Sponsor	Arcutis Biotherapeutics, Inc.
Protocol Title:	A Phase 2, Multicenter, Open-Label Study of the Long-Term Safety of ARQ-154 Foam 0.3% in Subjects with Seborrheic Dermatitis
Protocol Number:	ARQ-154-214
Premier Research PCN:	ARCU211880
Document Version:	Version 1.0
Document Date:	18-Feb-2022

Approvals

Role	Signatures
Biostatistician	Print Name:
	Sign Name:
Peer Reviewer	Print Name:
	Sign Name:

Role	Signatures
Arcutis Biotherapeutics, Inc. Representative	Print Name:
	Sign Name:
Arcutis Biotherapeutics, Inc. Representative	Print Name:
	Sign Name:

Document History

Not applicable.

Table of Contents

Contents

List of Abbreviations

Abbreviation	Definition
aCRF	annotated case report form
ADR	adverse drug reactions
AE	adverse event
AESI	adverse events special interest
ANCOVA	analysis of covariance
ASA	American Statistical Association
ATC	anatomical therapeutic chemical
BLQ	below limit of quantification
BMI	body mass index
BSA	Body Surface Area
CDI-2	Children's Depression Inventory 2
CI	confidence interval
CMH	Cochran Mantel Haenszel
COVID-19	novel coronavirus disease
CRF	case report form
CS	clinically significant
CSR	clinical study report
C-SSRS	Columbia Suicide Severity Rating Scale
DBL	database lock
DBP	diastolic blood pressure

Abbreviation	Definition
DLT	dose limiting toxicity
DMC	data monitoring committee
DOB	date of birth
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
ICH	International Conference on Harmonization
IGA	investigator global assessment
IP	investigational product
IRB	institutional review board
IRR	infusion related reactions
IRT	interactive response technology
ITT	intent-to-treat

Abbreviation	Definition
LLOQ	lower limit of quantification
MedDRA	medical dictionary for regulatory activities
MI	multiple imputation
MMRM	mixed model for repeated measurements
MTD	maximum tolerated dose
NA	not applicable
NCS	non-clinically significant
OLE	Open-Label
PHQ-A	Modified PHQ-8 for Adolescents
PHQ-8	Patient Health Questionnaire depression scale
PK	pharmacokinetic
PKAP	pharmacokinetic analysis plan
PP	per-protocol
PT	preferred term
RR	respiratory rate or relative rate
RSS	Royal Statistical Society
SAE	serious adverse event
SAP	statistical analysis plan
SAS	Statistical Analysis System
SBP	systolic blood pressure

Abbreviation	Definition
SD	standard deviation
SE	standard error
SOC	system organ class
SOP	standard operating procedure
SUSAR	suspected, unexpected, serious adverse (drug) reaction
TEAE	treatment-emergent adverse event
TESAE	treatment-emergent serious adverse event
TLF	Tables, Listings and Figures
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-154-214 (A Phase 2, Multicenter, Open-Label Study of the Long-Term Safety of ARQ-154 Foam 0.3% in Subjects with Seborrheic Dermatitis), dated 04-Nov-2021 Amendment 2. 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 inferential or descriptive analysis of data pertaining to Arcutis Biotherapeutics, Inc.'s study ARQ-154-214.

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

2. Study Objectives and Endpoints

2.1. Study Objectives

2.1.1. Primary Objective

The objective of this study is to assess long-term safety in a multicenter, open-label, single-arm 52-week study in subjects with seborrheic dermatitis treated with roflumilast foam 0.3%.

2.2. Study Endpoints

2.2.1. Efficacy Endpoints

2.2.1.1. Primary Efficacy Endpoint

Not applicable.

2.2.1.2. Secondary Efficacy Endpoints

The secondary efficacy endpoints of this study include the following:

- Achievement of an Investigator Global Assessment (IGA) of "completely clear" or "almost clear", as observed over time.
- "IGA success", defined as achievement of an IGA of "completely clear" or "almost clear" plus a 2-grade improvement in IGA from Baseline as observed over time.
- Duration of IGA success, defined as the time from the first observation of IGA success to the first subsequent time a subject's disease response does not meet the criteria for IGA success. The duration of IGA success for subjects who end treatment in IGA success will be censored at the last disease assessment date.
- Treatment-free interval, defined as time from when subject achieves disease clearance (IGA of 0, referred to as "completely clear") and stop treatment of all lesions to the time when treatment restarts due to IGA of 2 or above.

2.2.1.3. Other Secondary Efficacy Endpoints

Other secondary efficacy endpoints of this study include the following:

- Change and percent change in Scalpdex total score as observed over time.
- Change and percent change from baseline in Overall Assessment of Erythema score
- Change and percent change from baseline in Overall Assessment of Scaling score
- Achievement of an Overall Assessment of Erythema (0-3 scale) score of 0 or 1 plus a 2-grade improvement from Baseline as observed over time.
- Achievement of an Overall Assessment of Scaling (0-3 scale) score of 0 or 1 plus a 2-grade improvement from Baseline as observed over time.
- Achievement of an Overall Assessment of Erythema (0-3 scale) score of 0 as observed over time.
- Achievement of an Overall Assessment of Scaling (0-3 scale) score of 0 as observed over time.

