STATISTICAL ANALYSIS PLAN

LONGITUDINAL ENDOTYPING OF ATOPIC DERMATITIS THROUGH TRANSCRIPTOMIC SKIN ANALYSIS

LEADS STUDY

Protocol ADRN-12
IND Exemption
National Clinical Trial # NCT05436535

Study Sponsor: The National Institute of Allergy and Infectious Diseases (NIAID)
NIAID Funding Mechanism: Grant 1 UM1 AI151958-01 and Grant 1U01AI178772-01

Prepared By: Seibold Lab
[redacted]

ACKNOWLEDGEMENT AND SIGNATURE

Seibold Lab Approval	**Rho Approval**
[redacted]	[redacted]
Signature and Date	Signature and Date

VERSION HISTORY

Version	**Date**	**Change(s)**	**Author**
1.0	April 30, 2025	Initial Version	[redacted]

Table of Contents

1 LIST OF ABBREVIATIONS

AD	Atopic Dermatitis
ADR	Adverse Drug Reaction
ADRN	Atopic Dermatitis Research Network
AE	Adverse Event
AR	Adverse Reaction
CBC	Complete Blood Count
CFR	Code of Federal Regulations
CFU	Colony Forming Unit
CoNS	Coagulase Negative Staphylococcal Species
CPC	Clinical Product Center
CRA	Clinical Research Associate
CRSwNP	Chronic rhinosinusitis with nasal polyps
CTCAE	Common Terminology Criteria for Adverse Events
DAIT	Division of Allergy, Immunology, and Transplantation
DNA	Deoxyribonucleic Acid
DoR	Delegation of Responsibility Log
DSMB	Data Safety Monitoring Board
EASI	Eczema Area and Severity Index
eCRF	Electronic Case Report Form
EDC	Electronic Data Capture
EO	Exploratory Objective
EoE	Eosinophilic Esophagitis
FDA	Food and Drug Administration
FLG	Filaggrin
GCP	Good Clinical Practice
GLM	Generalized Linear Model
GWAS	Genome-wide Association Study
HIV	Human Immunodeficiency Virus
IGA	Investigator Global Assessment
IgE	Immunoglobulin E
IL	Interleukin
IRB	Institutional Review Board
ISR	Injection Site Reaction

JAK	Janus kinase
mAb	Monoclonal Antibody
MOP	Manual of Procedures
NBUVB	Narrowband Ultraviolet B
NCI	National Cancer Institute
NESS	Nottingham Eczema Severity Score
NIAID	National Institute of Allergy and Infectious Diseases
NIH	National Institutes of Health
OHRP	Office for Human Research Protections
PCR	Polymerase Chain Reaction
PI	[Site] Principal Investigator
PP-NRS	Peak Pruritus Numerical Rating Scale
PUVA	Psoralen and Ultraviolet A
Q2W	Every other week
Q4W	Every 4 weeks
RNA	Ribonucleic Acid
S. aureus	*Staphylococcus aureus*
SACCC	Statistical and Clinical Coordinating Center
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SAR	Suspected Adverse Reaction
SCORAD	Scoring Atopic Dermatitis
SO	Secondary Objective
SOP	Standard Operating Procedure
STS	Skin tape strip
TAA	Target Area Assessment
TCS	Topical Corticosteroids
Th	T Helper
TSB	Tryptic Soy Broth
US	United States
UVA1	Ultraviolet A1
UVB	Ultraviolet B
WGCNA	Weighted Gene Co-expression Network Analysis
WTS	Whole Transcriptome Sequence

2 PURPOSE OF THE ANALYSES

2.1 Primary objective

To determine if the type 2-high non-lesional skin (skin tape) endotype is associated with current mild versus moderate-to-severe AD disease

2.2 Secondary objectives

1. To determine how gene expression in the skin (skin tape) differs between non-AD participants and those with current mild or moderate-to-severe AD disease
2. To determine how gene expression in the skin (skin tape) changes with standard-of-care treatment (two timepoints) among the study outcome groups: (1) topical steroid responders, (2) dupilumab responders, (3) dupilumab non-responders, (4) non-AD, and (5) long-term dupilumab participants

2.3 Exploratory objectives

Not covered in this SAP, please refer to a separate Exploratory SAP.

