.3.3.3	ITT	Overall Assessment of Erythema Score of 0 at Week 8 by Age Group - Multiple Imputation (Subgroup Analysis)
14.2.3.3.4	ITT	Overall Assessment of Erythema Score of 0 at Week 8 by Baseline IGA Severity - Multiple Imputation (Subgroup Analysis)
14.2.3.3.5	ITT	Overall Assessment of Erythema Score of 0 at Week 8 by Gender - Multiple Imputation (Subgroup Analysis)
14.2.3.3.6	ITT	Overall Assessment of Erythema Score of 0 at 8 by Ethnicity - Multiple Imputation (Subgroup Analysis)
14.2.3.3.7	ITT	Overall Assessment of Erythema Score of 0 at Week 8 by Race - Multiple Imputation (Subgroup Analysis)
14.2.3.3.8	ITT	Overall Assessment of Erythema Score of 0 at Week 8 by Baseline BSA% - Multiple Imputation (Subgroup Analysis)
14.2.3.3.9	ITT	Overall Assessment of Erythema Score of 0 at Week 8 by Baseline BSA% Tertiles - Multiple Imputation (Subgroup Analysis
14.2.3.3.10	ITT	Overall Assessment of Erythema Score of 0 at Week 8 by Fitzpatrick Score - Multiple Imputation (Subgroup Analysis)
14.2.3.3.11	ITT	Overall Assessment of Erythema Score of 0 at Week 8 by Body Part Involvement - Multiple Imputation (Subgroup Analysis)
14.2.3.4	ITT	Summary of Overall Assessment of Erythema by Study Visit – Observed Data
14.2.3.5	ITT	Change from Baseline in Overall Assessment of Erythema by Study Visit – Observed Data (ANCOVA)
14.2.4.1	ITT	Overall Assessment of Scaling Severity Grades by Study Visit - Multiple Imputation (Primary Analysis)
14.2.4.1.s	ITT	Overall Assessment of Scaling Severity Grades by Study Visit - Multiple Imputation (Primary Analysis, Supplemental Estimand)
14.2.4.2.1	ITT	Overall Assessment of Scaling Severity Grades by Study Visit – Observed Data (Sensitivity Analysis) Categorical Results

Table Number	Population	Table Title/Summary
14.2.4.2.2	ITT	Overall Assessment of Scaling Score of 0 at Week 8 – Non-responder Imputation (Sensitivity Analysis) Categorical Results
14.2.4.3.1	ITT	Overall Assessment of Scaling Score of 0 at Week 8 by Response to Topical Corticosteroids - Multiple Imputation (Subgroup Analysis)
14.2.4.3.2	ITT	Overall Assessment of Scaling Score of 0 at Week 8 by Response to Topical Antifungals - Multiple Imputation (Subgroup Analysis)
14.2.4.3.3	ITT	Overall Assessment of Scaling Score of 0 at Week 8 by Age Group - Multiple Imputation (Subgroup Analysis)
14.2.4.3.4	ITT	Overall Assessment of Scaling Score of 0 at Week 8 by Age Group - Multiple Imputation (Subgroup Analysis)
14.2.4.3.5	ITT	Overall Assessment of Scaling Score of 0 at Week 8 by Gender - Multiple Imputation (Subgroup Analysis)
14.2.4.3.6	ITT	Overall Assessment of Scaling Score of 0 at Week 8 by Ethnicity- Multiple Imputation (Subgroup Analysis)
14.2.4.3.7	ITT	Overall Assessment of Scaling Score of 0 at Week 8 by Race- Multiple Imputation (Subgroup Analysis)
14.2.4.3.8	ITT	Overall Assessment of Scaling Score of 0 at Week 8 by Baseline BSA% - Multiple Imputation (Subgroup Analysis)
14.2.4.3.9	ITT	Overall Assessment of Scaling Score of 0 at Week 8 by Baseline BSA% Tertiles - Multiple Imputation (Subgroup Analysis)
14.2.4.3.10	ITT	Overall Assessment of Scaling Score of 0 at Week 8 by Fitzpatrick Score - Multiple Imputation (Subgroup Analysis)
14.2.4.3.11	ITT	Overall Assessment of Scaling Score of 0 at Week 8 by Body Part Involvement- Multiple Imputation (Subgroup Analysis)
14.2.4.4	ITT	Summary of Overall Assessment of Scaling by Study Visit - Observed Data
14.2.4.5	ITT	Change from Baseline in Overall Assessment of Scaling by Study Visit – Observed Data (ANCOVA)
14.2.5.1	ITT*	Summary of Scalpdex Total Score by Study Visit - Observed Data
14.2.5.2	ITT*	Change from Baseline in Scalpdex Total Score by Study Visit - Observed Data (ANCOVA)
14.2.5.3	ITT*	Summary of Scalpdex Score - Emotions Scale by Study Visit - Observed Data
14.2.5.4	ITT*	Change from Baseline in Scalpdex Score - Emotions Scale by Study Visit - Observed Data (ANCOVA)
14.2.5.5	ITT*	Summary of Scalpdex Score - Symptoms Scale by Study Visit - Observed Data
14.2.5.6	ITT*	Change from Baseline in Scalpdex Score - Symptoms Scale by Study Visit - Observed Data (ANCOVA)
14.2.5.7	ITT*	Summary of Scalpdex Score - Functioning Scale by Study Visit - Observed Data
14.2.5.8	ITT*	Change from Baseline in Scalpdex Score - Functioning Scale by Study Visit - Observed Data (ANCOVA)