- Change and percent change from baseline in % body surface area (BSA) affected by disease as observed over time.
- Change and percent change in Worst Itch – Numeric Rating Score (WI-NRS) as observed over time.
- In subjects with a Baseline WI-NRS pruritus score of ≥4, achievement of a ≥4-point improvement from Baseline in WI-NRS pruritus score as observed over time.
- Achievement of Itch-free state which is defined as post-baseline WI-NRS of 0 or 1

2.2.2. Safety Endpoints

The safety endpoints will be analysis of safety monitored through application site assessments in the clinic using the method of Berger and Bowman, clinical laboratory testing, Patient Health Questionnaire depression Scale (PHQ-8), Modified PHQ-Adolescents (PHQ-A), Children's Depression Inventory 2 (CDI-2), Columbia-Suicide Severity Rating Scale (C-SSRS), and adverse events (AEs). Two primary safety endpoint analyses are planned:

- Occurrence of treatment emergent AEs (TEAEs).
- Occurrence of serious AEs (SAEs).

3. Overall Study Design and Plan

3.1. Overall Design

This is an open-label, single-arm, long-term 52-week safety study of subjects with seborrheic dermatitis. This study will include both male and female subjects 9 years and older. Subjects rolling over from ARQ-154-116 (Cohort 1 Group 2) will not be summarized in the tables and would be listed only.

Cohort 1: This includes subjects of age 9 years or older from ARQ-154-203 and ARQ-154-116 studies and rolled into this extension study on the final visit from previous study. Subjects rolling over from ARQ-154-203 and ARQ-154-116 are separated into groups as indicated below.

Group 1: Subjects rolled over from ARQ-154-203 without gap

Group 2: Subjects rolled over from ARQ-154-116 (listings only and will not appear in summary tables)

Cohort 2: This includes subjects of age 12 years or older who may or may not have participated in a prior ARQ-154 study. The subjects could be either de novo subjects (requiring an IGA of disease severity of at least Moderate ["3"] at Baseline) or subjects who completed an ARQ-154 study before the initiation of this study (any IGA at Day 1 of this

study is allowed). Subjects in Cohort 2 are separated into groups as indicated below:

Group 3: Subjects rolled over from ARQ-154-203 with gap

Group 4: De novo subjects

Subjects will apply study drug once a day for up to 52 weeks. Subjects in both Cohorts cannot have more than 20% BSA affected by disease. All lesions on a subject will be treated including the scalp, face, trunk, and intertriginous areas.

3.2. Sample Size and Power

A sample size of approximately 410 subjects is planned for the study. This sample size is not based on any calculations of power or precision. This sample size will provide a sufficient population size to evaluate the long-term safety of roflumilast foam 0.3% over up to 52 weeks of treatment, and in the combination with other studies provide the development program with sufficient number of subjects to meet ICH exposure goals.

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.

3.4. Method of Assigning Treatment

Not applicable.

3.5. Treatment Administeration

All subjects will be treated with roflumilast foam 0.3% daily (QD).

3.6. Blinding and Unblinding

This is an open-label study and no blinding is planned.

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	Washout Cohort 2	Day 1 (Cohort 2)	Wk 4 D 29	Wk 12 D 85	Wk 18 D 127	Wk 24 D 169	Wk 30 D 211	Wk 36 D 253	Wk 44 D 309	Wk 52/ ET D 365
Visit	Screening	Final Visit of prior ARQ-154 Study l (Cohort 1)	1	2	Phone Visit	3	Phone Visit	4	Phone Visit	5
Visit Window	-35 days		+/- 3 days	+/- 5 days	+/- 7 days	+/- 7 days	+/- 7 days	+/- 7 Days	+/- 7 days	+/- 7 days
Informed consent/assent	X[k]	X[k]								
Physical examination[a]	X	X[l]				X				X
Medical history	X									
I/E criteria	X	X								
Hematology, Serum Chemistries, and Urine Analysis	X	X[l]	X	X		X		X		X
Vital signs, height, weight[b]	X	X[l]	X	X		X		X		X
IGA[c], Overall Assessment of Erythema[c], Overall Assessment of Scaling[c] , WI-NRS, Scalpdex	X	X[l]	X	X		X		X		X
BSA	X	X[l]				X		X		X
C-SSRS, PHQ-8/ PHQ-A/CDI-2[d]	X	X[l]	X	X		X		X		X
Pregnancy test[e]	X	X[l]	X	X		X		X		X
Local Tolerability Assessment[f]		X[l]	X	X		X		X		X
Pigmentation Assessment[g]	X	X[l]	X	X		X		X		X
Medical Photography[h]		X	X	X		X		X		X
Dispense study medication kit[i]		X	X	X		X		X		
IP application at the study site		X								
Dispense/review diary		X	X	X		X		X		X
Weigh study medication		X	X	X		X		X		X
Adverse event assessment[j]	X	X	X	X	X	X	X	X	X	X
Concomitant medications	X	X	X	X	X	X	X	X	X	X

Study Procedure	Washout Cohort 2	Day 1 (Cohort 2)	Wk 4 D 29	Wk 12 D 85	Wk 18 D 127	Wk 24 D 169	Wk 30 D 211	Wk 36 D 253	Wk 44 D 309	Wk 52/ ET D 365
Visit	Screening	Final Visit of prior ARQ-154 Study l (Cohort 1)	1	2	Phone Visit	3	Phone Visit	4	Phone Visit	5
Visit Window	-35 days		+/- 3 days	+/- 5 days	+/- 7 days	+/- 7 days	+/- 7 days	+/- 7 Days	+/- 7 days	+/- 7 days
PK Sampling[m]			X							
Malassezia test[n]			X	X		X		X		X

Footnotes from table above:

[a] Limited physical examination: skin, lungs, and heart only

[b] Height will be collected at Day 1, Week 24 and Week 52/ET only. Weight will be collected at all study visits. Subject to void prior to weight being taken. Remove any jackets, outerwear and shoes. Remove any objects of significant weight (i.e. cell phones, wallet, key chains). A 5% unintentional weight loss from Baseline should be reported to the medical monitor.