3 PROTOCOL SUMMARY

This is a multi-center, longitudinal study that will characterize the gene expression profiles and transcriptomic endotypes that underlie mild and moderate-to-severe AD and will determine changes in these expression patterns and endotypes in response to standard-of-care treatment as delineated in the protocol. The primary objective of this study is to determine whether the type 2-high non-lesional skin endotype is associated with current mild versus moderate-to-severe AD disease.

4 ANALYSIS SAMPLES

4.1 Primary Endpoints

Non-lesional skin tape transcriptome at Day 7 (long-term dupilumab participants excluded)

4.2 Secondary Endpoints

1. Non-lesional and lesional skin tape transcriptome at Day 7 and Day 168-224 (long-term dupilumab participants excluded) [Secondary Objective (SO)1]

2. Non-lesional and lesional skin tape transcriptome at Day 7 and Day 168-224 [SO2]

4.3 Exploratory Endpoints

Please refer to the Exploratory SAP.

5 ESTIMANDS FOR PRIMARY AND KEY SECONDARY EFFICACY ENDPOINTS

5.1 Primary Estimand

The primary estimand is the odds ratio describing the difference in frequency of mild (Day 7 EASI ≤ 7) vs. moderate-to-severe (Day 7 EASI > 7) AD between those with and without the type 2 endotype, as defined using gene expression in the non-lesional Day 7 skin tape transcriptome. The analysis will be carried out separately for the children (6-17 years old) and adults (18+ years old) populations in the study. Similarly, we will estimate the odds ratio for the change in the odds of having mild vs. moderate-to-severe AD associated with 1 unit increase in the geometric mean levels of type 2 gene expression from the non-lesional Day 7 skin tape transcriptome (again, separately in children and adults).

5.2 Secondary Estimands

The first of the two secondary estimands is the difference in mean gene expression from non-lesional skin tape transcriptome at Day 7, and separately at Day 168-224, between Non-AD vs. Mild DNAD and Non-AD vs. Moderate-to-severe DNAD children (6-17 years old), and separately adults (18+ year olds), enrolled in this study.

The second of the two secondary estimands is the mean change in normalized gene expression between Day 7 and Day 168-224 non-lesional skin tape transcriptomes, among topical steroid responders, dupilumab responders, dupilumab non-responders, non-AD, and long-term dupilumab responders in children (6-17 years old), and separately in adults (18+ year olds) enrolled in this study.

6 STUDY SUBJECTS

6.1 Study Subject Summary

The study population consists of children and adults with mild and moderate-to-severe AD, as well as children and adults with no AD. The total sample size will be approximately 600 children and adults (200 children and 400 adults). Formal intent-to-treat and per protocol populations do not apply to this study. Because we want a balanced population with regard to age, sex, and race/ethnicity between the three enrollment groups, we will monitor these variables during the enrollment period. Any detected skews accruing during enrollment will be addressed by instructing sites to target unrepresented groups. When assigning participants as child or adult, the age at baseline will be used (6-17 years at baseline = child; ≥ 18 at baseline = adult).

6.2 Disposition of Subjects

Study participants, whether child or adult, will be allocated to the following diagnostic groups prior to analysis:

I. *Non-AD*
Participants never diagnosed with AD.

II. *Dupilumab-naïve AD participants*
AD participants never taking dupilumab.

There are two possible AD severity groups:
1. Mild AD (EASI ≤7 at analysis timepoint)
2. Moderate-to-severe AD (EASI > 7 at analysis timepoint)

There are three possible outcome groups:
1. Topical steroid responders: Those on topical steroid who never obtain an EASI score > 7 by the 12-Week Steroid Assessment visit (EASI score > 7 obtained after that point will lead to early study termination).
2. Dupilumab responders: Steroid non-responders on dupilumab who obtain an EASI score ≤ 7 at the 15-Week Dupilumab Assessment visit (regardless of EASI scores obtained before or after that visit).
3. Dupilumab non-responders: Those on dupilumab who fail to obtain an EASI score ≤ 7 at the 15-Week Dupilumab Assessment visit (regardless of EASI scores obtained before or after that visit).

III. *Long-term dupilumab AD participants*
AD participants already on dupilumab (for ≥ 4 months) at the start of the study. These patients will continue dupilumab treatment as prescribed by their physician outside of the study.