Table Number	Population	Table Title/Summary
14.2.5.9	ITT*	Scalpdex Individual Questionnaire Responses by Study Visit (Categorical)
14.2.5.10	ITT*	Scalpdex Individual Questionnaire Responses by Study Visit (Continuous)
14.2.6.1	ITT	Summary of Dermatology Life Quality Index (DLQI) by Study Visit - Observed Data
14.2.6.2	ITT	Change from Baseline in Dermatology Life Quality Index (DLQI) by Study Visit - Observed Data (ANCOVA)
14.2.6.3	ITT	Summary of Children's Dermatology Life Quality Index (CDLQI) by Study Visit - Observed Data
14.2.6.4	ITT	Change from Baseline in Children's Dermatology Life Quality Index (CDLQI) by Study Visit - Observed Data (ANCOVA)
14.2.7.1	ITT	Summary of Body Surface Area (%) by Study Visit - Observed Data
14.2.7.2	ITT	Change from Baseline in Body Surface Area (%) by Study Visit - Observed Data (ANCOVA)
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 Strongest Relationship to Study Drug
14.3.1.5	Safety	Treatment Emergent Adverse Events of Preferred Term in Descending Order
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 Categorical Results
14.3.2.3	Safety	Categorical Summary of Investigator Local Tolerability Assessment (Other Effects) by Study Visit Categorical Results
14.3.2.4	Safety	Quantitative Summary of Subject Local Tolerability Assessment by Study Visit

Table Number	Population	Table Title/Summary
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 (Standard Units) 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	Shift from Baseline in BMI by Study Visit
14.3.4.4	Safety	Change in Weight from Baseline by Study Visit – by Weight Loss Intentional/Non-Intentional
14.3.4.5	Safety	Shift from Baseline in BMI by Study Visit - by Weight Loss Intentional/Non-Intentional
14.3.5	Safety	Shift from Baseline in Patient Health Questionnaire (PHQ-8)/Modified PHQ-Adolescents by Study Visit
14.3.6	Safety	Summary of CDI-2 by Study Visit
14.3.7	Safety	Summary of Columbia-Suicide Severity Rating Scale (C-SSRS) Suicidal Ideation and Suicidal Behavior by Study Visit
14.3.8	Safety	Summary of Physical Examination by Study Visit
14.3.9	Safety	Summary of Concomitant Medications by Anatomic Therapeutic Chemical (ATC) Class Level 4 and Preferred Term
14.3.10	Safety	Summary of Pigmentation Assessment by Race
14.3.10.1	Safety	Shift from Baseline in Hypopigmentation by Study Visit
14.3.10.2	Safety	Shift from Baseline in Hyperpigmentation by Study Visit
14.4	**Pharmacokinetic Data**	
14.4.1	PK	Summary of Pharmacokinetic Concentration Results by Study Visit
*Analyses of Scalpdex will be limited to subjects in the ITT population who have Seborrheic Dermatitis involvement on the scalp		

14.2. Planned Figure descriptions

The following are planned summary figures for protocol ARQ-154-304. The figure numbers are place holders only and may be changed when figures are produced.

Figure Number	Population	Figure Title/Summary
14.2.1.4	ITT	Forest Plot of IGA Success at Week 8 - Multiple Imputation (Subgroup Analysis)
14.2.1.3.1.1	ITT	Plot of Proportion (99% CI) of Investigator Global Assessment (IGA) Success at Week 8 by Site - Observed Data (Supporting Analysis)
14.2.1.3.2.1	ITT	Plot of Proportion (99% CI) of Investigator Global Assessment (IGA) Success at Week 8 - Multiple Imputation, Impact of Site (Supporting Analysis)
14.2.2.9.1	ITT	Plot of Mean (+/-SD) Change, and Percent Change from Baseline of Daily Worst Itch-Numeric Rating Scale (WI-NRS) by Treatment Group over Time
14.2.2.3	ITT	Forest Plot of Worst Itch-Numeric Rating Scale (WI-NRS) Success by Study Week - Multiple Imputation (Subgroup Analysis)
14.2.3.3	ITT	Forest Plot of Overall Assessment of Erythema Score of 0 at Week 8 - Observed Data (Subgroup Analysis)
14.2.4.3	ITT	Forest Plot of Overall Assessment of Scaling Score of 0 at Week 8 - Observed Data (Subgroup Analysis)
14.3.1.2.1	Safety	Plot of Most Frequent Treatment Emergent Adverse Event 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-154-304.

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
16.1.1	Screened Subjects	Subject Disposition
16.1.2	ITT	Protocol Deviations
16.1.3	ITT and Actual Stratification Factors	Randomization
16.1.4	ITT	Subject Demographics and Baseline Characteristics
16.1.5	Safety	Study Drug Application and Can weight at the Study Site
16.2.1	ITT	Efficacy Endpoints

Listing Number	Population	Listing Title/Summary
16.2.2	ITT	Daily Worst Itch Numerical Rating Scale (WI-NRS)
16.3.1	Safety	Adverse Events
16.3.2.1	Safety	Abnormal Clinical Laboratory Data: Clinical Chemistry
16.3.2.2	Safety	Abnormal Clinical Laboratory Data: Hematology
16.3.2.3	Safety	Abnormal Clinical Laboratory Data: Urinalysis
16.3.3	Safety	Abnormal Vital Signs
16.4.1	PK	Pharmacokinetic Blood Collection

15. TABLES, LISTINGS AND FIGURES SHELLS

15.1. Standard Layout for all Tables, Listings and Figures

Tables, listings and figures shells are provided in a separate document.

Note that programming notes may be added after each TLF shell if appropriate.