16.1.9 Statistical analysis plan

Table of Contents

NCT04718870

STATISTICAL ANALYSIS PLAN

OPEN-LABEL EXPLORATORY STUDY TO EVALUATE THE EFFECT OF DUPILUMAB ON SKIN BARRIER FUNCTION IN PEDIATRIC PATIENTS WITH MODERATE-TO-SEVERE ATOPIC DERMATITIS

STUDY NUMBER: LPS16764

STUDY NAME: PELISTAD

COMPOUND: SAR231893/REGN668

SANOFI STATISTICIAN: [redacted]

SERVICE PROVIDER STATISTICIAN: [redacted])

DATE OF ISSUE: 13-Dec-2022

VERSION: 1

Total number of pages: 38

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

AD:	atopic dermatitis
AE:	adverse event
AESI:	adverse event of special interest
AUC:	area under the curve
BMI:	body mass index
CDLQI:	children's dermatology life quality index
CI:	confidence interval
ClinRO:	clinical-reported outcome
DBP:	diastolic blood pressure
EASI:	eczema area and severity index
EoS:	end of study
EoT:	end of treatment
FLG:	filaggrin
FTIR:	fourier-transform infrared
HLGT:	high level group term
HLT:	high level term
ICH:	international conference on harmonization
IMP:	investigational medicinal product (synonymous with study drug)
ISS:	individual signs score
ITT:	intention-to-treat
LOCF:	last observation carried forward
MedDRA:	Medical Dictionary for Regulatory Activities
mITT:	modified intention-to-treat
NRS:	numerical rating scale
OCT:	optical coherence tomography
POEM:	patient oriented eczema measure
PRO:	patient reported outcome
PT:	preferred term
SAP:	statistical analysis plan
SAS:	statistical analysis system®
SBP:	systolic blood pressure
SC:	subcutaneous
SD:	standard deviation
SEM:	standard error of the mean
SOC:	system organ class
STS:	skin tape stripping
TEAE:	treatment-emergent adverse event
TEWL:	transepidermal water loss
V:	visit
W:	week
WHO-DD:	world health organisation drug dictionary

1 OVERVIEW AND STUDY PLAN

This statistical analysis plan (SAP) provides a comprehensive and detailed description of strategy and statistical technique to be used to realize the analysis of data for study LPS16764 (short title: Dupilumab-pediatric skin barrier function and lipidomics study in patients with atopic dermatitis – PELISTAD). It is applicable for the analysis for the treatment phase and the analysis for the follow-up period.

The analyses described are based upon the final clinical study protocol V1.0 dated 13 July 2020 and will be prepared in accordance with the international conference on harmonization (ICH) E9 Step 4 version dated 05 February 1998 (1).

Objectives, endpoints and statistical analyses related to biomarkers (ie, lipidomics, Filaggrin (FLG) genotyping, proteomics, and transcriptomics) will be detailed in a specific dedicated SAP.

The statistical analyses (except biomarkers analyses) will be performed by the Biostatistics department of BIOTRIAL BIOMETRICS in agreement with the Sponsor.

The clinical SAP will be validated and signed before the study database for the treatment phase analysis is locked.

1.1 STUDY DESIGN

Study design

This is a Phase 4 open-label, exploratory study in three sites evaluating the effect of dupilumab on skin barrier function (as measured by transepidermal water loss (TEWL) before and after skin tape stripping [STS]) in pediatric patients aged between ≥6 and <12 years old with moderate-to-severe atopic dermatitis (AD) with a healthy volunteer cohort as a reference comparator. Patients with moderate-to-severe AD will be included in the US site and patients with severe AD will be included in the EU site (UK). One more US site was initiated because the recruitment at the first US site was insufficient.

Sample size

Approximately 24 pediatric patients with moderate-to-severe AD will be enrolled. An approximately equal number of age, gender, location of targeted lesion area and site matched healthy volunteers serving as a reference comparator cohort will also be enrolled.

Randomization

Not applicable.

Treatment received

Study participants will be assigned to study intervention as described below:

- Up to 24 pediatric patients with moderate-to-severe AD will receive dupilumab.
- Dose regimen 1: children with baseline 15 kg ≤body weight <30 kg will receive a subcutaneous (SC) loading dose of dupilumab 600 mg (2 injections of dupilumab 300 mg) on Day 1 (Week 0), followed by every 4-week SC dosing of dupilumab 300 mg from Week 4 to Week 12.
- Dose regimen 2: children with baseline 30 kg ≤body weight <60 kg will receive a SC loading dose of dupilumab 400 mg (2 injections of dupilumab 200 mg) on Day 1 (Week 0), followed by bi-weekly SC dosing of dupilumab 200 mg from Week 2 to Week 14.

The treatment will be given to the pediatric patient (≥6 and <12 years of age) depending on its body weight.

Duration of study period

This is a 32-week study with a 4-week screening period, 16-week treatment phase designed to investigate dupilumab's effect on skin barrier function in approximately 20 pediatric patients with moderate-to-severe AD, and a 12-week follow-up period.

The graphical study design is detailed in Section 8 (Appendix A).

1.2 OBJECTIVES

1.2.1 Primary objective

The primary objective of this study is to evaluate changes in skin barrier function with TEWL assessed after STS in predefined lesional skin in pediatric patients with moderate-to-severe AD treated with dupilumab.

1.2.2 Secondary objectives

The secondary objectives are:

1. To evaluate changes in skin barrier function with TEWL assessed after STS in predefined lesional and non-lesional skin in pediatric patients with moderate-to-severe AD treated with dupilumab in reference to normal skin of healthy volunteers.
2. To evaluate time course of change in skin barrier function with TEWL assessed before and after STS in predefined lesional and non-lesional skin in pediatric patients with moderate-to-severe AD during dupilumab treatment phase and follow-up period in reference to normal skin of healthy volunteers.

1.2.3 Tertiary/exploratory objectives

The clinical tertiary/exploratory objectives are:

1. To evaluate dupilumab treatment effect on epidermal hypertrophy and structure of skin barrier changes measured by optical coherence tomography (OCT) and fourier-transform infrared (FTIR) spectroscopy in both predefined lesional and non-lesional skin in pediatric patients with moderate-to-severe AD during dupilumab treatment phase and follow-up period in reference to normal skin of healthy volunteers.
2. To explore the association of skin barrier function measured by TEWL with disease impact assessed by standard AD severity assessments (Eczema area and severity index [EASI], and individual signs score [ISS]), patient reported outcomes (PRO) (patient oriented eczema measure [POEM], children dermatology life questionnaire index [CDLQI], worst itch numerical rating scale [NRS], sleep disturbance NRS, skin pain NRS), standardized photos, and the skin barrier hypertrophy and structure profiles assessed by OCT and FTIR.

1.3 DETERMINATION OF SAMPLE SIZE

No formal sample size calculation was performed for this exploratory study. It was based on medical/clinical judgment and is consistent with the sample size from similar studies in the literature (2).

Allowing for drop-out rate of 15%, a total of approximately 24 pediatric patients with moderate-to-severe AD will be enrolled to achieve 20 evaluable patients. An approximately equal number of age, gender, location of targeted lesion area and site matched healthy volunteers serving as a reference comparator cohort will be enrolled. Any drop-out healthy volunteer for whom the matched patient is considered as evaluable will not be replaced.

1.4 STUDY PLAN

The study will comprise of:

- Screening period: Up to 4 weeks (Day -28 to Day -1) from signing the informed consent.
- Baseline visit (Week 0, Day 1): Participants who remain eligible will be enrolled.
- A 16-week open-label treatment phase for AD patients and a 16-week observation period for healthy volunteers.
- Follow-up period for 12 weeks after the end of treatment (EoT) visit (Week 16). Dupilumab may be used as a rescue therapy for AD patients, if needed during follow-up period.

Schedule of activities in pediatric patients and healthy volunteers are detailed in Section 8 (Appendix B and Appendix C).

1.5 MODIFICATIONS FROM THE PROTOCOL

The following changes are added compared to the planned analysis detailed in the study protocol:

- Overall addition: additional information to display on standard TFLs due to COVID19 flags.
- Addition of listing of participants and visits with reasons of exclusion for mITT population (see Section 4.3.1.3).
- Removing of the listing of other baseline characteristics (see Section 4.3.2.1).
- Due to strong variability between the 3 adjacent spots in a similar study conducted by the Sponsor, the mean of the 3 adjacent spots will not be appropriated. The analysis will be performed for each spot separately at a visit before/without STS measurement (see Section 4.3.2.2 and Section 4.3.3.2).
- Update in the Baseline definition for efficacy: the value at Day 1 will be considered as baseline value (or value at Screening if the one at Day 1 is not available). For participants with a value at Day 1 collected after and close to the investigational medicinal product (IMP), the value at this visit could be considered as baseline (judged by clinicians without impact on statistical analysis) (see Section 4.3.2.2).
- If a change in spot anatomical location changed versus Day 1 is identified and assessed by the study team as having a significant impact on skin barrier, only visits prior to the location change will be considered in the statistical analysis of TEWL data (see Section 4.3.3.1 and Section 4.3.3.2).
- Use of a repeated measures mixed model to analyze primary and main secondary endpoints (see Section 4.3.3.1).
- Addition of descriptive statistics on TEWL conditions (see Section 4.3.3.2).
- Except for AUC derivation, TEWL data performed at unplanned number of STS will not be used (see Section 4.3.3.2).
- Tertiary/exploratory endpoints will be analyzed also on the ITT population (see Section 4.3.3.2).
- The following analyses will be added on ITT population only (see Section 4.3.4):
  - The comparison in TEWL between Day 1 and Week 16 will be performed by skin group (lesional and non-lesional skin area only), STS and spot using a non-parametric Wilcoxon signed-rank test.
  - The comparison in TEWL AUC between Day 1 and Week 16 will be performed by skin group (lesional and non-lesional skin area only) and spot using a non-parametric Wilcoxon signed-rank test.
  - The comparison AD participants (lesional skin area/non-lesional skin area) versus Healthy participants will be performed on TEWL data at Week 16 using a non-parametric Wilcoxon signed-rank test.

- The comparison AD participants (lesional skin area/non-lesional skin area) versus Healthy participants will be performed on TEWL AUC data at Week 16 using a non-parametric Wilcoxon signed-rank test.

- Additional statistical analyses could be envisaged as the correlation analysis between the primary endpoint and quantitative global scores of standard AD severity assessments and PRO and the investigation of centre effect due to differences in clinical conditions observed in the similar study conducted by the Sponsor (see Section 4.3.4).

1.6 MODIFICATIONS FROM THE APPROVED STATISTICAL ANALYSIS PLAN

Not applicable.

2 COLLECTED DATA

Data collected on AD patients and healthy volunteers at screening and baseline are:

- Informed consent/assent form.
- Inclusion and exclusion criteria.
- Demographic data (age, sex, race and ethnicity).
- Medical/surgical history.
- Prior medications/procedures.
- Physical measurements (height and body weight).
- Urine pregnancy test for females with childbearing potential only.
- AD disease characteristics are assessed by clinical-reported outcome assessments (ClinROs) (EASI and ISS) and PROs via eDiary (POEM, CDLQI and Worst itch/Sleep disturbance/Skin pain NRS).
- Skin barrier function and structure evaluated by TEWL, STS, OCT, FTIR, and 2D and 3D photographs.
- Safety will be assessed by adverse events, vital signs (including heart rate, systolic and diastolic blood pressure, pulse rate, body temperature, and respiratory rate), physical examination.
- For AD patients only: attribution of the dose regimen for study treatment administration according to approved label.
- Standardized photographs of (predefined lesional/non-lesional or healthy) skin area used for TEWL and standardized full body photographs.

Data collected on AD patients and healthy volunteers after baseline (treatment phase (if applicable) and follow-up) are:

- Urine pregnancy test for females with childbearing potential only.
- Concomitant medications/procedures.
- Safety will be assessed by adverse events, vital signs (including heart rate, systolic and diastolic blood pressure, pulse rate, body temperature, and respiratory rate) and physical examination.
- AD disease characteristics are assessed by ClinROs (EASI and ISS) and PROs via eDiary (POEM, CDLQI and Worst itch/Sleep disturbance/Skin pain NRS).
- Skin barrier function and structure evaluated by TEWL, STS, OCT, FTIR, and 2D and 3D photographs.
- For AD patients only: study treatment administration according to approved label.
- Standardized photographs of (predefined lesional/non-lesional or healthy) skin area used for TEWL and standardized full body photographs.