7 STUDY OPERATIONS

7.1 Protocol Deviations

Protocol Deviation – The investigators and site staff will conduct the study in accordance with the protocol; no deviations from the protocol are permitted. Any change, divergence, or departure from the study design or procedures constitutes a protocol deviation. As a result of any deviation, corrective actions will be developed by the site and implemented promptly.

Major Protocol Deviation (Protocol Violation) - A Protocol Violation is a deviation from the IRB approved protocol that may affect the participant's rights, safety, or well-being and/or the completeness, accuracy, and reliability of the study data. In addition, protocol violations include willful or knowing breaches of human participant protection regulations or policies; any action that is inconsistent with the NIH Human Research Protection Program's research, medical, and ethical principles; and any serious or continuing noncompliance with federal, state, local or institutional human participant protection regulations, policies, or procedures. Examples of Major Protocol Deviations are described in the study MOP.

Non-Major Protocol Deviation - A non-major protocol deviation is any change, divergence, or departure from the study design or procedures of a research protocol that does not have a major impact on the participant's rights, safety or wellbeing, or the completeness, accuracy, and reliability of the study data.

7.2 Stratification, Randomization, and Blinding/Masking

Laboratory personnel conducting the assays will be blinded to diagnostic group (non-AD, mild AD, or moderate-to-severe AD) and study therapy regimen, until after samples are processed. Analytical personnel who are unblinded for quality control purposes or to complete interim analyses of the data will not be involved in processing samples.

8 GENERAL ANALYSIS AND REPORTING CONSIDERATIONS

8.1 Analysis Visits

The Day 7 Visit – for all relevant clinical and outcome groups, the Day 7 visit pertains to the 1st Week Clinical Assessment Visit that occurs 7 days (± 3 days) after the Baseline Assessment (Day 0) at study enrollment.

The End of Study Assessment (Day 168-224) Visit – for Non-AD, Topical Steroid Responders, and Long-term Dupilumab participants, their End of Study Assessment Visit occurs 168 days (± 3 days) after the Baseline Assessment (Day 0) at study enrollment. For Topical Steroid Non-Responders, their End of Study Assessment Visit occurs 133 days (130-136 days) after Dupilumab Initiation, which can occur any time between 35 days (-3, +6 days) and 91 days (± 3 days) after the Baseline Assessment (Day 0) at study

enrollment depending on when the participant obtains the qualifying EASI score (> 7) for initiating Dupilumab. Hence, the End of Study Assessment Visit can range from Day 168 to 224.

Unscheduled Visit - participants may be asked to return to the clinic for Unscheduled Visits, as needed, to provide additional skin swabs, STS, and/or skin biopsies for studies, or if samples are lost or destroyed, or if insufficient yields were obtained at a previous study visit. Thus, for the purposes of the analyses described in this SAP, the STS samples obtained at an unscheduled visit that occurred after the Day 7 or End of Study Assessment visits may be used for analysis instead if the samples obtained at the scheduled visits were lost, destroyed, or yielded insufficient gene expression.

8.2 Stratification and covariates

Throughout analysis, our approach will be to analyze pediatric and adult populations separately to better model and discover any cohort-specific endotypes and phenotypes. We will then compare how effects of interest differ between the two age cohorts. We will also fit combined pediatric/adult models using interaction terms with age group to help identify effects that are shared and unique to age groups. In regards to skin sampling, in addition to skin tape strip (STS) samples collected from all children and adults at each sample point, we will be collecting required and optional biopsy samples from adults, depending on time point and course of treatment. We may in certain instances compare biopsy and STS samples, as information from these samples can be complementary. For AD participants, skin samples will be collected from both lesional and non-lesional skin, as applicable. Although main analyses will focus on non-lesional skin, such that signatures of AD disease and severity can specifically be tied to systemic effects on skin tissue rather than to the unique character of lesional skin, we will also investigate how AD and treatment effects differ between the two skin types.