[c] IGA will be a 5-point scale ranging from completely clear (0) to severe (4). IGA should be completed prior to other physician assessments. Overall assessment of erythema (0-3 scale) and overall assessment of scaling (0-3 scale) will be completed. Total BSA affected by seborrheic dermatitis will be determined at each visit. A body diagram should be used to record areas of seborrheic dermatitis involvement.

[d] Adolescents and adults will complete the C-SSRS (12 years of age and older). Adults will complete the PHQ-8. Adolescents (ages 12 -17, inclusive) will complete the modified PHQ-A. Parents/caregivers will complete CDI-2 (parent report) for children 9-11 years of age, inclusive.

[e] A pregnancy test will be administered to all females of child-bearing potential. A serum pregnancy test will be performed at the Screening visit only for Cohort 2. A urine pregnancy test will be performed at Day 1, Weeks 4, 12, 24, 36 and 52. A negative result is required for continued participation in the study, and results must be available prior to dispensing of IP at each visit.

[f] Local tolerability will be assessed prior to IP application in the clinic for the Investigator assessment of skin (Berger and Bowman skin irritation score) and 10-15 minutes post IP application for the subject's "0-3" burning/stinging assessment at Baseline. Note for Investigator tolerability assessments: **reactions at the site of product application, which may occur post-Baseline, should be differentiated from the preexisting inflammation associated with the subject's seborrheic dermatitis**. At Weeks 4, 12, 24, 36, and 52 will be a recall assessment of burning/stinging experienced post drug application on the previous day of the clinic visit.

[g] An assessment for hypopigmentation and hyperpigmentation will be performed by the investigator at all clinic visits.

[h] At selected sites, medical photography will be obtained for target lesions. All efforts will be made to de-identify the subjects. Subjects who are unwilling to participate in the medical photography will be allowed to opt out of this procedure.

[i] Kits will be dispensed based on %BSA affected. See IP Handling Manual for details.

[j] Any emergent AEs will be followed in the clinic for up to one month at the Investigator's discretion until resolvedor otherwise judged as clinically stable.

[k] For Cohort 1, the consent will be signed after completion of the final visit in prior ARQ-154 study. Cohort 2 willsign consent prior to any study-related procedures at the Screening Visit.

[l] For Cohort 1, this data will be obtained from the final visit of the prior ARQ-154 Study and used as the Day 1 data for this long-term safety study (ARQ-154-214). If after the prior ARQ-154 study completion, subjects should complete the screening visit procedures. The Baseline values for Safety tabulations will be taken from the day that the subject received their first active IP (across prior ARQ-154 studies and ARQ-154-214). Baseline values for efficacy will be those recorded on Day 1 of the prior ARQ-154 Study. For Cohort 2,

Baseline evaluations will be obtained for all procedures and the Baseline/Screening version of the C-SSRS will be utilized at Screening.
For Cohort 2, Screening labs that are collected within 14 days of Day 1 do not need to be repeated.
[m]For Cohort 2 adolescent subjects PK samples will be collected on Week 4 for adolescent subjects between 12 – 16 years old. Collect PK sample with the safety lab collection. Ensure PK will not be drawn on the areawhere investigational product is applied.
[n] At select sites, the Investigator will collect samples for Malassezia testing on Cohort 1 Group 2 subjects at Weeks 4, 12, 24, 36, and 52. Samples will be collected from treated and untreated areas by swabbing the skin surface. Samples should be collected from the same areas as samples collected during the ARQ-154-116 Baseline visit.

4. Statistical Analysis and Reporting

4.1. Introduction

Data processing, tabulation of descriptive statistics, calculation of inferential statistics, and graphical representations will 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 the treatment group unless otherwise specified. The denominator for by-visit displays will be the number of subjects in the relevant study population with non-missing data at each visit.

The minimum and maximum will be reported with the same degree of precision (i.e., the same number of decimal places) as the observed data. Measures of location (mean, median, Q1, and Q3) 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, unless otherwise specified or modified upon request of the sponsor.

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

4.2. Interim Analysis and Data Monitoring

No interim analyses are planned. Once about 400 subjects finish the treatment period in this study, the final analysis will be performed for the full analysis CSR. When all subjects rolled over from study ARQ-154-116 (Cohort 1 Group 2), who enrolled into this ARQ-154-214 study

later than the initial about 400 subjects, finish the treatment in this (ARQ-154-116) study, only listings will be generated for these subjects and these listings will be added as a CSR addendum. No further tables or figures will be generated for these subjects.

5. Analysis Populations

The following analysis populations are planned for this study:

- **All Enrolled (ENRL):** All enrolled population includes subjects who are enrolled into the study (i.e, signed informed consent) and not screen failures.
- **Safety Population (SAF):** The Safety population includes all subjects who are enrolled and received at least 1 confirmed dose of roflumilast foam 0.3% in this extension study.
- **Pruritus Population Primary (PRU4-P):** The PRU4-P population is a subset of the Safety population and includes subjects with WI-NRS pruritus score ≥ 4 at Primary Baseline. This population will be used for the analysis of achievement of a 4-point reduction in WI-NRS pruritus score as compared to Primary Baseline.
- **Pruritus Population Open Label Extension (PRU4-OLE):** The PRU4-OLE population is a subset of the Safety population and includes subjects with WI-NRS pruritus score ≥ 4 at OLE Baseline. This population will be used for the analysis of achievement of a 4-point reduction in WI-NRS pruritus score as compared to OLE Baseline.
- **Pharmacokinetic Population (PK):** The PK population will include all subjects with sufficient plasma concentrations of roflumilast to define a profile, as determined by the pharmacokineticist. This population will be used for analyses of PK parameters.