3 GENERAL STATISTICAL APPROACH

Two final analyses will be performed:

- One for the treatment phase: based on all patients and performed on data collected until the EoT visit (Week 16). This will be done once the last evaluable participant has completed his/her EoT visit.
- One for the follow-up period: based on all participants and all data.

The quantitative variables will be summarized using the following parameters:

- Number of non-missing data.
- Mean.
- Standard deviation (SD).
- Standard Error of the Mean (SEM).
- Median.
- Minimum.
- Maximum.

The qualitative variables will be summarized using the following parameters:

- Number of non-missing data,
- Counts and percentages.

Missing data or unknown responses will not be counted in the percentages.

The following definitions will be used in this document:

- Cohort: the study includes 2 cohorts: patients with moderate-to-severe AD cohort and a healthy volunteers cohort.
- Group: 3 groups are identified: patients' lesional skin area group, patients' non-lesional-skin area group, and healthy volunteers' normal skin group.

4 ANALYSIS OF PATIENT DATA

4.1 ANALYSIS VARIABLES

4.1.1 Baseline variables

Baseline characteristics (demographics, medical history, prior medications, vital signs, physical examination and disease characteristics) are collected.

No electrocardiogram and clinical safety laboratory assessments are required in this study.

4.1.2 Main evaluation variable(s)

The primary endpoint is the percent change from baseline in TEWL after 5 STS assessed on lesional skin at Week 16 in AD patients.

4.1.3 Secondary evaluation variable(s)

The secondary endpoints are:

- Change (percent and absolute) from baseline in TEWL after 5, 10, 15, 20 STS assessed on lesional skin at Week 16 in AD patients.
- Change (percent and absolute) from baseline in TEWL after 5, 10, 15, 20 STS assessed on non-lesional skin at Week 16 in AD patients.
- Change (percent and absolute) from baseline in TEWL after 5, 10, 15, 20 STS assessed on normal skin at Week 16 in healthy volunteers.
- Change (percent and absolute) from baseline in TEWL before STS on lesional skin in AD patients over time.
- Change (percent and absolute) from baseline in TEWL before STS on non-lesional skin in AD patients over time.
- Change (percent and absolute) from baseline in TEWL before STS on normal skin in healthy volunteers over time.
- Change (percent and absolute) from baseline in TEWL area under the curve (TEWL AUC: a composite measure before and after 5, 10, 15, and 20 STS) for skin barrier function in lesional skin in AD patients over time.
- Change (percent and absolute) from baseline in TEWL AUC (a composite measure before and after 5, 10, 15, and 20 STS) for skin barrier function in non-lesional skin in AD patients over time.
- Change (percent and absolute) from baseline in TEWL AUC (a composite measure before and after 5, 10, 15, and 20 STS) for skin barrier function in normal skin in healthy volunteers over time.

- Change (percent and absolute) from baseline in TEWL after STS assessed on lesional skin in AD patients over time.
- Change (percent and absolute) from baseline in TEWL after STS assessed on non-lesional skin in AD patients over time.
- Change (percent and absolute) from baseline in TEWL after STS assessed on normal skin in healthy volunteers over time.

4.1.4 Tertiary/exploratory evaluation variable(s)

The tertiary/exploratory endpoints are:

- Change (percent and absolute) from baseline in epidermal hypertrophy parameters including epidermal thickness, superficial plexus depth, blood vessel diameter and density measured by OCT over time.
- Change (percent and absolute) from baseline in stratum corneum lipid structure measured by FTIR spectroscopy in conjunction with tape-stripping over time.
- Change (percent and absolute) from baseline in EASI (total score) over time.
  - The EASI is a composite index with scores ranging from 0 to 72. Four AD disease characteristics (erythema, thickness [induration, papulation, and edema], scratching [excoriation], and lichenification) will be assessed for severity by the Investigator or designee on a intensity scale of "0" (absent) through "3" (severe). In addition, the area of AD involvement will be assessed as a percentage by body area of head and neck, trunk, upper limbs, and lower limbs, and converted to a area score of 0 to 6. In each body region, the area score is expressed as 0, 1 (1% to 9%), 2 (10% to 29%), 3 (30% to 49%), 4 (50% to 69%), 5 (70% to 89%), or 6 (90% to 100%).
  - Even if the EASI scores are available in the SDTM database, the intermediate and total EASI scores will be derived as follows. However, only the total EASI score will be analyzed.

    i. Severity scores: for each region (Head and neck, Trunk, Upper limbs and Lower limbs), the severity score will be derived as follows, ranging from 0 to 12:

    $$\text{Severity score}_{R} = \text{Intensity score}_{\text{Erythema, R}} + \text{Intensity score}_{\text{Edema/Papulation, R}} + \text{Intensity score}_{\text{Excoriation, R}} + \text{Intensity score}_{\text{Lichenification, R}},$$

    where R = region.

    ii. Region scores: for each region, the region score will be derived as follows:

    $$\text{Region score}_{R} = \text{Severity score}_{R} * \text{Area score}_{R} * \text{Multiplier}$$

    where R = region, Multiplier = 0.1 (or 0.2 for children <8 years) for Head and Neck region, 0.3 for Trunk, 0.2 for Upper Limbs, 0.4 (or 0.3 for children <8 years) with the age calculated at each visit.

    iii. Total EASI score will be derived as the sum of the 4 region scores, ranging from 0 to 72.

- Change (percent and absolute) from baseline in ISS (total score) over time:
  - Total ISS score will be derived as the sum of the four target lesions (Erythema, Infiltration/Papulation, Excoriations and Lichenification), ranging from 0 to 12.
- Change (percent and absolute) from baseline in POEM (total score) over time.
- Change (percent and absolute) from baseline in CDLQI (total score) over time.
- Change (percent and absolute) from baseline in assessment of worst itch NRS (scale) over time.
- Change (percent and absolute) from baseline in sleep disturbance NRS (scale) over time.
- Change (percent and absolute) from baseline in skin pain NRS (scale) over time.
- Change (percent and absolute) from baseline in photograph outputs (TiVi and skin roughness indexes) obtained from skin imaging over time.
- Correlation between baseline values of TEWL before and after STS, and TEWL AUC in lesional and non-lesional skin of AD patients with the following baseline measures:
  - EASI.
  - Targeted lesion ISS.
  - Image-derived severity score in targeted lesional skin.
  - Quantified epidermal hypertrophy and structure changes in OCT and FTIR.
- Correlation between percent change from baseline in TEWL before and after STS, and TEWL AUC in lesional and non-lesional skin of AD patients at Week 8, Week 16 and Week 28 with corresponding change from baseline in the following measures:
  - EASI.
  - Targeted lesion ISS.
  - Image-derived severity score in targeted lesional skin.
  - Quantified epidermal hypertrophy and structure changes in OCT and FTIR.

4.2 ANALYSIS POPULATIONS

The following analysis populations will be defined:

4.2.1 Enrolled population

The enrolled population will be defined by all participants who sign the informed consent form.

4.2.2 Intent-to-treat population

The intent-to-treat (ITT) population will include all enrolled patients, who received at least 1 dose of investigational medicinal product (IMP) and all enrolled healthy volunteers who have at least 1 TEWL/STS assessment performed, irrespective of compliance with the study protocol and procedures.

The ITT population will be used for efficacy analyses.

4.2.3 Modified Intent-to-treat population

The modified intent-to-treat population (mITT) will include all ITT participants. If prohibited therapies for atopic dermatitis (AD) are used and assessed by the study team as having a significant impact on skin barrier only visits prior to rescue treatment use are considered.

The mITT population will be used for efficacy analyses.

4.2.4 Safety population

The safety population will include all patients, who actually received at least 1 dose of IMP or had at least 1 TEWL/STS assessment and all healthy volunteers who have at least 1 TEWL/STS assessment performed.

4.3 STATISTICAL METHODS

All analysis variables will be described according to the methods defined in Section 3 for the description of continuous and categorical variables.

The reporting of the efficacy and safety statistical analyses will be programmed according to the recommendations and templates of the BTD-009094 “Summarizing and reporting standard Early Development clinical trial data (except oncology)” Version 4.0 dated 02-Feb-2021.

4.3.1 Disposition of patients

4.3.1.1 Participant disposition

A detailed description of participant accountability including count of participants by analysis populations, screen failure participants with reasons for screen failure, participants who did not complete the study observation period along with the main reason for study discontinuation (including the relationship to COVID19 for the reasons “Other” if any), and participants who requested permanent treatment discontinuation with the main reason for permanent treatment discontinuation (including the relationship to COVID19 for reasons “Adverse event”, “Withdrawal by subject” and “Other” if any), will be generated by cohorts.

Participant disposition by centre/site will also be performed.

All withdrawals from the study, taking place on or after study intervention intake, will be listed (including the column “Related to COVID19” if applicable).

A listing of patients with treatment discontinuation will be provided (including the column “Related to COVID19” if applicable).

4.3.1.2 Protocol deviations

Individual deviations to inclusion and exclusion criteria as reported by the Investigator will be listed. Major and critical deviations, other than those involving inclusion/exclusion, will be listed by participant (including the column "Related to COVID19" if applicable).

4.3.1.3 Analysis populations

The number of participants included in each study population (safety population and efficacy populations (ITT and mITT) will be provided. Participants (and visits for mITT population) with reasons of exclusion will be listed.

4.3.2 Analyses of baseline characteristics

4.3.2.1 Participant demographic characteristics, medical history, and diagnoses

Continuous variables (age, weight and body mass index [BMI]) and qualitative variables (gender, race, ethnicity, and BMI [<20 / ≥20]) will be summarized by cohorts, by descriptive statistics (summary tables) for the mITT population and for additional population if relevant (eg, if many participants from the mITT population are not part of the safety population).

Medical/surgical histories will be coded using the version of medical dictionary for regulatory activities (MedDRAMedDRA) in effect at Sanofi at the time of each database lock. Medical/surgical histories will be summarized by cohort and by system organ class (SOC), high level term (HLT) and preferred term (PT) and listed.

Disease characteristics (ie, years since AD diagnosis at baseline) will be summarized by descriptive statistics for the mITT population.

Matched participants will be listed.

4.3.2.2 Baseline efficacy parameters

For efficacy parameters, baseline is defined as the last available and evaluable value at Day 1 (Week 0) (or at Screening if the value at Day 1 is not available) for patients and healthy volunteers.

For TEWL data, the baseline value will be defined for each skin group and each number of STS (before/without STS, after 5 STS, after 10 STS, after 15 STS and after 20 STS). In addition for before/without STS, a baseline will be defined for each spot, due to strong variability between the 3 adjacent spots in a similar study conduct by the Sponsor.

Baseline efficacy values will be presented along with the subsequent efficacy values assessed during the on-treatment period.

4.3.2.3 Baseline safety parameters

For vital sign parameters, baseline will be defined as the last available and evaluable value before and closest to the first dose of IMP for patients and as the last available and evaluable value at Day 1 (Week 0) for healthy volunteers.

Baseline safety values will be presented along with subsequent safety values assessed during the study.

4.3.2.4 Extent of study treatment exposure and compliance

The extent of study treatment exposure and compliance will be assessed and summarized for patients' cohort within the safety population.

A listing of patients receiving IMP from specified batch (patient, IMP, and IMP batch number) will be performed by patient.

Extent of IMP exposure

Duration of IMP exposure is defined regardless of unplanned intermittent discontinuations as:

- Dose regimen 1: the last dose date - first dose date +28 days.
- Dose regimen 2: the last dose date - first dose date +14 days.

If the patient's date of last dose is unknown, his/her last IMP dispensing date will be used in its place.

Duration of exposure will be summarized in patients' cohort using descriptive statistics as quantitative and qualitative variables (1 day / 2 to 15 days / 16 to 29 days / 30 to 43 days / 44 to 57 days / 58 to 71 days / 72 to 85 days / 86 to 99 days / 100 to 113 days / 114 to 155 days / 156 to 197 days). This description could be performed by trial impact (disruption) due to COVID19 (without and with).

Compliance

A given administration will be considered noncompliant if the patient did not take the planned dose of treatment as required by the protocol. No imputation will be made for patients with missing or incomplete data.

Percentage of compliance for a patient will be defined as the number of administrations the patient was compliant divided by the total number of administrations the patient was planned to take (the number of doses missed due to interruptions at Investigators' judgment will not be subtracted) on or before the last IMP administration date.

Treatment compliance, above-planned and under-planned dosing percentages will be summarized descriptively. The percentage of patients with compliance <80% will be summarized. In addition, number and percentage of patients with at least 1 above-planned dosing administration will be

given, as well as the number and percentage of patients with 0,]0, 20%], and >20% under planned dosing administrations.