Finally, as transcriptional responses to AD or AD treatments can also be influenced by individual traits, such as age, sex, race/ethnicity, body mass index (BMI), concomitant therapies, concomitant asthma, age of asthma diagnosis, asthma controller usage, presence and frequency of eczema herpeticum, history of Staph infection, presence of recurrent or chronic sinusitis, modified rhinitis symptom utility index (MRSUI) total score, presence and type of food and animal allergies, presence and treatment of viral/bacterial infections, smoking status and pack years of the participant or caregiver, level of compliance for the different treatments, and existence and type of adverse events, we may include these as covariates in the models below where appropriate, to ensure that such variables don't confound our inferences. Where significant, effects of these covariates on responses will be investigated by running sensitivity analyses on datasets stratified based on these covariates.

8.3 Descriptive analyses

Descriptive analyses will be reported separately by baseline AD severity group and study outcome group. Continuous baseline measures will be reported as 1) mean (or geometric mean) with standard deviation, or 2) median with 1st and 3rd quartile, as appropriate. Categorical baseline and demographic characteristics and participant disposition will be reported as frequencies and proportions.

9 ENDPOINT EVALUATION

9.1 Primary Analysis of Primary Endpoint

9.1.1 Determining if the type 2-high non-lesional skin (skin tape) endotype is associated with current mild versus moderate-to-severe AD disease [Primary Objective]

This analysis will utilize WTS data from the Day 7 Visit and will be performed separately for children and adults, but also in a combined analysis (described below). We will first define the type 2-high skin endotype in the transcriptome dataset by identifying a network of co-expressed genes enriched in known markers of type 2-inflammation using weighted gene co-expression analysis (Weighted Gene Co-expression Network Analysis (WGCNA); Langfelder and Horvath 2008), or similar analysis. These networks should capture distinct biological processes taking place in these participants, including type 2-high inflammation. The type 2-high network will be identified as the network with strongest enrichment of published type 2 inflammation gene sets (Dyjack et al. 2018; Leung et al. 2019), based on hypergeometric tests of over-representation. Alternatively, if a type 2 network is not well defined, we may use a known skin type 2 inflammatory gene set.

Specifically, we will test whether expression of the type 2 endotype is associated with AD and AD severity using the following 8 logistic regression models:

Regressor	Population	Response Groups	Model #
Type 2 Inflammation in Non-lesional STS collected at Day 7	Adults (18+ year olds)	Non-AD vs. DNAD (all severity) at Day 7	1
		Non-AD vs Mild DNAD at Day 7	2
		Non-AD vs Moderate-to-Severe DNAD at Day 7	3
		Mild vs. Moderate-to-Severe DNAD at Day 7	4*
	Children (6-17 year olds)	Non-AD vs. DNAD (all severity) at Day 7	5
		Non-AD vs Mild DNAD at Day 7	6
		Non-AD vs Moderate-to-Severe DNAD at Day 7	7
		Mild vs. Moderate-to-Severe DNAD at Day 7	8*

** Model # 4 and 8 will evaluate our primary estimand that will directly answer our primary objective. The remaining models will help support our primary findings.*

The 8 models will be fit twice, once treating type 2 inflammation as a dichotomous variable and the second time treating type 2 inflammation as a continuous variable (see below).

1. *Type 2 inflammation as a dichotomous variable:*
   We will divide individuals into type 2-high and type 2-low expression categories by clustering all participants based on gene-level expression of all genes in the type 2 inflammation gene network. Thus, these dichotomous group assignments will emerge from the transcriptome data based on individuals' relative expression of the network of genes most enriched in known type 2 markers. We will then test whether type 2-high status (high vs. low) is associated with disease group identity (AD vs. non-AD, mild AD vs. non-AD, moderate-to-severe AD vs. non-AD, and mild AD vs. moderate-to-severe AD) using logistic regression, with type 2 high status as the predictor and disease group as

the outcome. We will also include relevant covariates, such as age, sex, race/ethnicity, and concomitant therapies as discussed above (Section 8.2).

2. *Type 2 inflammation as a continuous variable:*
   We will calculate the geometric mean of expression across the type 2 inflammation gene network and then test whether this mean expression is associated with disease group identity using logistic regression, with mean expression (and covariates) as the predictor and disease group as the outcome.

Moreover, in addition to considering the categorical grouping of mild and moderate-to-severe AD disease (which is based on body-wide EASI as defined in section 6.2), we will also consider modeling for the association between continuous EASI scores and type 2 inflammation (both dichotomous and continuous as described above). Finally, in addition to using body-wide EASI to measure AD severity, we will also consider measuring sampling site-specific local AD severity based on local EASI score (sum of severity scores of the four clinical signs: Erythema, Edema, Lichenification, and Excoriation that were assessed at the target area sampling site) and target area assessment (TAA). We will test for association between these continuous local AD severity scores versus dichotomous or continuous type 2 inflammation measures.