6. General Issues for Statistical Analysis

For analysis purposes, the tables and figures will be summarized by groups as indicated below.

Cohort 1 Group 1: Subjects rolled over from ARQ-154-203 without gap (Roflumilast Foam 0.3%, Vehicle, Total)

Cohort 2 Group 3: Subjects rolled over from ARQ-154-203 with gap (Roflumilast Foam 0.3%, Vehicle, Total)

Cohort 2 Group 4: De novo subjects

Cohort 2 Total

Overall (includes all Cohorts and groups)

6.1. Statistical Definitions and Algorithms

6.1.1. Baseline

There are three baseline definitions for the analyses:

- Primary baseline is defined as the last observation before the first dose of roflumilast foam 0.3% in either ARQ-154 parent study or this study for subjects administered vehicle previously.
- Open-label baseline (OLE baseline) is defined the last observation before the first dose of roflumilast foam 0.3% in this study. For subjects in Cohort 1 who choose to enroll in this study, study procedures will not be repeated; study sites will carry over the information from the prior ARQ-154 study to the Day 1 visit for this study. For subjects in Cohort 2 Group 4, the definition for Primary and OLE baseline is the same.
- Safety baseline is defined as primary baseline for Cohort 1 Group 1, and OLE baseline for Cohort 2 Group 3 and Group 4.

The following table described the baseline used in the analysis:

Cohort Groups	**Prior therapy, medical history**	**Demog and Efficacy Analysis**	**Safety Analysis for AE, PHQ8, PHQ-A, Prior and Concomitant medications**	**Safety Analysis for lab, vital signs and tolerability assessments**
Cohort 1 Group 1	OLE baseline	Primary baseline and OLE baseline	OLE baseline	Primary baseline
Cohort 2 Group 3	OLE baseline	Primary baseline and OLE baseline	OLE baseline	OLE baseline
Cohort 2 Group 4	OLE baseline	Primary baseline and OLE baseline	OLE baseline	OLE baseline
Cohort 2 Total	OLE baseline	Primary baseline and OLE baseline	OLE baseline	OLE baseline

Overall	OLE baseline	Primary baseline and OLE baseline	OLE basseline	Safety baseline of roflumilast foam 0.3% in this study for subjects in the safety analysis population. For subjects in Cohort 2 Group 4, the definition for Primary and OLE Baselines is the same

Baseline values will most commonly be from Day 1.
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. Multiple Comparisons

No adjustments for multiple comparisons will be made for this study. No inferential statistical analyses are planned for this study.

6.1.3. Handling of Dropouts or Missing Data

Other than what is mentioned in 6.1.8, no data will be imputed in the analyses for this study.

6.1.4. 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 and Table 3. The windows follow the Schedule of Events in Table 1. For subjects in Cohort 2 Group 3, the gap between completion of parent study and starting of the study ARQ-154-214 should not be counted in the analysis visit window. The windowing will be applied using the days on study as shown in the Table 2. Unscheduled/repeated measurements will be listed.

Table 2: Analysis Visit Windows (Primary Baseline for Cohort 1 Group 1 & Cohort 2 Group 3)

	Primary Baseline for Group 1 - Roflumilast & Group 3 - Roflumilast	
Scheduled Visit	**Target Days on Study**	**Window (Days)**

ARQ-154-203 Week 2	14	2 to 22
ARQ-154-203 Week 4	28	23 to 42
ARQ-154-203 Week 8	56	43 to 61
Week 4 OLE	85	71 to 113
Week 12 OLE	141	114 to 183
Week 24 OLE	225	184 to 267
Week 36 OLE	309	268 to 365
Week 52 OLE	421	≥ 366

Table 3: Analysis Visit Windows

	Primary Baseline for Cohort 1 Group 1 - Vehicle & Cohort 2 Group 3 - Vehicle Cohort 2 Group 4 + OLE Baseline for Cohort 1 Group 1 - Roflumilast & Cohort 2 Group 3 – Roflumilast	
Scheduled Visit	**Target Study Day**	**Window (Days)**
ARQ-154-203 Week 2	n.a.	n.a.
ARQ-154-203 Week 4	n.a.	n.a.
ARQ-154-203 Week 8	n.a.	n.a.
Week 4 OLE	29	15 to 57
Week 12 OLE	85	58 to 127
Week 24 OLE	169	128 to 211
Week 36 OLE	253	212 to 309
Week 52 OLE	365	≥ 310

n.a. = not applicable

6.1.5. Pooling of Sites

Not applicable.

6.1.6. Subgroup

The subgroup analyses of selected safety and efficacy endpoints will be based on the below subgroup variables at both primary baseline and OLE baseline:

- Age groups (9-11 years, 12-17 years, 18-64 years, and ≥65 years)
- Gender (Male vs. Female)
- Race (White, Black vs. Other)
- Ethnicity (Hispanic vs. Non-hispanic)
- Baseline IGA (3 vs. 4 for primary baseline, 0 or 1 vs. 2 vs. 3. vs. 4 for OLE baseline)
- Baseline BSA% (<5% vs. ≥5%)
- Baseline BSA% tertiles based on overall group

6.1.7. Derived Variables

- **IGA success =** IGA of "completely clear" or "almost clear" plus a 2-grade improvement from Baseline. For IGA success analysis, the subjects who achieved IGA of 0 or 1 at baseline will be excluded from the analysis.