4.3.2.5 Prior/Concomitant medication/therapy

All medications taken during the study will be coded using the version of world health organization drug dictionary (WHO-DD) in effect at Sanofi at the time of each database lock.

Medications will be classified into three categories, prior, concomitant and post-treatment medications, according to the following rules:

- Prior medications are those the participant used before the first IMP administration. Prior medications can be discontinued before the first IMP administration or can be ongoing during the treatment phase.
- Concomitant medications are any medications received by the participant concomitantly to the IMP, from the first IMP administration to the last IMP administration + 28 and 14 days for regimen dose 1 and regimen dose 2 respectively.
- Post-treatment medications are those the participant took in the period running from the end of the concomitant medications period up to the end of the study.

A given medication can be classified simultaneously as a prior, concomitant and post-treatment medication. If a medication date or time is missing or partially missing and it cannot be determined whether it was taken prior or concomitantly or post, it will be considered as prior, concomitant, and post-treatment medication.

For each category of medications, medications will be summarized by cohort according to the WHO-DD dictionary, considering the first digit of the anatomic therapeutic class (ATC) (anatomic category) and the first three digits of the ATC class (therapeutic category). All ATC codes corresponding to a medication will be summarized, and participants will be counted once in each ATC category (anatomic or therapeutic) linked to the medication. Therefore, participants may be counted several times for the same medication. Concomitant medications will be listed separately by participants.

4.3.3 Analyses of evaluation variables

4.3.3.1 Analysis of main evaluation variable(s)

The primary endpoint will be analyzed on the mITT population. As a sensitivity analysis, it will also be analyzed based on the ITT population.

The calculation of changes (absolute and percent) from baseline will be based on the same number of STS (ie, for absolute change from baseline after 5 STS at W16: Change TEWL at W16 after 5 STS = TEWL value at W16 after 5 STS – TEWL value at baseline after 5 STS).

If a change in spot anatomical location changed versus Day 1 is identified and assessed by the study team as having a significant impact on skin barrier, only visits prior to the location change will be considered in the statistical analysis of TEWL data.Raw data and change (absolute and percent) from baseline will be summarized with descriptive statistics by timepoint.

Profiles over time (from baseline to Week 16) will be generated for individual values (spaghetti plot).

For information purpose as study is not powered (no multiplicity correction will be applied):

- For each number of STS, the evolution of the percent change from baseline over time will be analyzed using a repeated measures analysis of variance with visit, type of skin, interaction between visit and type of skin, age, and baseline value as fixed effects and subject identifier as random effect, at a Type 1 error level of alpha=0.05. In case of heterogeneity at baseline, additional covariates (ie, site, race …) could be added in this statistical model.
- The following covariance structures will be tested in the REPEATED statement of the PROC MIXED: un, vc and cs at least. The structure with the smallest Bayesian information criteria will be selected for the final model.
- Point estimate and two-sided 95% confidence interval and corresponding two-sided p-value will be reported for each effect. No comparison will be performed.
- If assumption of normal distribution is strongly violated, transformation method could be used: eg, log transformation; normality assumption will be assessed through quantile-quantile plot and Shapiro-Wilk test at 5% Type I error.

4.3.3.2 Analyses of secondary evaluation variables

The secondary endpoints will be analyzed on the mITT populations. Secondary endpoints analysis might be run on the ITT population as sensitivity analysis in case an important number of patients using rescue therapy is observed.

Due to high variability between the 3 adjacent spots for TEWL data in a similar study conducted by the Sponsor, the average of these 3 adjacent spots will not be appropriated. Consequently, TEWL data collected at the same visit from without or before STS measurement on each of the 3 spots within the predefined skin areas will be analyzed by spot separately. The calculation of changes (absolute and percent) from baseline at a visit before/without STS measurement will be based on the same spot (ie, for absolute change from baseline at W2 before STS in spot 1: Change TEWL at W2 before STS in spot 1 = TEWL value at W2 before STS in spot 1 – TEWL value at baseline before STS in spot 1). The calculation of changes (absolute and percent) from baseline at a visit after STS measurement will be based on the same number of STS, as for the primary endpoint.

If a change in spot anatomical location changed versus Day 1 is identified and assessed by the study team as having a significant impact on skin barrier, only visits prior to the location change will be considered in the statistical analysis of TEWL data.

For the analysis of change (percent and absolute) from baseline in TEWL after 5, 10, 15 and 20 STS assessed on each skin group at Week 16, the similar approach, as for the analysis of primary endpoint, will be used. For each number of STS, point estimate, two-sided 95% confidence interval and corresponding two-sided p-value will be reported for each effect. No comparison will be performed.

For the analyses of change (percent and absolute) from baseline in TEWL before/without and after each planned number of STS on each skin group over time:

- Raw data and change (absolute and percent) from baseline will be summarized with descriptive statistics by skin group, number of STS and spot (if applicable), and timepoint. For follow-up period data, change from Week 16 (or last visit prior to use of rescue treatment) will also be reported.
- Profiles over time will be generated for individual values (spaghetti plot) and group means (±SEM) for raw data and absolute change from baseline/Week 16 by skin group, number of STS and spot (if applicable).
- As an exploratory analysis, the time to improvement of skin barrier function for lesional and non-lesional skin area measured by TEWL will be summarized for the treatment phase. Improvement is defined as time to the first post-baseline day for which at least 1 TEWL measure of a patient is equal or lower than the TEWL median in the matched heathy volunteers over time. Results might be reported through Kaplan-Meier plot in case of high number of events. Time to improvement of skin barrier function for lesional and non-lesional skin area will also be listed.

For the analyses of change (percent and absolute) from baseline in TEWL AUC (a composite measure before and after 5, 10, 15, and 20 STS) for skin barrier function in each skin group over time:

- As all TEWL values are expected to be positive, area under the TEWL curve will be calculated using the trapezoidal method over the 20 STS (measurement at every 5 STS and all other unplanned number of STS) at each visit. For AUC derivation only, TEWL data performed at unplanned number of STS will be used.
- Raw data and change (absolute and percent) from baseline will be summarized with descriptive statistics by skin group and timepoints. For follow-up period data, change from Week 16 (or last visit prior to use of rescue treatment) will also be reported.
- Box plots over time will be generated for absolute change from baseline/Week 16.
- As exploratory analysis, time to improvement of skin barrier function for lesional and non-lesional skin area measured by TEWL AUC will be summarized for the treatment phase. Improvement is defined as time to the first post-baseline day for which TEWL AUC of a patient is equal or lower than the median, of the corresponding TEWL AUC computed in his/her matched heathy volunteer over time. Results might be reported through Kaplan-Meier plot in case of high number of events.

TEWL conditions (ie, room temperature and humidity, etc.) will be summarized with descriptive statistics by timepoint.

4.3.3.3 Analyses of tertiary/exploratory evaluation variables

The tertiary/exploratory endpoints will be analyzed on the mITT and ITT populations.

For the analysis of change (percent and absolute) from baseline in epidermal hypertrophy parameters and in stratum corneum lipid structure over time, raw data and change (absolute and percent) from baseline will be summarized with descriptive statistics by skin group and timepoint. For follow-up period data, change from Week 16 (or last visit prior to use of rescue treatment) will also be reported. Profiles over time will be generated for individual values and group means for raw data and absolute change from baseline.

For the analysis of the change (percent and absolute) from baseline in CDLQI, EASI, ISS for target lesion, POEM, assessment of worst itch NRS, sleep disturbance NRS, skin pain NRS, and photograph outputs (TiVi and skin roughness indexes) obtained from skin imaging over time, raw data and change (absolute and percent) from baseline will be summarized with descriptive statistics by skin group and timepoint. For follow-up period data, change from Week 16 (or last visit prior to use of rescue treatment) will also be reported. Profiles over time will be generated for individual values (spaghetti plot only for PRO measures [POEM, CDLQI, worst itch NRS, sleep disturbance NRS, and skin pain NRS]) and group means for raw data and absolute change from baseline (only for ClinROs measures [EASI and ISS]).

A correlation analysis will also be investigated between baseline values of TEWL (before and after each number of STS), and TEWL AUC in lesional and non-lesional skin of AD patients with the following baseline measures: EASI, targeted lesion ISS, image-derived severity score in targeted lesional skin, and quantified epidermal hypertrophy and structure changes in OCT and FTIR measures. For this analysis, scatterplot to visualize the form of the relationship between variables will be produced for patients' cohort at specific time points: baseline, Week 8, Week 16 and Week 28. Pearson correlation will be assessed at each defined time point for EASI, ISS, photograph outputs, OCT outputs: epidermal thiceasikness, supervidial plexus depth, blood vessel diameter and density, and FTIR measures. Pearson correlation coefficient and p-value will be reported. Spearman's rank correlation will be assessed at each defined timepoint for ISS. Spearman correlation coefficient and p-value will be reported. As an exploratory analysis repeated measures correlation method might be used to assess the overall correlation pooling data from baseline, Week 8, Week 16 and Week 28 (3).

The same approach will be used to investigate the correlation between percent change from baseline in TEWL (before and after each number of STS), and TEWL AUC in lesional and non-lesional skin of AD patients at Week 8, Week 16 and Week 28 with corresponding change from baseline of the aforementioned measures (ie, EASI, targeted lesion ISS …).

4.3.4 Additional exploratory analyses

In a similar study conducted by the Sponsor, differences in clinical conditions were also identified between centres (ie, room temperature in a specific centre higher than required by the study protocol). Consequently, the centre effect will be investigated and some covariates (ie, room temperature …) could be included in the statistical models.

Additionally, the following comparisons will be investigated on ITT population only:

- The comparison in TEWL between Day 1 and Week 16 will be performed by skin group (lesional and non-lesional skin area only), STS and spot using a non-parametric Wilcoxon signed-rank test.
- The comparison in TEWL AUC between Day 1 and Week 16 will be performed by skin group (lesional and non-lesional skin area only) and spot using a non-parametric Wilcoxon signed-rank test.
- The comparison between AD participants (lesional skin area/non-lesional skin area) and Healthy participants will be performed on TEWL data at Week 16 using a non-parametric Wilcoxon signed-rank test.
- The comparison AD participants (lesional skin area/non-lesional skin area) versus Healthy participants will be performed on TEWL AUC data at Week 16 using a non-parametric Wilcoxon signed-rank test.

4.3.5 Analyses of safety parameters

All the safety analyses will be performed using the safety population. When applicable, results will be by cohort and overall.

The safety analysis will be conducted according to the Sponsor's document "Analysis and reporting of safety data from Clinical Trials through the Clinical Study Report" (BTD-009536 Version 3.0 dated on 24 May 2014).

4.3.5.1 Adverse events

Adverse events will be coded to a lower level term (LLT), PT, HLT, high level group term (HLGT) and primary SOC using the version of MedDRA in effect at Sanofi at the time of each database lock.

For patients' cohort, they will be classified into predefined standard categories according to chronological criteria:

- Pre-treatment AEs: AEs that occurred, worsened or became serious during the pre-treatment period defined as the time between informed consent signature and the first IMP administration.
- Treatment emergent AEs (TEAEs): AEs that occurred, worsened or became serious during the TEAE period defined as the time from the first IMP administration up to the EoT visit (EoT included).
- Post-TEAEs: AEs that occurred, worsened or became serious during the post-TEAE period defined as the time starting after the TEAE period (including the follow-up period).

Treatment-emergent AEs will be assigned to the treatment received at the time of the AE onset. If the onset date (or time) of an AE (occurrence, worsening or becoming serious) is incomplete or missing, then the AE will be considered as a TEAE unless a partial date (or time) shows it as a pre- or post-treatment event.

For all healthy volunteers, safety data will be considered as part of a single period starting with the signature of the informed consent and ending with the EoS visit (EoS included).

All AEs reported in the study will be listed, sorted by participant, onset date and time.

The following TEAEs summaries will be provided for the patients' cohort of the safety population:

- Overview of TEAEs: number and percentage of participants with any TEAE, any serious TEAE, any TEAE leading to death (if any occurred), any TEAE leading to permanent treatment discontinuation, and any TEAE of special interest.
- Summary of TEAEs by primary SOC and PT:
  - Number and percentage of patients with at least one TEAE.
  - Number of occurrences of TEAEs.

Patients presenting TEAEs will be listed sorted by primary SOC and PT. By definition, no TEAEs summary will be displayed for healthy volunteers. AEs that occur outside the treatment-emergent period or in healthy volunteers will be summarized separately.

In addition, the following listings/summary will be provided:

- Listing of any deaths and serious AEs.
- Listing of any AEs leading to permanent treatment discontinuation.
- Number (%) of patients experiencing treatment emergent adverse events of special interest (AESI) will be presented by AESI category and PT, sorted by decreasing incidence of PT within each AESI category.
- Listing of AESIs.
- Number (%) of patients experiencing at least one treatment emergent COVID19 related adverse event by primary SOC and PT.