As stated in 8.2, in addition to the models above, we will also run an additional 8 models (four participant-comparisons by two type 2 outcome measures) that test for association between type 2 outcome measures and disease group identity in pediatric and adult patients combined. These models will contain interactions terms involving disease status and age group to help identify disease effects that may differ by age group.

The age-stratified models defined above will also be repeated, including additional covariates as described in Section 8.2. If a covariate is found to be statistically significant in the model and if deemed appropriate, we will carry out sensitivity analyses by repeating analysis in participants stratified into covariate-based groups to determine whether effects qualitatively differ among stratified groups. For continuous covariates, this will require discretization. If sufficient samples are available for covariate-based groups, we may also fit a single model that includes the covariate and disease group with covariate interaction terms, enabling us to quantitatively determine whether disease effects differ by the covariate.

9.2 Primary Analysis of Secondary Endpoints/Outcomes

9.2.1 Determining how gene expression in the skin (skin tape) differs between non-AD participants and those with current mild or moderate-to-severe AD disease [SO #1]

This analysis will utilize WTS data from non-lesional and lesional skin tape samples collected at the Day 7 and End of Study Assessment (Day 168-224) Visits. To investigate whether skin gene expression differs due to AD status and severity of disease at each time point, we will carry out single-gene differential expression (DE) analysis using negative-binomial generalized linear models (GLM) to compare expression between

participants in different disease severity groups (non-AD, mild AD, and moderate-to-severe AD), while accounting for covariates. To assess effects of AD status, we will compare AD vs. non-AD, mild AD vs. non-AD, and moderate-to-severe AD vs. non-AD. We will compare mild AD vs. moderate-to-severe AD to explicitly test for differences based on AD severity. For all comparisons above and throughout this SAP, the Benjamini-Hochberg (BH) method will be used to adjust p-values to control the false discovery rate (FDR) at 5%, which we will apply across all tests that fall under a specific hypothesis (e.g., when testing whether gene expression is associated with non-AD vs AD, we will apply an adjustment for all genes tested. A separate adjustment will be applied when testing whether gene expression is associated with mild vs moderate-to-severe AD.

Although the above analyses will focus on non-lesional skin samples, which will be available for all participants, we will also examine differences in disease signatures between lesional and non-lesional skin in participants from whom we have both types of samples (AD participants only) using a linear mixed model. Variance stabilizing transformation (VST)-normalized gene expression (Love et al. 2014) will be used as the outcome and models will include fixed effects for skin type (lesional or non-lesional), disease severity (mild or moderate-to-severe), and skin type x disease severity interaction, along with a random intercept for each participant to account for pairing of lesional and non-lesional samples. The interaction term will allow us to explicitly test for different changes between mild and moderate-to-severe AD in lesional vs. non-lesional skin. Using this model, we will also be able to compare gene expression differences between disease groups within lesional skin. Note that each of the above models will be fit for each visit separately. The above models will also be fit separately for children and adults, after which we will fit a combined model using interaction terms with age group to identify severity changes that are shared and unique to age groups.

In total, there will be 22 distinct models:

Response	Population	Regressors	Model #
Single-gene differential expression from Non-lesional STS collected at Day 7	Adults (18+ year olds)	Non-AD vs. DNAD (all severity) at Day 7	1
		Non-AD vs Mild DNAD at Day 7	2*
		Non-AD vs Moderate-to-Severe DNAD at Day 7	3*
		Mild vs. Moderate-to-Severe DNAD at Day 7	4
	Children (6-17 year olds)	Non-AD vs. DNAD (all severity) at Day 7	5
		Non-AD vs Mild DNAD at Day 7	6*
		Non-AD vs Moderate-to-Severe DNAD at Day 7	7*
		Mild vs. Moderate-to-Severe DNAD at Day 7	8
Single-gene differential expression from Non-lesional STS collected at End of Study (Day 168-224)	Adults (18+ year olds)	Non-AD vs. DNAD (all severity) at End of Study	9
		Non-AD vs Mild DNAD at End of Study	10*
		Non-AD vs Moderate-to-Severe DNAD at End of Study	11*
		Mild vs. Moderate-to-Severe DNAD at End of Study	12
	Children (6-17 year olds)	Non-AD vs. DNAD (all severity) at End of Study	13
		Non-AD vs Mild DNAD at End of Study	14*
		Non-AD vs Moderate-to-Severe DNAD at End of Study	15*
		Mild vs. Moderate-to-Severe DNAD at End of Study	16