- **Erythema success =** Erythema score of 0 or 1 plus a 2-grade improvement from Baseline as observed over time. For erythema success analysis, the subjects who achieved erythema score of 0 or 1 at baseline will be excluded from the analysis.

- **Scaling success =** Scaling score of 0 or 1 plus a 2-grade improvement from Baseline as observed over time. For scaling success analysis, the subjects who achieved scaling score of 0 or 1 at baseline will be excluded from the analysis.

- **Compliance** = number of applications divided by the expected number of investigational product (IP) applications for each subject. Compliance will be calculated using drug accountability data over the entire treatment period for each subject, up to treatment completion or discontinuation.

- **Number of expected IP applications** = calculated as last treatment date - first treatment date + 1. Excluding treatment-free interval time (defined in this section).

- **Number of IP applications =** number of expected IP applications – missed IP applications as collected in the CRF. Missed IP applications will not be taken into consideration in determining number of IP applications if the reason for missing dose is per investigator confirmation of clear.

- **Days on IP** = last date of IP application – first date of IP application +1. Subjects that stop taking IP per investigator confirmation of clear will be excluded from this calculation.

- **Treatment Gap** = This will be calculated for subjects belonging to Cohort 2 Group 3 that rolled into the extension study (ARQ-154-214) with a gap. This is derived start treatment start date in ARQ-154-214 – last treatment date in ARQ-154-203 + 1.

- **Treatment Duration (weeks)** = (last date of IP application – first date of IP application +1)/7.

- **Duration of IGA success** = The date of firstobservation where IGA does not meet success criteria - the date of first observation of IGA success + 1. The duration of IGA success for subjects who end treatment in IGA success will be censored at the last disease assessment date. Only the first duration of IGA success will be analyzed.

- **Duration of IGA of 0 =** the date of first observation of IGA > 0 - the date of first observation of IGA of 0 and stop treatment. The duration of IGA of 0 for subjects who end treatment in IGA success will be censored at the last disease assessment date. This is for subjects who achieve IGA of 0 and stop the treatment. Subjects who achieve IGA of 0 but never stop treatment will be excluded from the analysis. For subjects who achieve IGA of 0 but stop the treatment later, the duration starts from the date of stop treatment. Only the first IGA of 0 duration will be analyzed.

- **Duration of IGA of 0 or 1=** the date of first observation of IGA > 1-the date of first observation of IGA of 0 or 1 + 1. The duration of IGA of 0 or 1 for subjects who end treatment in IGA of 0 or 1 will be censored at the last disease assessment date. Only the first IGA of 0 or 1 duration will be analyzed.

- **Disease recurrence** = the date of first observation of IGA of 2 or above after achieving IGA of 0 and stop treatment - the date of first observation of IGA of 0 + 1. Only the first disease recurrence will be analyzed.

- **Treatment-free interval** = the date of re-starting study drug due to IGA score above 0 - the date of stop treatment due to attainment of score of IGA of 0 +1, among subjects who achieve a "completely clear" IGA and stop treatment to all lesions. If subjects never re-start study drug after stopping treatment, the treatment-free interval is censored at the end of study date. Only the first treatment-free interval will be analyzed.

- **Weight of IP (g)** = dispensed can weight – returned can weight.

- **BMI (kg/m^2)** = (weight in kg)/[(height in cm/100)^2]. Day 1 height will be used for BMI derivation upto Week 24. For visits including week 24 through week 36, week 24 height

will be used to derive BMI. Week 52 BMI is derived using the height and weight collected at the same visit. This derivation is applied 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

- **WI-NRS 4-point reduction** = achievement of a 4-point reduction in WI-NRS pruritus score post-baseline compared to baseline, calculated only for subjects with a pruritus score of ≥ 4 at baseline.

- **Itch-free state** = achievement of in WI-NRS of 0 or 1.

- **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.

- **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.
- **Scalpdex 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.

- **Scalpdex Symptoms Scale =** average of (Q1, Q3, Q8) after transforming to 0 to 100 scale as mentioned above. Q refers to question number.

- **Scalpdex 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.

- **Total % BSA (body surface area)** = sum of coverage areas of head, limbs, upper limbs, and lower limbs.

- **Change from baseline** = value at current time point – value at baseline. This is derived separately for primary baseline and OLE baseline.

- **Percent change from baseline** = (value at current time point – value at baseline)/value at baseline. This is derived separately for primary baseline and OLE baseline.

- **TEAE** = any AE that started after application of IP at the Baseline Visit in this extension study, or in parent study through completion.

- **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).

6.1.8. Data Adjustments/Handling/Conventions

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

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

Adverse events will be coded using the MedDRA version 24.1 thesaurus. 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 related, probably related, likely related, 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

The impact of COVID-19 on this clinical trial and participants will be assessed by including in summaries/listings relevant information as captured on the eCRF. Information collected will include the following:

- Any visit disruption caused by COVID-19
- Early termination due to COVID-19
- Weight loss due to complications associated with COVID-19
- Local tolerability assessment and WI-NRS disruption caused by COVD-19
- Study drug interruption (missed dose) due to COVID-19
- Delay in any assessment caused by COVID-19
- Protocol deviation due to COVID-19

7. Study Subjects and Demographics

The disposition, demographics and baseline characteristics, medical/surgical history, prior/concomitant medication, protocol deviations, and exposure will be summarized as indicated below in two separate tables for the associated baseline type defined in section 6.1.1:

Cohort 1 Group 1 (ARQ-154-203 roflumilast foam 0.3%, ARQ-154-203 vehicle, and total for subjects that consented to rollover into this extension study without gap),

Cohort 2 Group 3 (ARQ-154-203 roflumilast foam 0.3%, ARQ-154-203 vehicle, and total for subjects that enrolled into this study with a gap),

Cohort 2 Group 4 (de novo subjects),

Cohort 2 Total, and

Overall (includes all the cohorts and groups listed above).