4.3.5.2 Vital signs

Heart rate, systolic and diastolic blood pressures (SBP and DBP), body temperature and respiratory rate will be analyzed as raw parameter value and change from baseline.

Body weight will be analyzed as raw parameter value and percent change from baseline. BMI will be analyzed as raw parameter value.

Baseline definition

The values to be used as baseline will be the Day 1 (Visit 2, Week 0) predose assessment. If any of the scheduled baseline tests are repeated for any participant, the last rechecked values will be considered as baseline, provided, for patients, they were done before the IMP administration of the respective treatment phase, and in the same condition.

Abnormalities analyses

For vital sign parameters, analysis will be performed using all post-baseline assessments done during the TEAE period/safety period, including all unplanned and rechecked values. Counts of participants with potentially clinically significant abnormal (PCSA) values will be presented by cohort, regardless of the baseline status. PCSA values are defined as abnormal values considered medically important by the Sponsor according to predefined criteria/thresholds based on literature review and defined by the Sponsor.

A listing of individual data from participants with post-baseline PCSAs will be provided; values will be flagged when reaching the PCSA criteria.

PCSA criteria are detailed in Section 8 (Appendix D).

Descriptive statistics and plots

For heart rate, blood pressures, body temperature and respiratory rate, raw data and absolute changes from baseline will be summarized in descriptive statistics for patients' cohort and scheduled time of measurement.

For body weight, raw data and percent change from baseline will be summarized in descriptive statistics by cohort and scheduled time of measurement.

4.3.5.3 Clinical laboratory parameters

Positive results as well as all individual data for urine pregnancy tests (for women with childbearing potential only) will be listed.

5 DATA HANDLING CONVENTIONS

5.1 DATA HANDLING CONVENTIONS FOR PATIENT DATA

In all tables, figures, listings and in-texts, the following cohort labels will be used:

- Healthy Participant.
- AD Participant.

5.2 DATA HANDLING CONVENTIONS FOR PATIENT REPORTED OUTCOMES

When multiple answers are reported by the patient, the data will be considered as missing.

5.3 MISSING DATA

Missing data imputation (last observation carried forward (LOCF) method or another complex method) could be envisaged due to a high number of missing data and high variability. Sensitivity analyses could also be planned.

5.4 DEFINITION OF REGIONS

Not applicable.

5.5 WINDOWS FOR TIME POINTS

Except for the baselines, only observations from visits and/or timepoints planned in the protocol will be used in descriptive statistics. Usually, all timepoints planned in the protocol should be described.

All the data collected from eDiary (Patient-reported outcome) were linked to a unique unscheduled visit in the SDTM database. To perform the analysis by visit, the following visit window algorithm is defined.

VISIT WINDOW ALGORITHM		
Patient-Reported Outcomes	**Planned visits in the protocol**	**Time interval for analysed visits**
Worst itch NRS	D1	-1 to 1 day before the first IMP administration
Sleep disturbance NRS	From D2 to D28	x days after the first IMP administration for each Dx
Skin pain NRS	D36	29 to 39 days after the first IMP administration

VISIT WINDOW ALGORITHM		
Patient-Reported Outcomes	**Planned visits in the protocol**	**Time interval for analysed visits**
	D43	40 to 50 days after the first IMP administration
	D57	51 to 71 days after the first IMP administration
	D85	72 to 99 days after the first IMP administration
	D113	100 to 134 days after the first IMP administration
	D155	135 to 176 days after the first IMP administration
	D197	177 to 200 days after the first IMP administration

If more than one record occurs within the same visit window where only one assessment is expected, then the following rule will be applied: for D1, the last non-missing result prior to study drug administration will be used; for post-treatment assessments the closest non-missing result to the scheduled visit will be used.

5.6 STATISTICAL TECHNICAL ISSUES

No hypothesis testing is predefined in this exploratory study.

6 INTERIM ANALYSIS

An analysis of the primary and secondary endpoints was performed after the inclusion of 10 AD patients who have completed the TEWL assessments at Day 1. Descriptive statistics and plots were provided only.

The interim analysis was performed using the safety population for safety parameters, the mITT and ITT populations for TEWL parameters. Only primary and secondary objectives for which data are available were analyzed.

As it is an open-label study with no predefined hypothesis testing there is no issue regarding blinding or multiplicity adjustment.

As specified in Section 3, two analyses will be conducted (one for the treatment phase and one for the follow-up period) but none of these will be considered as interim analysis. The analysis for the treatment phase will be descriptive and will include the following analyses:

- Disposition of patients: participant disposition, analysis population, participant demographic characteristics, medical history, and diagnoses.
- Efficacy analysis: descriptive statistics and the repeated measures mixed model of primary endpoint and secondary endpoints, and tertiary/exploratory endpoints [ClinROs (EASI and ISS) and PROs via eDiary (POEM, CDLQI and Worst itch/Sleep disturbance/Skin pain NRS) only].
- Safety analysis: TEAEs summaries only.

7 SOFTWARE DOCUMENTATION

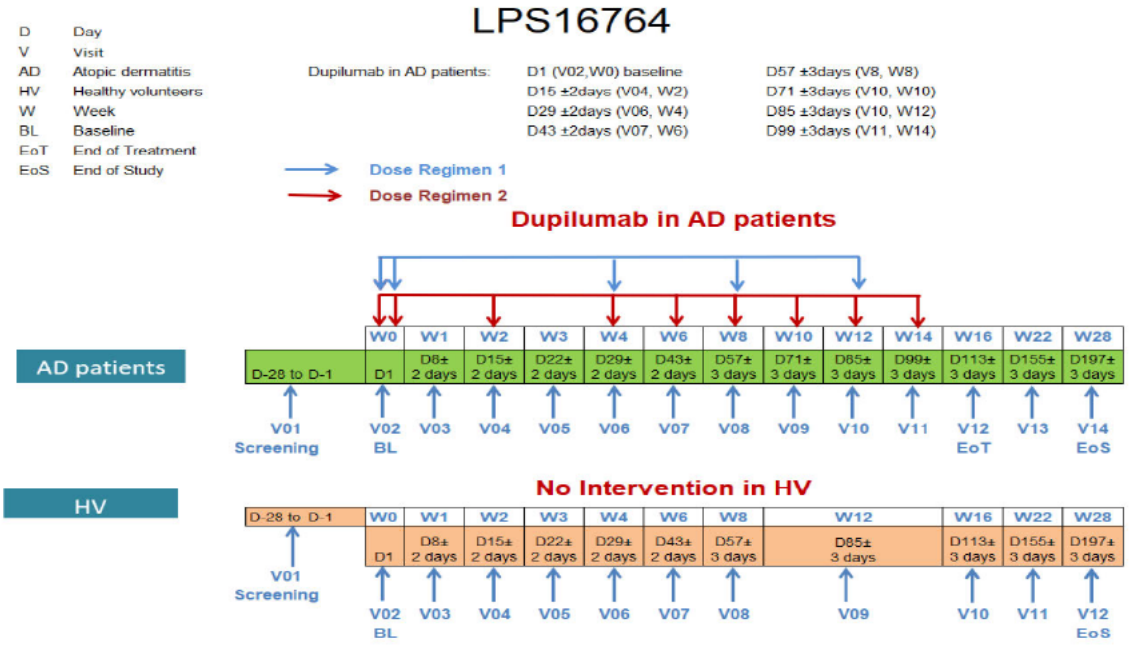
All summaries and statistical analyses will be generated using SAS® software Version 9.4 (SAS institute Inc. Cary NC USA).

8 LIST OF APPENDICES

Appendix name	Title
Appendix A	Graphical study design
Appendix B	Schedule of activities in pediatric patients
Appendix C	Schedule of activities in healthy volunteers
Appendix D	Potentially clinically significant abnormalities criteria

8.1 APPENDIX A GRAPHICAL STUDY DESIGN

Figure 1 - Graphical Study Design



8.2 APPENDIX B SCHEDULE OF ACTIVITIES IN PEDIATRIC PATIENTS

Table 1 - Schedule of activities in pediatric patients

Phase	Screening	Baseline	Treatment phase									EoT	Follow-up period		Unscheduled visit	Premature EoT
Day	D-28 to D-1	D 1	D8 ±2 days	D15 ±2 days	D22 ±2 days	D29 ±2 days	D43 ±2 days	D57 ±3 days	D71 ±3 days	D85 ±3 days	D99 ±3 days	D113 ±3 days	D155 ±3 days	EoS D197 ±3 days	UNSCH	
Week		W0	W1	W2	W3	W4	W6	W8	W10	W12	W14	W16	W22	W28		
Visits	V01	V02	V03	V04	V05	V06	V07	V08	V09	V10	V11	V12	V13	V14		
Informed consent/assent form	X															
Inclusion/exclusion criteria	X	X														
Medical/surgical history/IGA/demographics	X															
Pregnancy test (urine), WOCBP only[a]	X	X	X	X	X	X	X	X	X	X	X	X	X	X		
Prior/concomitant medications/procedures	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Confirm emollient and washing compliance[b]	X	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	X[b]	X	X[b]
Study treatment administration																
SC administration of dupilumab, dose regimen 1[c]		X				X		X		X					(X)	
SC administration of dupilumab, dose regimen 2[c]		X		X		X	X	X	X	X	X				(X)	

Phase	Screening	Baseline	Treatment phase									EoT	Follow-up period		Unscheduled visit	Premature EoT
Day	D-28 to D-1	D 1	D8 ±2 days	D15 ±2 days	D22 ±2 days	D29 ±2 days	D43 ±2 days	D57 ±3 days	D71 ±3 days	D85 ±3 days	D99 ±3 days	D113 ±3 days	D155 ±3 days	EoS D197 ±3 days	UNSCH	
Week		W0	W1	W2	W3	W4	W6	W8	W10	W12	W14	W16	W22	W28		
Visits	V01	V02	V03	V04	V05	V06	V07	V08	V09	V10	V11	V12	V13	V14		
Skin barrier function test in lesional and non-lesional skin[d]																
TEWL before STS[e, f]		X	X	X	X	X	X	X		X		X	X	X	(X)	X
TEWL and STS assessment[f]		X[g]		X[g]		X[g]		X[g]		X[g]		X[g]	X[g]	X[g]	(X)[h]	X[h]
Lipidomics, proteomics, FLG breakdown products, and transcriptomics assessment from skin tape strips		X[g]		X[g]		X[g]		X[g]		X[g]		X[g]	X[g]	X[g]	(X)[h]	X[h]
Optical Coherence Tomography (OCT)[i]		X	X	X	X	X	X	X		X		X		X	(X)	X
Fourier-Transform Infrared (FTIR) Spectroscopy[i]		X						X				X		X	(X)	X
Standardized photographs of predefined lesional and non-lesional skin areas used for TEWL		X	X	X	X	X	X	X		X		X	X	X	(X)	X
Standardized full body photographs		X	X	X	X	X	X	X		X		X	X	X	(X)	X
EASI	X	X	X	X	X	X	X	X		X		X	X	X	(X)	X
ISS	X	X	X	X	X	X	X	X		X		X	X	X	(X)	X

Phase	Screening	Baseline	Treatment phase									EoT	Follow-up period		Unscheduled visit	Premature EoT
Day	**D-28 to D-1**	**D 1**	**D8 ±2 days**	**D15 ±2 days**	**D22 ±2 days**	**D29 ±2 days**	**D43 ±2 days**	**D57 ±3 days**	**D71 ±3 days**	**D85 ±3 days**	**D99 ±3 days**	**D113 ±3 days**	**D155 ±3 days**	**EoS D197 ±3 days**	**UNSCH**	
Week		**W0**	**W1**	**W2**	**W3**	**W4**	**W6**	**W8**	**W10**	**W12**	**W14**	**W16**	**W22**	**W28**		
Visits	**V01**	**V02**	**V03**	**V04**	**V05**	**V06**	**V07**	**V08**	**V09**	**V10**	**V11**	**V12**	**V13**	**V14**		
Patient reported outcomes (PRO)																
POEM		X	X	X	X	X	X	X		X		X	X	X	(X)	X
CDLQI		X	X	X	X	X	X	X		X		X	X	X	(X)	X
Worst itch NRS[j]	X	X	X	X	X	X	X	X		X		X	X	X	(X)	X
Sleep disturbance NRS[j]	X	X	X	X	X	X	X	X		X		X	X	X	(X)	X
Skin pain NRS[j]	X	X	X	X	X	X	X	X		X		X	X	X	(X)	X
eDiary[k]	X	X	X	X	X	X	X	X		X		X	X	X	(X)	X
Safety																
AE/SAE collection	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Height		X										X		X		X
Body weight	X	X										X		X		X
Vital signs[l]	X	X		X		X	X	X		X		X		X	(X)	X
Physical examination[m]	X	X[m]										X[m]		X[m]		X[m]
Saliva swab sample[n]		X														
Optional assessments																
Optional Biomarker research samples (blood)[o]		X										X				

a For the purpose of this study, any female who has had her first menstrual period (menarche) and is sexually active will be considered to be of childbearing potential. Female patients who are not of childbearing potential at the start of the study but have the onset of menarche during the course of the study and are sexually active will also have to follow adequate birth control methods to continue participation in the study. These females must have a negative urine β-HCG pregnancy test at screening and at each visit (except patients who assigned to dose regimen 1 will not attend visits V9 (W10) and V11 (W14), therefore there's no pregnancy test for these patients on both visits). Pregnancy will lead to definitive treatment discontinuation in all cases.

b Emollients should NOT be applied from Day -7 to the EoS (Week 28) to the targeted, predefined skin areas that will be used for TEWL assessment. Participants should not take showers or soaking in a bathtub within 6 hours before TEWL assessment. Emollients for use during the study will be provided by the study site to study participants. Use of emollients should be documented in a diary provided by the study site.

c Dose regimen 1: Children with baseline 15 kg ≤ body weight <30 kg will receive a SC loading dose of dupilumab 600 mg (2 injections of dupilumab 300 mg) on Day 1 (V02), followed by every 4-week SC dosing of dupilumab 300 mg until Week 12 (at Week 0 [(Day 1], Weeks 4, 8, and 12).