Single-gene differential expression from matching set of Non-lesional AND Lesional STS collected at Day 7	Adults (18+ year olds)	Mild vs. Moderate-to-Severe DNAD at Day 7, Non-lesional vs Lesional, Interaction	17
	Children (6-17 year olds)	Mild vs. Moderate-to-Severe DNAD at Day 7, Non-lesional vs Lesional, Interaction	18
	Adults AND Children	Mild vs. Moderate-to-Severe DNAD at Day 7, Non-lesional vs Lesional, Adults vs. Children, Interactions	19
Single-gene differential expression from matching set of Non-lesional AND Lesional STS collected at End of Study (Day 168-224)	Adults (18+ year olds)	Mild vs. Moderate-to-Severe DNAD at End of Study, Non-lesional vs Lesional, Interaction	20
	Children (6-17 year olds)	Mild vs. Moderate-to-Severe DNAD at End of Study, Non-lesional vs Lesional, Interaction	21
	Adults AND Children	Mild vs. Moderate-to-Severe DNAD at End of Study, Non-lesional vs Lesional, Adults vs. Children, Interactions	22

** Models 2,3,6,7,10,11,14, and 15 will evaluate the first of our secondary estimands that will directly answer the first of our secondary objectives. The remaining models serve as supportive analyses.*

The 22 models defined above will then be repeated with additional covariates as described in Section 8.2. If a covariate is found to be statistically significant in the model and if deemed appropriate, we will carry out sensitivity analyses by repeating analysis in participants stratified into covariate-based groups to determine whether effects qualitatively differ among stratified groups. For continuous covariates, this will require discretization. If sufficient samples are available for covariate-based groups, we may also fit a single model that includes the covariate and disease group x covariate interaction terms, enabling us to quantitatively determine whether disease effects differ by the covariate.

Pathway and enrichment analyses will be performed on differentially expressed gene (DEG) sets to determine the biological mechanisms that underlie transcriptional differences.

9.2.2 Determining how gene expression in the skin (skin tape) changes with standard-of-care treatment (two timepoints) among the study outcome groups: (1) topical steroid responders, (2) dupilumab responders, (3) dupilumab non-responders, (4) non-AD, and (5) long-term dupilumab responders [SO #2]

To investigate changes in skin gene expression due to standard-of-care treatment as delineated in the protocol and how these changes may differ with study outcome groups (i.e., groups based on treatment regimen and response), we will model gene expression at both Day 7 and End of Study (Day 168-224) Visits using linear mixed models. Gene expression data will be normalized using the VST method. Models will include fixed effect variables for visit, study outcome group (topical steroid responders, dupilumab-naïve responders, dupilumab-naïve non-responders, non-AD, and long-term dupilumab participants), and visit x

study outcome group interaction, along with relevant covariates and a random intercept for each participant to account for pairing of Day 7 and End of Study samples. Among the covariates, there will be a binary variable that further distinguishes topical steroid responders who had mild AD to begin with since baseline (set covariate = 1) versus those who actually improved from having moderate-to-severe AD at baseline (set covariate = 0) in order to consider them as "true" topical steroid responders to compare with among the other outcome groups in the model (all remaining groups would set covariate = 0 as well). Using this model, we will examine responses to standard-of-care treatment as delineated in the protocol by contrasting Day 7 and End of Study Assessment Visits within outcome groups. We will also examine differences in response to standard-of-care treatment between any two outcome groups by testing the interaction term for visit effect in one outcome group versus another.