Due to few number of subjects rolling over from study ARQ-154-116 (Cohort 1 Group 2), information collected from ARQ-154-116 will only be included in the listings.

7.1. Subject Disposition

Disposition summary will include tabulation of the number of subjects who enrolled into each cohort/group, the number of subjects who received treatment, number of subjects who completed the study, number of subjects who prematurely discontinued the study, the reasons for

discontinuation from the study overall and due to COVID-19 disruption, and number of subjects in each analysis population as identified in Section 5.

7.2. Demographics and Baseline Characteristics

Summary statistics for age, age groups (9-11 years old, 12-17 years old, 18-64 years old, and ≥65 years old), gender (including child-bearing potential), ethnicity, race, height, weight, baseline disease characteristics (IGA, Overall Assessment of Erythema, Overall Assessment of Scaling, Scalpdex total score, WI-NRS, percent BSA [total]), and BMI will be presented.

A summary of treatment history, including history of response, intolerance, or contraindication to topical corticosteroids and/or topical antifungals, and subjects having facial involvement and current scalp involvement will be provided.

For the continuous variables, the number of non-missing values and the mean, SD, minimum, Q1, Q3, median, 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 WI-NRS, summary statistics including the mean, median, Q1, Q3, and range of the ordinal variable will be presented as well as frequency counts of each level of the ordinal variable.

These analyses will be conducted for the Safety population. Baseline characteristics will be presented for both primary and OLE baselines.

7.3. Prior and Concomitant Medications

Prior and concomitant medications will be summarized descriptively by Anatomical Therapeutic Chemical (ATC) level 4, and preferred term using counts and percentages for Safety population.

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 IP. 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. This will include only medications that started after the first dose of IP in this extension study (ARQ-154-214).

Medications will be coded using WHODrug Global B3, version September 2021.

7.4. Surgical and Medical History

Surgical and medical history will be coded according to the MedDRA, Version 24.1.

Surgical and medical history will be summarized by system organ class (SOC) and preferred term (PT) using the Safety population. A subject who experienced the same surgical and medical history event multiple times will be counted only once for the corresponding PT. Similarly, if a subject experienced multiple surgical and medical history events within the same SOC, the subject will be counted only once for that SOC. Surgical and medical history events will be sorted alphabetically by SOC and within each SOC, the PT will be presented by descending order.

7.5. Protocol Deviations

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

The number of subjects with important protocol deviations and/or eligibility deviations in this study will be summarized, including protocol deviations that are related to COVID-19 disruptions and COVID-19 infections for the subjects in Safety population by groups and overall.

All of the protocol deviations (including important or non-important protocol devations) will be provided in the data listings.

7.6. Exposure and Compliance

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

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

The number of days on IP will be derived as indicated in Section 6.1.7 and summarized descriptively.

The gap between completing study ARQ-151-203 and enrolling into this study will not be included in the above definition and will be summarized separately for subjects in Cohort 2 Group 3.

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. Doses missed due to IGA clearance were not considered as noncompliance.

The number and percent of subjects that missed more than 3 consecutive doses will be tabulated except for subjects who withheld IP application due to IGA clear per investigator.

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 using the following categories:

> 100%

≥ 80% - ≤100%

< 80%

Drug exposure and compliance will be presented using primary and OLE baselines. The number of subjects with treatment duration (≥1 day, ≥8 weeks, ≥12 weeks, ≥24 weeks, ≥32 weeks, ≥54 weeks, and ≥62 weeks) will also be summarized similarly. Study duration will be summarized from primary and OLE baseline.

Similar exposure and compliance information will be summarized for subjects who completed or discontinued study before protocol amendment 1.

8. Efficacy Analysis

This study is open label without a comparator group; therefore, there will be no formal hypothesis testing performed in the study.

For the continuous variables, the number of nonmissing values and the mean, SD, minimum, Q1, Q3, median, and maximum will be tabulated by study visit; similarly, categorical variables will have the counts and proportions of each value will be tabulated by study visit.

For efficacy summaries 2 baselines will be used (primary and OLE) using Safety population as defined in Section 6.1.1.

All efficacy summaries for primary and OLE baselines will be presented as indicated below:

Cohort 1 Group 1 (ARQ-154-203 roflumilast foam 0.3%, ARQ-154-203 vehicle, and total for subjects that consented to rollover into this extension study without gap),

Cohort 2 Group 3 (ARQ-154-203 roflumilast foam 0.3%, ARQ-154-203 vehicle, and total that enrolled into this study with a gap),

Cohort 2 Group 4 (de novo subjects),

Cohort 2 Total, and

Overall (includes all Cohorts and groups).

For vehicle subjects in Cohort 1 Group 1, Cohort 2 Group 3, de nova subjects in Group 4, primary baseline is the same as OLE baseline.

Due to the few number of subjects rolling over from the study ARQ-154-116, information collected from ARQ-154-116 will be included in the listings only.