Dose regimen 2: Children with baseline 30 kg ≤ body weight <60 kg will receive a SC loading dose of dupilumab 400 mg (2 injections of dupilumab 200 mg) on Day 1 (V02), followed by bi-weekly SC dosing of dupilumab 200 mg until Week 14 (at Week 0 (Day 1), Weeks 2, 4, 6, 8, 10, 12, and 14).

d Assessments on treatment days have to be conducted before administration of dupilumab.

e Predefined lesional and non-lesional skin areas. TEWL will be measured at each of the 3 closely adjacent spots within the predefined skin areas at all visits without STS and before STS at visits with STS.

f TEWL will be conducted before STS and then after 5, 10, 15, and 20 STS in lesional skin and non-lesional skin.

g At Weeks 0, 8, 16, and 28, STS will be conducted on the first spot within the predefined skin area. At Week 2 STS will be conducted on the second spot within the predefined skin area. At Weeks 4, 12, and 22 STS will be conducted on the third spot within the predefined skin area. TEWL will be assessed in lesional skin and non-lesional skin before STS and after 5, 10, 15, and 20 STS. All skin tape strips should be collected and stored.

h In case STS assessment is to be conducted at an unscheduled visit or at a premature EoT visit the assessment should be conducted at that spot of the skin area, for which the period passed since the last STS assessment is the longest.

i OCT will be done before STS/TWL, and FTIR will be done before and during STS.

j Worst itch NRS, Sleep disturbance NRS and Skin pain NRS will be assessed daily from Day -7 to Day -1; daily from Day 1 to Week 4 and then 7 days prior to Week 6, Week 8, Week 12, Week 16, Week 22, and to Week 28 by e-diary.

k eDiary will be dispensed to the patients at screening (Visit 01) and will be collected on Week 28 (Visit 14). The PROs will be administered to patients via eDiary. Patients will bring the eDiary to the site at each visit, and it will be reviewed for any questionnaires omission and dispensed back to patients at each visit.

l Vital signs, including heart rate, systolic and diastolic blood pressure (mmHg), pulse rate (beats per minute), body temperature (°C), and respiratory rate will be measured.

m Limited to skin-related physical examination at baseline, EoT visit (Week 16) and EoS (Week 28).

n Saliva swab samples for DNA isolation should be collected on Day 1/Baseline (predose) for FLG gene sequencing analysis, but may be collected at any study visit.

o Please refer to Section 8.8.6 of protocol for details.

AE = Adverse Events; AD = Atopic Dermatitis; CDLQI = Children's Dermatology Life Quality Index; D= Day; EASI = Eczema Area and Severity Index; EoT = End of Treatment; EoS = End of Study; FLG = Filaggrin; IGA = Investigator Global Assessment; ISS = Individual Signs Score; NRS = Numerical Rating Scale; POEM = Patient Oriented Eczema Measure; SAE = Serious Adverse Events; SC = Subcutaneous; STS = Skin Tape Stripping; TEWL = Transepidermal Water Loss; UNSCH = unscheduled; V = Visit; W = Week; W = Women of Childbearing Potential.

8.3 APPENDIX C SCHEDULE OF ACTIVITIES IN HEALTHY VOLUNTEERS

Table 2 - Schedule of activities in healthy volunteers

Phase	Screening	Baseline								EoT	Follow-up period		Unscheduled visit	Premature EoT
Day	**D-28 to D-1**	**D 1**	**D8 ±2 days**	**D15 ±2 days**	**D22 ±2 days**	**D29 ±2 days**	**D43 ±2 days**	**D57 ±3 days**	**D85 ±3 days**	**D113 ±3 days**	**D155 ±3 days**	**EoS D197 ±3 days**		
Week		**W0**	**W1**	**W2**	**W3**	**W4**	**W6**	**W8**	**W12**	**W16**	**W22**	**W28**		
Visits	**V01**	**V02**	**V03**	**V04**	**V05**	**V06**	**V07**	**V08**	**V09**	**V10**	**V11**	**V12**		
Informed consent/assent form	X													
Inclusion/exclusion criteria	X	X												
Matching by age (±2 years) and gender	X													
Medical/surgical history/demographics	X													
Pregnancy test (urine), WOCBP only[a]	X	X	X	X	X	X	X	X	X	X	X	X		
Prior/concomitant medications/procedure	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Confirm emollient and washing compliance[b]	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Skin barrier function test in predefined normal skin														
TEWL before STS[c, d]		X	X	X	X	X	X	X	X	X	X	X	X	X
TEWL and STS assessment[d]		X[e]		X[e]		X[e]		X[e]	X[e]	X[e]	X[e]	X[e]	X[f]	X[f]
Lipidomics, proteomics, FLG breakdown products and transcriptomics		X[e]		X[e]		X[e]		X[e]	X[e]	X[e]	X[e]	X[e]	X[f]	X[f]

Phase	Screening	Baseline								EoT	Follow-up period		Unscheduled visit	Premature EoT
Day	**D-28 to D-1**	**D 1**	**D8 ±2 days**	**D15 ±2 days**	**D22 ±2 days**	**D29 ±2 days**	**D43 ±2 days**	**D57 ±3 days**	**D85 ±3 days**	**D113 ±3 days**	**D155 ±3 days**	**EoS D197 ±3 days**		
Week		**W0**	**W1**	**W2**	**W3**	**W4**	**W6**	**W8**	**W12**	**W16**	**W22**	**W28**		
Visits	**V01**	**V02**	**V03**	**V04**	**V05**	**V06**	**V07**	**V08**	**V09**	**V10**	**V11**	**V12**		
assessment from skin tape strips[e]														
Standardized photographs of healthy skin area used for TEWL		X	X	X	X	X	X	X	X	X	X	X	(X)	X
Standardized full body photographs		X	X	X	X	X	X	X	X	X	X	X	(X)	X
Optical Coherence Tomography (OCT)[g]		X	X	X	X	X	X	X	X	X		X	(X)	X
Fourier-Transform Infrared (FTIR) Spectroscopy[g]		X						X		X		X	(X)	X
Safety														
AE/SAE collection	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Height		X								X		X		X
Body weight	X	X								X		X		X
Vital signs[h]	X	X						X		X				X
Physical examination[i]	X	X[i]								X[i]		X[i]		X[i]
Saliva swab sample[j]		X												
Optional assessments														
Biomarker research samples (blood)[k]		X								X				

a For the purpose of this study, any female who has had her first menstrual period (menarche) and is sexually active will be considered to be of childbearing potential. Female patients who are not of childbearing potential at the start of the study but have the onset of menarche during the course of the study and are sexually active will also have to follow adequate birth control methods to continue participation in the study. These females must have a negative urine ß-HCG pregnancy test at screening and at each visit. Pregnancy will lead to definitive treatment discontinuation in all cases.

b Emollients should NOT be applied from Day -7 to the EoS (Week 28) to the targeted, predefined skin areas that will be used for TEWL assessment. Participants should not take showers or soak in a bathtub within 6 hours before TEWL assessment. Emollients for use during the study will be provided by the study site to study participants. Use of emollients should be documented in a diary provided by the study site.

c TEWL will be measured at each of the 3 closely adjacent spots within the predefined skin areas at all visits without STS and before STS at visits with STS.

d TEWL will be conducted before STS and then after 5, 10, 15, and 20 STS for skin barrier function assessment in normal skin and at the same location as for the lesional skin in the matching AD patient. All skin tape strips will be collected and stored.

e At Weeks 0, 8, 16, and 28, STS will be conducted on the first spot within the predefined skin area. At Week 2 STS will be conducted on the second spot within the predefined skin area. At Weeks 4, 12, and 22 STS will be conducted on the third spot within the predefined skin area. TEWL will be assessed in NORMAL skin before STS and 5, 10, 15, and 20 STS. All skin tape strips should be collected and stored.

f In case STS assessment is to be conducted at an unscheduled visit or at a premature end of treatment visit the assessment should be conducted at that spot of the skin area, for which the period passed since the last STS assessment is the longest.

g OCT will be done before STS/TWL, and FTIR will be done before and during STS.

h Vital signs, including heart rate, systolic and diastolic blood pressure (mmHg), pulse rate (beats per minute), body temperature (°C), and respiratory rate will be measured.

i Limited to skin-related physical examination at baseline, EoT visit (Week 16), and EoS (Week28).

j Saliva swab samples for DNA isolation should be collected on Day 1/Baseline (predose) for FLG gene sequencing analysis, but may be collected at any study visit.

k Please refer to Section 8.8.6 of protocol for details.

AE = Adverse Events; AD = Atopic Dermatitis; CDLQI = Children's Dermatology Life Quality Index; D = Day; EASI = Eczema Area and Severity Index; EoT = End of Treatment; EoS = End of Study; FLG = Filaggrin; IGA = Investigator Global Assessment; NRS = Numerical Rating Scale; POEM = Patient Oriented Eczema Measure; SAE = Serious Adverse Events; SC = Subcutaneous; STS = Skin Tape Stripping; TEWL = Transepidermal Water Loss; V = Visit; W = Week; W = Women of Childbearing Potential.

8.4 APPENDIX D POTENTIALLY CLINICALLY SIGNIFICANT ABNORMALITIES (PCSA) CRITERIA

Table 3 - PCSA criteria for vital signs parameters for studies in children (24 May 2014)

CRITERIA for POTENTIALLY CLINICALLY SIGNIFICANT ABNORMALITIES For Studies in Children			
Parameter	**Age range**	**PCSA**	**Comments**
Vital Signs			Ref. : Kidney Disease Outcomes Quality Initiatives (KDOQI) Guideline 13; 1996; The fourth report on the diagnosis, evaluation, and treatment of high blood pressure in children and adolescents, Pediatrics 2004; Bowman E & Fraser S Neonatal Handbook 2012; Mulberg AE et al. Pediatric Drug Development Concepts and applications. John Wiley & sons, Inc. 2009; Pediatric respiratory rates http://www.health.ny.gov/
SBP	6 to <12 years old (Children)	≤80 mmHg and decrease from baseline ≥20 mmHg ≥108 mmHg and increase from baseline ≥20 mmHg	Based on definition of Hypertension as average SBP or DBP ≥95th percentile for gender, age, and height on ≥3 occasions
	12 to 16/18 years old (Adolescents)	≤90 mmHg and decrease from baseline ≥20 mmHg ≥119mmHg and increase from baseline ≥20 mmHg	
DBP	6 to <12 years old (Children)	≤48 mmHg and decrease from baseline ≥10 mmHg ≥72mHg and increase from baseline ≥10 mmHg	
	12 to 16/18 years old (Adolescents)	≤54 mmHg and decrease from baseline ≥10 mmHg ≥78mHg and increase from baseline ≥10 mmHg	
Temperature	All age ranges	Rectal, ear or temporal artery: ≥100.4 °F/38.0 °C Oral or pacifier: ≥99.5 °F/37.5 °C Axillary or skin infrared: ≥99 °F/37.2 °C	Ear temperature not accurate below 6 months of age
Respiratory rate	6 to <12 years old (Children)	<18 per minutes >30 per minutes	Based on normal range
	12 to 16/18 years old (Adolescents)	<12 per minutes >20 per minutes	
Weight	All ranges	≥5 % weight loss from baseline	Based on identification of trends in the child's growth with a series of visits WHO Multicentre Reference Study Group, 2006; Center for Disease Control. Growth chart 2007

9 REFERENCES

1. International Conference on Harmonisation of technical requirements for registration of pharmaceuticals for human use. ICH Harmonised tripartite guideline. Statistical Principles for Clinical Trials E9 – Current step 4 dated 5 February 1998.