As noted in Section 9.2.1, these analyses will focus on non-lesional skin samples. However, for each outcome group separately, we will also examine differences in response to standard-of-care treatment between lesional and non-lesional skin in participants from whom we have both types of samples (AD participants only) by fitting mixed models based on VST-normalized data that include a random intercept for each participant and fixed effect terms for skin type (lesional or non-lesional), visit, and skin type x visit interaction. Using these models, we will also be able to investigate gene expression responses to standard-of-care treatment for each outcome group within lesional skin. Note that for lesional skin, we will compare target skin sampled at End of Study that corresponds to the original lesional skin sampled in the target area at Day 7, even if the lesion has resolved over the course of treatment.

The above tests will be performed separately for children and adults. We will also fit a combined model, using interaction terms with age group, to identify changes with outcome group that are shared and unique to age groups.

In total, there will be 34 models:

Response	**Population**	**Regression Analysis Group(s)**	**Group Comparisons**	**Model #**
VST-normalized gene expression from Non-lesional STS collected at BOTH Day 7 AND End of Study	Adults	Topical Steroid Responders, Dupilumab Responders, Dupilumab Non-Responders, Non-AD, and Long-Term Dupilumab Responders	For the 5 analysis groups, Day 7 vs. End of Study Indicator, Interactions	1*
		1. Non-AD vs. Topical Steroid Responders 2. Non-AD vs. Dupilumab Responders 3. Non-AD vs. Dupilumab Non-Responders 4. Non-AD vs. Long-Term Dupilumab Responders 5. Topical Steroid Responders vs. Dupilumab Responders 6. Topical Steroid Responders vs. Dupilumab Non-Responders 7. Topical Steroid Responders vs.	Between the select 2 groups, Day 7 vs. End of Study Indicator, Interaction	2-11

		Long-Term Dupilumab Responders 8. Dupilumab Responders vs. Dupilumab Non-Responders 9. Dupilumab Responders vs. Long-Term Dupilumab Responders 10. Dupilumab Non-Responders vs. Long-Term Dupilumab Responders		
	Children	Topical Steroid Responders, Dupilumab Responders, Dupilumab Non-Responders, Non-AD, and Long-term Dupilumab Responders	For the 5 outcome groups, Day 7 vs. End of Study Indicator, Interactions	12*
	Children	1. Non-AD vs. Topical Steroid Responders 2. Non-AD vs. Dupilumab Responders 3. Non-AD vs. Dupilumab Non-Responders 4. Non-AD vs. Long-Term Dupilumab Responders 5. Topical Steroid Responders vs. Dupilumab Responders 6. Topical Steroid Responders vs. Dupilumab Non-Responders 7. Topical Steroid Responders vs. Long-Term Dupilumab Responders 8. Dupilumab Responders vs. Dupilumab Non-Responders 9. Dupilumab Responders vs. Long-Term Dupilumab Responders 10. Dupilumab Non-Responders vs. Long-Term Dupilumab Responders	Between the select 2 outcome groups, Day 7 vs. End of Study Indicator, Interaction	13-22
VST-normalized gene expression from matching set of Non-lesional AND Lesional STS collected at BOTH Day 7 AND End of Study (for the same target area on the lesional skin)	Adults	Topical Steroid Responders	Non-lesional vs Lesional Indicator, Day 7 vs. End of Study Indicator, Interaction	23
		Dupilumab Responders	Non-lesional vs Lesional Indicator, Day 7 vs. End of Study Indicator, Interaction	24
		Dupilumab Non-Responders	Non-lesional vs Lesional Indicator, Day 7 vs. End of Study Indicator, Interaction	25
		Long-term Dupilumab Responders	Non-lesional vs Lesional Indicator, Day 7 vs. End of Study Indicator, Interaction	26
	Children	Topical Steroid Responders	Non-lesional vs Lesional Indicator, Day 7 vs. End of Study Indicator, Interaction	27
		Dupilumab Responders	Non-lesional vs Lesional Indicator, Day 7 vs. End of Study Indicator, Interaction	28
		Dupilumab Non-Responders	Non-lesional vs Lesional Indicator, Day 7 vs. End of Study Indicator,	29