Missing data for efficacy variables will not be imputed.

8.1. Primary Efficacy Analysis

Not applicable.

8.2. Secondary Efficacy Analysis

8.2.1. Investigator Global Assessment (IGA)

The IGA is an ordinal scale with five severity grades which is reported only in integers. Table 4 illustrates the description of each severity grade.

Table 4: IGA

Score	Description
0	**Completely clear:** No erythema, no scaling (hypo-hyperpigmentation can be present)
1	**Almost clear:** Residual slight erythema and/or trace amounts of scaling
2	**Mild:** Pink to red color and/or slight scaling
3	**Moderate:** Distinct 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)

IGA grade

IGA grade and change from baseline will be summarized using descriptive statistics at each visit by groups and overall.

The number and percentage of subjects in each IGA grade category will be summarized by cohort, group, and overall.

The number and percentage of subjects who achieve IGA 0 or 1 will be tabulated at each scheduled visit by groups and overall. In addition, the achievement of IGA of 0, the achievement of IGA success (i.e., IGA of 0 or 1 plus a 2-grade improvement from Baseline) also will be tabulated. Furthermore, 95% CIs for proportion of subjects who meet the criteria will be provided along with proportions. IGA success from OLE baseline should exclude subjects who already achieved IGA of 0 or 1 at OLE baseline.

IGA success, IGA of 0 or 1, IGA of 0 at any time post baseline including the assessments at the unscheduled visit will be summarized by groups and overall, along with 95% CIs for proportion of people who meet the criteria.

These analyses are repeated for both primary and OLE baselines. The subjects who achieved IGA of 0 or 1 at baseline will be excluded from the analysis of IGA success from baseline.

Subgroup analyses will be performed for both primary and OLE baseline for the subgroups mentioned in Section 6.1.6.

Duration of IGA success, Duration of IGA of 0 or 1, and Duration of IGA of 0

This analysis will be performed using IGA grade assessed in the open-label study.

Duration of IGA success, for subjects whose IGA grade is not 0 or 1 at OLE baseline, is defined as the time from the first observation of IGA success to first subsequent time a subject's disease response does not meet the criteria for IGA Success. The duration of IGA success for subjects who end treatment in IGA success will be censored at the last disease assessment date. Only the first IGA success duration will be analyzed.

If there is no IGA assessment after the first observation of IGA success, the duration will be censored at the end of study date.

The median time (number of days) for IGA success will be analyzed using the Kaplan–Meier method. Subjects who discontinue study without relapse will be considered censored in the Kaplan–Meier analyses. The Kaplan-Meier estimate will be stratified by the IGA success time (IGA success by Week 4, between Week 4 and Week 12, between Week 12 and Week 24, between Week 24 and Week 36, and between Week 36 and Week 52).

Duration of IGA success will be summarized by Cohort, Groups, and overall.

The analysis of duration of IGA of 0 or 1 will be performed in the similar fashion.

Duration of IGA of 0 will be performed for subjects who achieve IGA of 0 post OLE baseline and stop treatment. Subjects who achieve IGA of 0 but never stop treatment will be excluded from analysis. If subject achieved IGA of 0 but stop treatment several days later, the duration

will start at the IP stop date. The analysis of duration of IGA of 0 will be in the similar fashion as the analysis for duration of IGA success.

Similar analysis from primary baseline will be performed.

Disease recurrence

This analysis will be performed using IGA assessments during the open-label study.

Time to disease recurrence is defined as time from first observation of IGA of 0 and stop treatment to to first observation of IGA of 2 or higher after stop treatment. Subjects who achieve IGA of 0 but never stop treatment will be excluded from analysis. If subject achieve IGA of 0 but stop treatment several days later, the duration will start at the IP stop date. If there is other IGA assessment after achieving IGA of 0, the time to disease recurrence will be censored at the last disease assessment date. If there is no IGA assessment after the first observation of IGA success, the duration will be censored at the end of study date. The analysis of time to disease recurrence will be in the similar fashion as the analysis of durationof IGA success.

Similar analysis from primary baseline will be performed.

Treatment Free Interval

Time to restarting study drug (treatment free interval), defined as time from when subject achieves disease clearance (IGA of 0) and stop treatment to all lesions to time when treatment restarts due to IGA of 2 or above, will be calculated and summarized in the same manner as duration of IGA success. The number (%) of subjects who achieved disease clearance (IGA of 0) and stopped treatment will be presented as well. If the subjects never restart the study treatment after achieving IGA of 0, the subject will be censored at the end of study date.

8.3. Other Secondary Efficacy Analysis

The other secondary efficacy endpoints are summarized using OLE baseline, except for WI-NRS which will be summarized at both primary and OLE baseline. These are summarized as discussed in Section 8 and include the following:

- Change and percent change in Scalpdex total score as observed over time.
- Change and percent change in Overall Assessment of Erythema score
- Change and percent change in Overall Assessment of Scaling score
- Achievement of an Overall Assessment of Erythema (0-3 scale) score of 0 or 1 plus a 2-grade improvement from Baseline as observed over time.
- Achievement of an Overall Assessment of Scaling (0-3 scale) score of 0 or 1 plus a 2-grade improvement from Baseline as observed over time.