2. Danby SG, Al-Enezi T, Sultan A, Chittock J, Kennedy K, Cork MJ. The effect of aqueous cream BP on the skin barrier in volunteers with a previous history of atopic dermatitis. Br J Dermatol. 2011;165(2):329-34.

3. Bakdash JZ, Marusich LR. Repeated measures correlation. Front Psychol. 2017;8:456.

Signature Page for VV-CLIN-0636740 v2.0
lps16764-16-1-9-sap

Approve & eSign	[redacted]
Approve & eSign	[redacted]
Approve & eSign	[redacted]

STATISTICAL ANALYSIS PLAN

FOR TRANSLATIONAL RESEARCH

PEdiatric skin barrier function and LIpidomics Study in patients with Atopic Dermatitis - PELISTAD

Open label exploratory study to evaluate the effect of dupilumab on skin barrier function in pediatric patients with moderate to severe atopic dermatitis

STUDY NUMBER: LPS16764

STUDY NAME: PELISTAD

COMPOUND: SAR231893/REGN668

SANOFI STATISTICIAN: [redacted]

DATE OF ISSUE: 19-Dec-2022

VERSION: 1.0

Total number of pages: 24

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

AD:	atopic dermatitis
ANOVA:	analysis of variance
ATR:	attenuated total reflectance
AUC:	area under the curve
BMI:	body mass index
CDLQI:	children's dermatology life quality index
ClinRO:	clinician reported outcome assessments
cpm:	count per million
CSR:	clinical study report
EASI:	eczema area and severity index
FLG:	filaggrin
FTIR:	Fourier transform infrared
IL:	interleukin
IPA:	ingenuity pathway analysis
IQR:	interquartile range
ISS:	individual signs score
LLOQ:	lower limit of quantification
MDS:	multidimensional scaling
NRS:	numerical rating scale
OCT:	optical coherence tomography
PCA:	pyroglutamic acid, principal component analysis
POEM:	patient oriented eczema measure
PRO:	patient reported outcome
SD:	standard deviation
SEM:	standard error of mean
STS:	skin tape stripping
TEWL:	transepidermal water loss
t-SNE:	t-distributed stochastic neighbor embedding
UCA:	urocanic acid
VST:	variance stabilizing transformation

1 OVERVIEW AND INVESTIGATIONAL PLAN

1.1 INTRODUCTION

Atopic dermatitis (AD) is a chronic systemic inflammatory skin disease with a prevalence of up to 25% in children and up to 7% in adults. AD is caused by the complex interplay between epithelial dysfunction and dysregulated/over-activated Type 2 immune response in the skin, with a special role for IL-4/IL-13-driven signaling in AD pathogenesis. This Type 2 hyperactivation blocks terminal differentiation of skin keratinocytes and formation of a mature stratum corneum that is mainly responsible for the skin barrier function.

Dupilumab is a human monoclonal antibody that blocks the shared receptor subunit for interleukin (IL) -4 and IL-13, thus inhibiting signaling of both IL-4 and IL-13, cytokines that are key drivers of Type 2 inflammatory diseases. Dupilumab is recently approved in the US for the treatment of pediatric patients ≥6 and <12 years of age with moderate-to-severe AD whose disease is not adequately controlled with topical prescription therapies or when those therapies are not advisable.

The LPS16764 study is an open label exploratory phase 4 study designed for assessing the effect of subcutaneous injection of dupilumab on skin barrier function measured as transepidermal water loss (TEWL) in AD pediatric patients. The effect of dupilumab treatment on skin barrier function is evaluated using TEWL in conjugation with skin lipidomics, proteomics, and transcriptomics generated from skin tape stripping (STS) samples. Skin barrier structure changes is evaluated using optical coherence tomography (OCT) and Fourier-Transform Infrared (FTIR) spectroscopy in both predefined lesional and non-lesional skin areas.

This document provides a statistical analysis plan for all biomarker data repeatedly collected during the LPS16764 study. The term biomarker in this document will then refer to lipidomics, proteomics, transcriptomics, or filaggrin (FLG) breakdown products. TEWL may refer to a measurement before STS; after a given number of STS samplings or can be AUC (area under the curve) of TEWL before STS and after STSs performed at the specified STS visits. The term study group or group in general refers to lesional, non-lesional or normal health skin group. This biomarker statistical analysis plan is based on clinical trial protocol Version 1 (13-Jul-2020).

1.2 STUDY DESIGN

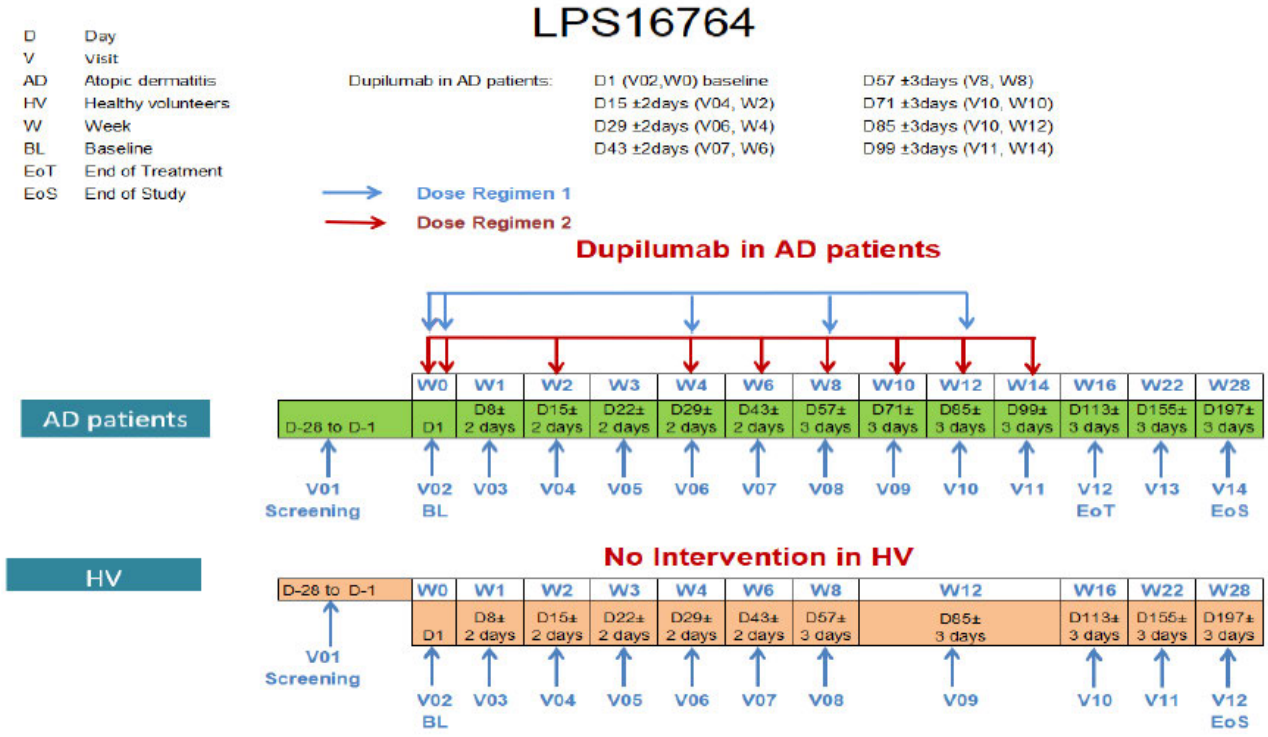
LPS16764 is a Phase 4 open-label, exploratory study in two sites evaluating the effect of dupilumab on skin barrier function in pediatric patients aged between ≥6 and <12 years old with moderate-to-severe AD. A healthy volunteer cohort will be enrolled as a reference comparator matched to AD for age-, gender-, study-site and sampling will be matched on targeted lesion area. Approximately 24 pediatric patients with moderate-to-severe AD will be enrolled to achieve 20 evaluable patients. These enrolled AD patients will receive Investigational medicinal product (IMP) dupilumab treatment during the study. Approximately 24 healthy volunteers will be enrolled to achieve 20 evaluable subjects. These enrolled healthy volunteers will not receive any IMP treatment during the study.

The maximum study duration per participant will be 32 weeks. The study will comprise of:

- Screening period 1 (Day -28 to Day -1): Participants will be evaluated according to
- Inclusion and exclusion criteria.
- Baseline visit (Week 0, Day 1): Participants who remain eligible will be enrolled.
- A 16-week treatment phase for AD patients and a 16-week observation period for healthy volunteers.
- A 12-week follow-up period.

Biomarker assessment based on lipidomics, proteomics, and transcriptomics in STS samples will be conducted at baseline (Week 0, Day 1), Week 2, Week 4, Week 8, Week 12, EoT at Week 16, and Follow-ups at Week 22 and Week 28.

1.2.1 Graphical study design



1.3 ANALYSIS ENDPOINTS

1.3.1 Biomarkers

The followings are the list of biomarkers to be explored:

- Lipidomics.
- Proteomics.

- Transcriptomics.
- FLG breakdown products.

1.3.2 Demographic and baseline characteristics

Baseline is defined as the last available and evaluable value before and closest to the first dose of the IMP: i.e., at Day 1 (Week 0), for patients and as the last available and evaluable value at Day 1(Week 0) for healthy volunteers. For TEWL data, the average of the values of 3 spots within each predefined skin area will be used as baseline, unless a large difference between the spots is observed, if so, the 3 spots will be analyzed separately.

Demographic characteristics

- Age.
- Weight.
- Sex.
- Race.
- Body Mass Index (BMI).

Disease characteristics

- TEWL:
  - TEWL before STS at each study visit
  - TEWL AUC, a composite measure before and after 5, 10, 15, and 20 STS
  - TEWL after each set of 5 STSs for up to 20 STSs
- Clinician-reported outcome assessments (ClinROs) and patient reported outcome assessments (PROs).
- Eczema area and severity index (EASI).
- Individual signs score (ISS).
- Worst itch numerical rating scale.
- Sleep disturbance numerical rating scale.
- Skin pain numerical rating scale.
- Patient oriented eczema measure (POEM).
- Children’s dermatology life quality index (CDLQI).

1.3.3 Efficacy endpoints

Primary efficacy endpoint is percent change from baseline in TEWL after 5 STS assessed on lesional skin at Week 15 in AD patients.

Secondary efficacy endpoints are:

- Change (percent and absolute) from baseline in TEWL after 20 STS (after 15 STS, after 10 STS) assessed on:
  - Lesional/non-lesional skin at Week 16 in AD patients
  - Normal skin at Week 16 in healthy volunteers
- Change (percent and absolute) from baseline in TEWL after 5 STS assessed on:
  - Non-lesional skin at Week 16 in AD patients
  - Normal skin at Week 16 in healthy volunteers
- Change (percent and absolute) in TEWL before STS (after STS) on:
  - Lesional/non-lesional skin in AD patients over time
  - Normal skin in healthy volunteers over time
- Change (percent and absolute) in TEWL AUC (a composite measure before and after 5, 10, 15, and 20 STS) for skin barrier function in:
  - Lesional/non-lesional skin in AD patients over time
  - Normal skin in healthy volunteer over time

Tertiary (Exploratory) efficacy endpoints are:

- Change (percent and absolute) in lipidomics parameters in lesional and non-lesional skin including the ratio of EOS CER and NS CER, and FLG breakdown products of UCA and PCA concentrations over time.
- Global characterization of protein-bound CER over time.
- Change (percent and absolute) in the expression of proteins associated with skin barrier function including keratin intermediate filaments, proteins associated with inflammatory response, and glycolysis and oxidative stress response proteins in STS protein extracts over time.
- Change (percent and absolute) in epidermal hypertrophy parameters including epidermal thickness (µm), superficial plexus depth (µm), blood vessel diameter (µm) and density (segments/mm2) measured by OCT over time.
- Change (percent and absolute) in stratum corneum lipid structure measured by FTIR spectroscopy in conjunction with tape-stripping over time.
- Changes in expression of genes associated with epidermal differentiation, barrier and lipid metabolism, and Type 2 inflammation over time.
- Change (percent and absolute) in CDLQI over time.
- Change (percent and absolute) in EASI over time.
- Change (percent and absolute) in ISS for target lesion over time.
- Change (percent and absolute) in POEM over time.