			Interaction	
		Long-term Dupilumab Responders	Non-lesional vs Lesional Indicator, Day 7 vs. End of Study Indicator, Interaction	30
	Adults AND Children	Topical Steroid Responders	Non-lesional vs Lesional Indicator, Day 7 vs. End of Study Indicator, Adults vs. Children Indicator, Interactions	31
		Dupilumab Responders	Non-lesional vs Lesional Indicator, Day 7 vs. End of Study Indicator, Adults vs. Children Indicator, Interactions	32
		Dupilumab Non-Responders	Non-lesional vs Lesional Indicator, Day 7 vs. End of Study Indicator, Adults vs. Children Indicator, Interactions	33
		Long-term Dupilumab Responders	Non-lesional vs Lesional Indicator, Day 7 vs. End of Study Indicator, Adults vs. Children Indicator, Interactions	34

** Models 1 and 12 will evaluate the second of our secondary estimands that will directly answer the second of our secondary objectives. The remaining models serve as supportive analyses.*

The 34 models defined above will then be repeated with additional covariates as described in Section 8.2. If a covariate is found to be statistically significant in the model and if deemed appropriate, we will carry out sensitivity analyses by repeating analysis in participants stratified into covariate-based groups to determine whether effects qualitatively differ among stratified groups. For continuous covariates, this will require discretization. If sufficient samples are available for covariate-based groups, we may also fit a single model that includes the covariate and disease group with covariate interaction terms, enabling us to quantitatively determine whether disease effects differ by the covariate.

Pathway and enrichment analyses will be performed on DEG sets.

9.3 Analysis of Exploratory Endpoints/Outcomes

Please refer to the Exploratory SAP.

10 SAFETY EVALUATION

Safety data are collected, graded, recorded, and reported according to section 12 of the study protocol. The DAIT/NIAID medical monitor receives, reviews, and makes decisions on monthly reports from the Rho statistical team compiling new and accumulating information on AEs, SAEs, and pregnancies recorded by the study site(s) on appropriate eCRFs. The NIAD Allergy and Asthma DSMB annually reviews safety data

that include, at a minimum, a listing of all reported AEs and SAEs. No further safety evaluation has been planned for the interim or final analyses.

11 PHARMACOKINETIC EVALUATION

None

12 OTHER ANALYSES

None

13 INTERIM ANALYSES AND DATA MONITORING

13.1 Quality Control of RNA-seq Data

Beginning with the first batch of RNA-seq data, we will generate metrics of sample quality, including percent of transcripts mapped (to the genome, to genes, to the mitochondrial genome, etc.), and number of expressed read counts and detected genes per sample. Samples with poor or outlier values for these metrics will be removed or re-sequenced. For AD and non-AD groups, we will examine the most highly expressed genes, most variant genes, enriched pathways thereof (using Enrichr), and identified co-expression networks (using WGCNA) to determine whether expected genes and pathways are expressed within the skin tape and skin biopsy data. Furthermore, once we have at least 10 samples per comparison group, and with every new batch of sequence data thereafter, we will use differential expression analysis to test for gene expression differences based on disease status (AD vs. Non-AD) and severity (AD mild vs. AD moderate-to-severe), to ensure the expected gene expression patterns are observed.

13.2 [redacted]

[redacted]

13.2 Interim Analysis of Safety Data

No formal interim analyses are planned for safety.

13.3 Futility Analysis

No formal interim analyses are planned for futility.

14 CHANGES TO THE ANALYSES PLANNED IN THE PROTOCOL

If WGCNA does not capture well type 2 inflammatory gene sets, we may instead use pre-existing type 2 gene sets to define the type 2-high endotype used in carrying out the primary objective. If local EASI and TAA do not well capture local AD severity, we may either modify these measures or use different local severity measures that better capture variation in local AD severity in our population.

15 REFERENCES

Dyjack N., Goleva E., Rios C., Kim B.E., Bin L.H., Taylor P., Bronchick C., Hall C.F., Richers B.N., Seibold M.A., Leung D.Y.M. 2018. Minimally invasive skin tape strip RNA sequencing identifies novel characteristics of the type 2-high atopic dermatitis disease endotype. Journal of Allergy and Clinical Immunology. 141:1298-1309.

Langfelder P., Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 9:559.

Leung D.Y., Calatroni A., Zaramela L.S., LeBeau P.K., Dyjack N., Brar K., David G., Johnson K., Leung S., Ramirez-Gama M. 2019. The nonlesional skin surface distinguishes atopic dermatitis with food allergy as a unique endotype. Science Translational Medicine. 11:eaav2685.

Love M.I., Huber W., Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome. Biol. 15.