- Achievement of an Overall Assessment of Erythema (0-3 scale) score of 0 as observed over time.
- Achievement of an Overall Assessment of Scaling (0-3 scale) score of 0 as observed over time.
- Change and percent change from baseline in % BSA affected by disease as observed over time.
- Change and percent change in WI-NRS as observed over time.
- In subjects with a Baseline WI-NRS pruritus score of ≥4, achievement of a ≥4-point improvement from Baseline in WI-NRS pruritus score as observed over time.
- Achievement of itch-free state which is defined as WI-NRS of 0 or 1

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. Observed, change, and percent change from baseline for Scalpdex scale scores along with total score are summarized descriptively.

Observed, change, and percent change from baseline for WI-NRS and % BSA (total) affected by disease will be summarized descriptively.

In subjects with WI-NRS pruritus score ≥ 4 at baseline (primary and OLE), achievement of a 4-point reduction in WI-NRS pruritus score at each scheduled visit will be summarized. The number and percentage of subjects for each scale score will be summarized for OLE baseline.

The number and percentage of subjects who achieve Overall Assessment of Erythema Success, Overall Assessment of Scaling Success, Erythema score of 0, and Scaling score of 0 will be tabulated at each scheduled visit. In addition, categorical summaries (number and percentage of subjects) in each category will be summarized for Erythema and Scaling.

9. Safety and Tolerability Analysis

Safety will be evaluated from reported AEs, local tolerability assessments, changes in clinical laboratory values, changes in vital signs/weight, C-SSRS, PHQ-8/modified PHQ-A, and CDI-2 results.

9.1. Adverse Events

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

A TEAE is defined as an AE that started after application of IP at the Baseline Visit in this extension study. All TEAEs will be summarized from OLE baseline.

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 SAE (TESAE), TEAE leading to study drug withdrawal, TEAE leading to study withdrawal, or TEAE resulting in death.

The number and percent of subjects reporting TEAEs, grouped by MedDRA system organ class (SOC) and preferred term (PT), will be tabulated by all TEAEs, non-serious TEAEs, maximum severity, or greatest relationship to study IP; Cohort and by groups and overall. In the case of multiple occurrences of the same TEAE within the same subject, each subject will only be counted once for each preferred term.

The number and percent of subjects reporting TEAEs will be tabulated by PT and presented by groups and overall.

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).

All the AEs will be listed. The AEs that are treatment emergent will be flagged. Any AEs related to the application site will be flagged in the listing.

9.1.1. Adverse Events Leading to Withdrawal

A summary of incidence rates (frequencies and percentages) of TEAEs leading to withdrawal of study drug and TEAEs leading to withdrawal of study, by SOC and preferred term will be prepared for the Safety population.

9.1.2. Deaths and Serious Adverse Events

TESAEs will be tabulated by SOC and PT and presented by groups and overall. Fatal AEs will be flagged in the AE listings.

9.2. Local Tolerability Assessments

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

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

Primary baseline will be used for Cohort 1 Group 1 and OLE baseline for Cohort 2. Overall group will be summarized using Safety baseline.

9.3. Clinical Laboratory Evaluations

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

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 (shift table).

Primary baseline will be used for Cohort 1 Group 1 and OLE baseline for Cohort 2. Overall group will be summarized using Safety baseline.

Abnormal laboratory values will be presented in the listings.

9.4. Vital Signs

Descriptive summaries of observed values, changes, and percent 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 group and study visit.

Categorical percent changes in weight (i.e., shifts from baseline) by group will be summarized by 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 groups and study visit.

BMI is derived as specified in Section 6.1.7. Shift tables 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 groups and study visit.

Shift tables for weight and BMI will also by summarized by intentional and unintentional weight loss (defined from question on vital sign eCRF). Vital signs summaries will use primary baseline for Cohort 1 Group 1 and OLE baseline for Cohort 2. Overall group will be summarized using Safety baseline.

9.5. PHQ-8

PHQ-8 will be completed by adult subjects and modified PHQ-A will be completed 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 study visit will be presented by group. As PHQ-A is not assessed in ARQ-154-203 study, only OLE baseline will be used for this summary.

9.6. CDI-2

Parents/caregivers will complete the CDI-2 for subjects belonging to 9-11 years of age inclusive. For CDI-2, descriptive summaries of observed values, changes and percent changes from baseline will be calculated for the total score and the 2 scales emotional problems and functional problems by groups and study visit.

As CDI-2 is not assessed in ARQ-154-203 study, only OLE baseline will be used for this summary.

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 for Cohort 2 only; otherwise, the "Since Last Visit" 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.

Suicidality data collected on the C-SSRS with positive responses will be listed for all subjects if there is any. 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. OLE baseline will be used for this summary.

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 by group for OLE data only.

9.9. Pigmentation Assessment

The hypopigmentation and hyperpigmentation are scored on a 0-3 scale where 0 = none, 1 = mild, 2 = moderate and 3 = severe. A shift table summarizing each category and study visit will be summarized for subjects belonging to Cohort 2 Group 4 for OLE baseline. This summary is repeated for race subgroups as well (white, black, vs others).

10. Changes from Planned Analysis

Per the sponsor request, the following changes have made to the SAP:

- The 2-grade improvement in IGA from Baseline as observed over time endpoint has been removed.
- TEAE by PT, TESAE by SOC, PT table summaries have been added.
- Shift in pigmentation assessment summary for Cohort 2 Group 4 is added.
- Subgroup analysis tables for TEAE by severity have been added.
- The Durations of IGA Success, Duration of IGA 0/1 and other similar duration variables have been updated to be defined from first instance of the event to first subsequent instance of loss of the event, rather than the time from the first instance of the event to the last subsequent instance of the event.

11. Other Planned Analysis

11.1. Pharmacokinetic Analysis

All PK data will be listed and summarized descriptively.

12. References

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/.