- Change (percent and absolute) in assessment of worst itch NRS over time.
- Change (percent and absolute) in sleep disturbance NRS over time.
- Change (percent and absolute) in skin pain NRS over time.
- Change (percent and absolute) in photograph outputs (eg, severity score) obtained from skin imaging over time.
- Correlation between baseline values of TEWL before and after STS, and TEWL AUC in lesional and non-lesional skin of AD patients with the following baseline measures:
  - EASI, Targeted lesion ISS, PRO measures (POEM, CDLQI, worst itch numerical rating scale (NRS), sleep disturbance NRS, and skin pain NRS), Lipidomics in STS (ratio of EOS CER to NS CER), FLG breakdown products of Urocanic Acid (UCA) and Pyroglutamic Acid (PCA) concentrations in STS, Image-derived severity score in targeted lesional skin, Key components of skin proteomics in STS (expression of proteins associated with skin barrier function), Quantified epidermal hypertrophy and structure changes in OCT and FTIR measures, Key components of gene expression from transcriptomics.
- Correlation between percent change from baseline in TEWL before and after STS, and TEWL AUC in lesional and non-lesional skin of AD patients at Week 8, Week 16 and Week 28 with corresponding change from baseline in the following measures:
  - EASI, Targeted lesion ISS, PRO measures (POEM, CDLQI, worst itch NRS, sleep disturbance NRS, and skin pain NRS), Lipidomics in STS (ratio of EOS CER to NS CER), FLG breakdown products of UCA and PCA concentrations in STS, Image-derived severity score in targeted lesional skin, Key components of skin proteomics in STS (expression of proteins associated with skin barrier function), Quantified epidermal hypertrophy and structure changes in OCT and Fourier-transform infrared (FTIR) measures, Key components of gene expression from transcriptomics.

1.4 PATIENT AND SAMPLE DISPOSITION

A detailed description of participant accountability including count of participants by analysis populations, screen failure participants with reasons for screen failure, participants who did not complete the study observation period along with the main reason for permanent treatment discontinuation, and participants who requested permanent treatment discontinuation, will be generated by cohorts.

All withdrawals from the study, taking place on or after study intervention intake, will be fully documented in the body of the clinical study report (CSR).

A listing of subjects with treatment discontinuation will be provided.

1.5 ANALYSIS POPULATIONS

Table 1 - Populations for analyses

Population	Description
Baseline biomarker	All participants who were enrolled in the study, who had a sample drawn for biomarker measurement and successfully analyzed at baseline.
Pharmacodynamic (PD)	All participants who were enrolled in the study, who had a sample drawn for biomarker measurement and successfully analyzed at baseline and at any post-baseline visit.

1.6 BIOMARKER ANALYSIS OBJECTIVES

The followings are the analysis objectives of biomarkers (lipidomics, proteomics, transcriptomics, FLG):

- Evaluate dupilumab treatment effect on skin biomarkers using STS samples in both predefined lesional and non-lesional skin in pediatric patients with moderate-to-severe AD during dupilumab treatment phase and follow-up period in reference to normal skin of healthy volunteers.
- Explore the association of skin barrier function measured by TEWL with disease impact assessed by standard AD severity assessments (EASI, and ISS), PRO (POEM, CDLQI, worst itch NRS, sleep disturbance NRS,)patient reported outcomes (PRO) (Patient Oriented Eczema Measure [POEM], Children Dermatology Life Questionnaire Index [CDLQI], worst itch Numerical Rating Scale [NRS], sleep disturbance NRS, skin pain NRS), standardized photos, the biomarker profiles of lipidomics, proteomics, FLG, and transcriptomics assessed in the STS samples, and the skin barrier hypertrophy and structure profiles assessed by OCT and FTIR.

1.7 BIOMARKER ASSESSMENTS AND SCHEDULE

1.7.1 Lipidomics, proteomics, and transcriptomics

Samples for skin lipidomic assessment are required and will be collected from all participants in this study.

Targeted skin lipidomics will be performed by means of a validated mass spectrometry method from skin tapes that have been collected during tape stripping at the TEWL assessment. The detailed assessment plan for lipidomics is listed in Section 4.1.1.

Protein extracts will be prepared from STS samples by LC-MS/MS. The detailed assessment plan for proteomics is listed in Section 4.1.2.

Skin tapes will be collected for transcriptomic analysis. The detailed assessment plan for transcriptomics is listed in Section 4.1.3.

For both AD pediatric patients and healthy volunteers, samples are collected at Week 0, 2, 4, 8, 12, 16, and 22. All STS will be conducted on the predefined spots. TEWL will be assessed before STS and after 5, 10, 15, and 20 STS in lesional and non-lesional skin (AD patients) and normal skin (healthy volunteer). Weeks and spot on which STS is conducted are

- At Weeks 0, 8, 16, and 28, STS conducted on the first spot.
- At Week 2, STS conducted on the second spot.
- At Weeks 4, 12, 22, STS conducted on the third spot.

All skin tape strips should be collected and stored.

1.7.2 Optical coherence tomography (OCT)

OCT is a non-invasive imaging modality conceptually like ultrasound but uses near-infrared radiation rather than sound. The detailed assessment plan for OCT is in Section 4.1.4. For both AD pediatric patients and healthy volunteers, samples for OCT are generated at Week 0, 1, 2, 3, 4, 6, 8, 12, 16, and 28. OCT will be done before STS/TEWL.

1.7.3 Attenuated total reflectance (ATR) FTIR spectroscopy

ATR FTIR spectroscopy is a form of molecular spectroscopy useful for the analysis of surfaces, including the lipid, water, and carboxylate content of the skin barrier and for analyzing lipid arrangement/structure. The detailed assessment plan for FTIR spectroscopy is in Section 4.1.5. For both AD pediatric patients and healthy volunteers, samples for FTIR spectroscopy are generated at Week 0, 8, 16 and 28. FTIR will be done before and during STS.

1.7.4 Optional samples for biomarker research

The following optional serum and plasma samples for biomarker research should be collected from participants in the study, where possible. The detailed collection plan for optional samples is in Section 4.1.6. For both AD pediatric patients and healthy volunteers, optional blood samples are collected at Week 0 and 16.

2 STATISTICAL AND ANALYTICAL PROCEDURES

2.1 PREPROCESSING AND QUALITY CONTROL

2.1.1 Data normalization

2.1.1.1 Lipidomics

STS samples processing for lipid extraction will be done as described in Section 4.1.1. Lipidomic data will be provided as absolute concentration in pmol/mg protein or in ug/mg protein requiring no normalization step. Values below limit of quantification (LLOQ), missing values, and outliers will be treated as described in Section 2.1.2.

2.1.1.2 Proteomics

There is no normalization step performed after sample analyses as specified by National Jewish Health (NJH). Polar metabolite samples will be normalized based on sample protein values before running the assay then signal intensities are expressed relative to healthy volunteers' group.

2.1.1.3 Transcriptomics

Raw sequencing data (fastq files) will be taken through successive steps to generate expression measures at gene level (counts): quality control, adapter trimming and filtering, mapping of reads to the human reference genome or pseudo-alignment to the transcriptome, quantification. These steps will be performed by NJH. The raw and the preprocessed transcriptomic data will be sent to Sanofi.

Samples filtering

Total RNA from the skin tapes will be normalized for cDNA generation and preparation of sequencing libraries. Both the cDNA and sequencing libraries generated will be quantified and qualified using the Bioanalyzer High-Sensitivity DNA chip. Samples that will yield low quality cDNA and sequencing libraries will not be sequenced.

Gene filtering

Genes having very low counts across all libraries will be filtered out. For this purpose, the count per million (cpm) values will be computed. Counts per million are counts scaled by the number of sequenced fragments in a sample (library size) times one million. Genes with less than 1 cpm in more than half of the samples will be filtered out. This filtering rule will be applied separately for count data matrices in HV and AD cohorts. Genes that successfully pass the filter in at least one data set will be kept. The specific threshold of 1 cpm may be adjusted depending on count distribution in the dataset.

Normalization and transformation of counts data

Normalization is a process designed to remove systematic technical differences between samples to ensure that technical bias has minimal impact on the results. One can envisage within-sample and between-samples normalization. In this study since we are interested in differential expression between experimental conditions rather than comparing expression of different genes in a sample, only a between-sample normalization will be performed. Using the DESeq2 R package, variance stabilizing transformation (VST) will be applied for analyses other than differential testing, for example clustering of samples or other machine learning applications. For differential testing the DESeq function will be applied to raw counts.

Included in VST, a log2 transformation is performed after addition of a small offset to avoid taking the log of zero. The quality of the normalization will be assessed by plotting the densities of variance-stabilized counts for the different samples. Since it is assumed that most of the genes are unaffected by the experimental conditions, a successful normalization should lead to overlapping densities. The mean-variance association will also be investigated before and after normalization. Samples boxplots will be plotted before and after normalization.

2.1.2 Missing data, outlier, and batch effect assessment

2.1.2.1 Data imputation

For each biomarker, the number of values below LLOQ will be computed by group (lesional, non-lesional and healthy skin). Values below LLOQ will be imputed according to the following rules:

- If for a given biomarker, less than 25% of all samples have values below LLOQ, all values below LLOQ will be imputed by a random number draw from a uniform distribution in [0, LLOQ]. The minimum computed across all samples will be used if the LLOQ is not provided.
- If more than 50% of samples have values below LLOQ, the biomarker will be discarded from statistical analyses
- If the percentage of values below LLOQ is between 25% and 50%, it will be checked if the values below LLOQ occur mostly in a specific study group. The Fisher test will be used to check the dependency between the occurrence of below LLOQ and group.
- The biomarker will be discarded from statistical analysis if the LLOQ occurrence is group independent. Values below LLOQ will be imputed by the minimum computed across all samples, otherwise.

Missing values not due to measurements being below LLOQ maybe be imputed by the mean value computed across all samples, as appropriate.

2.1.2.2 Outlier detection

Biomarker will be investigated for outliers regardless of group or visit. A value will be declared outlier if it falls out of the interval [Q1-3*IQR, Q3+3*IQR], where Q1 denotes the first quartile,

Q3 the third quartile and IQR the interquartile range. Outliers will be imputed by Q1-3*IQR and Q3+3*IQR, respectively. Samples will be investigated for possible contamination that may lead to sample being outlier. For this purpose, samples boxplots, principal component analyses (PCA) hierarchical cluster dendrogram will used. Outlier samples will be scrutinized and maybe discarded from the statistical analyses.

The distribution of biomarkers will be assessed by histogram and QQ-plot. The values may be log-transformed to fit a Gaussian distribution prior to some statistical analyses. The standardization may be performed for multivariate analyses.

2.1.2.3 Batch effect assessment

Samples maybe analyzed or transferred to NJH in different batches. PCA will be computed and a visual inspection of the projection of the samples on the first principal components will enable a first view of the potential batch's effect. The batch effect will be assessed using a linear model estimating a batch effect for each biomarker. This model will include the batch effect, the visit, the health status and interaction between visit and health status. The relevance of the correction for batch effect will be assessed by the p-values of the batch effect.

2.2 DESCRIPTIVE STATISTICS AND VISUALIZATION

2.2.1 Descriptive statistics

For lipidomics and proteomics, the summary statistics (min, Q1, mean, median, Q3, max, SD and SEM) will be computed by group (lesional, non-lesional and healthy skin) and by visit. The summary statistics will be computed by age categories as appropriate.

For OCT, raw data and change from baseline (absolute and percent changes) will be summarized with descriptive statistics (mean, median, SD, min, max) by groups and time points. For follow-up period data, change from Week16 (or last visit prior to use of rescue treatment) will also be reported.

For FTIR, profiles over study days will be generated for individual values and group means for raw and absolute change from baseline and Week 16.

2.2.2 Data Visualization

Spaghetti plots

The biomarker profiles will be represented graphically using line plots (one line per subject). The lines will be colored by group. Additional plots will be generated by age categories as appropriate.

Profile plots

For each biomarker, the mean ±SEM will be represented graphically by group and by visit for the absolute values and percent change from baseline. Additional plots will be generated by age categories as appropriate.

Boxplots

Longitudinal boxplots for absolute values will be provided by group and by visit. Additional plots will be generated by age categories as appropriate.

Heatmaps

A heatmap representation will be provided for each biomarker data matrix highlighting the 2 study groups (AD lesional, non-lesional and healthy volunteers) and different study visits. The age categories will be represented as appropriate. Biomarkers will be clustered according to an unsupervised clustering on Pearson correlation with Ward's agglomeration technique and a dendrogram will be added to the map. Each biomarker will be centered to the mean and reduced to the unit standard deviation. To improve the readability a separate heatmap will be generated with only the baseline biomarker levels and different groups will be highlighted.

Principal component plot

PCA plot will be conducted on log-transformed and scaled data for each biomarker data matrix. The data will be visualized on a 2-dimensional plot of first principal component (PC1) versus PC2 highlighting potential group differences and change during the study. Confidence ellipses will be added as appropriate. A scatterplot with only the baseline samples may be generated to ease the readability. Depending on the percentage of variability explained on the PC1, a second technique for deriving a composite biomarker will be applied.

Median polish

A two-way decomposition (row effect + column effect + overall effect) of the data matrix with each biomarker type will be performed using the Tukey's median polish procedure. The biomarker will be on columns and the samples on rows. The row effect will be extracted as composite biomarker.

Composite biomarker

PC1 and the row effect from the median polish algorithm will be analyzed as composite biomarkers analog to single biomarker. This includes the correlation with TEWL values, the difference between the study groups and the change during the study.

MDS and t-SNE

Multidimensional scaling (MDS) and t-distributed stochastic neighbour embedding (t-SNE) will be applied for further visualization of the data on 2 dimensions highlighting the study groups.

2.3 CLUSTER ANALYSIS

2.3.1 Patients clustering

A hierarchical clustering of AD patients will be performed on baseline biomarker data using Euclidean distance and Ward agglomeration technique. This will enable the identification of

subgroups of homogeneous patients. The identified patients' subgroups will be further investigated regarding the disease activity, disease progression and the response to the treatment (eg, longitudinal plots for mean ±SEM of TEWL by cluster, ANOVA model for TEWL at baseline with cluster as independent predictor), cross table and Chi2-test for clusters versus categorical parameters.

For clusters definition, an elbow approach (silhouette or inertia by number of clusters) will be used for determining the optimal number of subgroups of patients using at least two different algorithms to ensure robustness (K-means and agglomerative clustering).

2.3.2 Biomarker clustering at baseline

Clustering of biomarker will be performed on biomarker expressions at baseline (using correlation-based distance metric). Biomarker clusters may be suggested to pathway analysis as appropriate. Heatmaps will be generated considering the biomarker clusters.

2.4 CORRELATION ANALYSIS

2.4.1 Correlation between biomarker and TEWL

In addition to the correlation analysis between each the composite biomarkers and the TEWL, the association between each single biomarker (lipidomics, proteomics, transcriptomics) and TEWL will be investigated at baseline in a linear model adjusted for the study group and age. TEWL at baseline will be evaluated as TEWL. Biomarker will be ranked according to their importance in explaining TEWL using RandomForest for regression. The Correlation between change in biomarker and change in TEWL, the correlation between baseline biomarker and change in TEWL will be investigated.

2.4.2 Correlation between biomarker and clinical and disease parameters

Analog to TEWL, the correlation analyses will be done between each biomarker and the relevant clinical and disease parameters. Analysis of Variance (ANOVA) will be performed for categorical variables.

2.4.3 Intra-class Correlation Coefficient

Stability of repeated measurements will be assessed for each biomarker (lipids and proteins). For each biomarker, its measurement will be modeled as:

$$Yij = \mu + \alpha i + \beta j + \epsilon ij$$

for subject i and visit j of the subject i in a single arm (e.g., control arm) (1). Here we assume fixed visit effect βj but random subject effect αi and random error ϵij. If we assume $\alpha i \sim N(0, \sigma^2$ subject) and $\epsilon ij \sim N(0, \sigma^2$ error) and independence between them, the correlation of

measurements within a subject (also known as intra-class correlation coefficient or ICC) can be derived as

$$\mathrm{ICC} = \frac{\sigma^2_{\mathrm{subject}}}{\sigma^2_{\mathrm{total}}}$$

where σ^2 total = σ^2 subject + σ^2 error. Large ICC value indicates strong stability of repeated measurements of the biomarker of a subject.

2.5 IDENTIFICATION OF BIOMARKERS

2.5.1 Differentially expressed biomarkers

The biomarker potentially linked to the disease status (lesional, non-lesion or healthy skin) and the biomarker changing with the treatment will be identified using a linear mixed model. For each biomarker, the following model will be used to test for any interaction between group and visit for group i, visit j and subject k:

$y_{ijk} = \mu + \mathrm{group}_i + \mathrm{Visit}_j + \mathrm{group}_i\,\mathrm{Visit}_{ij} + \mathrm{Subj}_k + \varepsilon_{ijk}$

with group, visit and interaction between group and visit considered as fixed effects and subject as random effect. Baseline will be used as reference visit and healthy skin as reference group. The data will be log-transformed. A contrast matrix will be defined for extracted the comparison of interest such as group difference at baseline or at Week 16. The fold-changes will be derived, and volcano-plots will be provided as appropriate. The intersection between biomarker differentially expressed at baseline between lesional vs. non-lesional, lesional vs healthy skin and non-lesional vs. healthy skin will be represented graphically using Venn diagrams. The proportion of biomarker changing significantly from baseline to Week 16 will be derived for each study group. The direction of the changes will be highlighted.

2.5.2 Discriminative biomarkers

A multivariate predictive panel will be identified used Elastic-net penalized multinomial logistic regression and the predictive value of the panel will be estimated as 10-fold cross-validated as accuracy. Biomarkers will be ranked according to their importance in discriminating between Lesional, non-lesional and healthy skin groups as provided by RandomForest.

2.5.3 Predictive biomarkers

In case of any binary response defined a clinical endpoint, the predictive characteristics for each biomarker will be evaluated using a logistic regression model with the baseline biomarker and the response variable as binary outcome at Week 16. For the binary response at Week 16, a multivariate predictive panel will be identified used Elastic-net penalized logistic regression and the predictive value will be estimated as 10-fold cross-validated Area Under the ROC curve. These analyses will be computed for AD patients (lesional skin). Biomarkers will be ranked according to their importance in discriminating between responders and non-responders as

provided by RandomForest. A linear model will be used for change from baseline in clinical responses as continue parameters and baseline biomarker level.

2.6 MULTIPLICITY ISSUE

All analyses will be exploratory. No formal hypothesis testing will be conducted.

2.7 BIOINFORMATICS METHODS AND RESULTS INTERPRETATION

2.7.1 Analyses based on biomarkers regulated by treatment from longitudinal analysis of RNA-seq

- Downstream effect analyses (GO terms, biological processes and diseases) of regulated gene sets in Ingenuity Pathway Analysis (IPA) to get mechanistic insights into disease-affected biological processes.
- Pathway Enrichment Analyses (Over-representation analysis, t-test based) aimed to the identification of curated biological pathways involved in regulated gene patterns (MetaCore [Clarivate Analytics], IPA Core Analyses [Qiagen])

2.7.2 Functional protein association networks

For identifying proteins sets, functional protein association networks will be built and interpreted using the String database.

3 SOFTWARE DOCUMENTATION

All analyses of clinical data will be performed under the responsibility of Sanofi Biostatistics Department using R software Version 4.0.0 (2).

4 SUPPLEMENTARY INFORMATION

4.1 DETAILED BIOMARKER ASSESSMENTS

4.1.1 Lipidomics

Skin lipidomics samples may be stored for a maximum of 5 years following the last participant's last visit in the study at a facility selected by the Sponsor. Detailed guidelines and requirements for samples preparation, handling of skin tapes after stripping, and storage and shipment of such skin tapes will be provided in the study reference manual.

4.1.1.1 Filaggrin breakdown products

Filaggrin is a structural protein that plays an important role in controlling water retention in the skin, and the hygroscopic properties of FLG breakdown products have an important role as natural moisturizing factor ingredients. Filaggrin deficiency leads to impaired lipid profile and altered acidification pathways. Lipid abnormalities have been reported in patients with FLG mutations. Filaggrin breakdown products, cis/trans-urocanic acid (total UCA) and pyrrolidone carboxylic acid (PCA), also known as pyroglutamic acid, will be quantified via a liquid chromatography electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS) approach in the same tape strips that are used for lipidomics assessment.

4.1.1.2 Analysis of stratum corneum lipids

It has been reported that lesional and non-lesional skin of AD patients has decreased EOS CER and other CER with very long-chain fatty acids (C22-C30) and increased short-chain fatty acids NS CER (C16-C22). Highly hydrophobic omega-esterified acid sphingosine CER (EOS CER) and NS CER will be analyzed as follows.

STS processing for lipid extraction

STS will be processed through a Bligh and Dyer procedure. The bottom chloroform layer from the skin tape extraction will be used for lipid analyses while upper water-methanol phase will be used for polar component analyses by the Liquid Chromatography with tandem Mass Spectrometry (LC-MS/MS). All mass spectrometric data will be normalized to the total amount of hydrolyzed protein determined. EOS CER and NS CER will be identified and quantified using a targeted ultra-high performance liquid chromatography electrospray ionization tandem mass spectrometry (UHPLC-ESI-MS/MS).

4.1.1.3 Quantification of protein bound ceramides

Protein bound ceramides are formed from EOS CER and represent a portion of EOS CER without linoleic acid, thus having a terminal hydroxyl group N-linked fatty acid. Protein-bound ceramides

are released from proteins during protein hydrolysis. After protein determination in hydrolysates, freed omega-OH CER will be extracted and omega hydroxy fatty acid containing CER will be detected by the UHPLC-ESI-MS/MS. Values obtained will be normalized against protein content in the samples.

4.1.2 Proteomics

The following 3 major functional groups of proteins associated with skin barrier function will be examined:

- Keratin intermediate filaments.
- Proteins associated with inflammatory response (S100 proteins, alarmins, protease inhibitors).
- Glycolysis and oxidative stress response proteins (glycolytic enzymes, oxidative stress response enzymes).

Skin proteomics samples may be stored for a maximum of 5 years following the last participant's last visit in the study at a facility selected by the Sponsor. Detailed sample preparation for proteomic analysis will be provided in a separate study reference manual.

4.1.3 Transcriptomics

Skin tapes will be collected and stabilized in RLT buffer with DTT and stored frozen at -80°C for transcriptomic analysis. Skin tape transcriptome samples will be shipped to National Jewish Health, where total RNA will be isolated and RNA sequencing will be performed using qualified methods to support whole transcriptome analysis and data generation. Skin transcriptomics samples and any RNA isolated from the samples may be stored for a maximum of 5 years following the last participant's last visit in the study at a facility selected by the Sponsor. Transcriptome sample preparation and analysis details will be provided in a separate study reference manual.

4.1.4 Optical coherence tomography

Structural and Angiographic OCT scans using the Vivosight OCT machine

OCT is a non-invasive imaging modality conceptually similar to ultrasound but uses near-infrared radiation rather than sound. It has a 2 to 10 μ depth resolution compared with 100 to 1 000 μ typical for clinical ultrasound; and 1 to 2 mm imaging depth vs. 10 to 100 mm for clinical ultrasound. It is thus ideal for imaging the surface layers of accessible tissues such as the skin. It is attracting interest throughout the medical community as a scanning tool and for diagnosis of illnesses such as epithelial cancer, connective tissue disorders, and atherosclerosis. The Vivosight OCT device will be used to provide information for both structural and angiographic OCT outcomes.

1. *Structural* OCT: With a depth focus of 1.0 mm and optical resolution of 7.5 × 5.0 μm the Vivosight provides structural images of the skin comparable to histology sections. From

these images the thickness of the epidermis, the suprapapillary epidermis, and the papillary region can be extrapolated. Epidermal thickness has been identified using histology, as a useful biomarker of disease activity and treatment effects in AD.

2. *Angiographic* OCT: Structural OCT performed at high image frame rates can map areas of temporal decorrelation due to moving red blood cells. Based on this, demonstrated that the Vivosight can acquire high-quality maps of the superficial vasculature in living participants and is an ideal tool to monitor and quantify erythema, hyperplasia, and TCS-induced tissue remodeling.

The OCT assessments will be performed before STS/TEWL assessments at each designated visit from Week 0 to Week 28 (see Section 1.3, Section 1.4 and Figure 6 of Clinical Trial Protocol). Two scans will be taken from each target area (separate, adjacent, sites for each scan) at each designated visit. All images will be checked for quality upon collection, and images with visible imaging artifacts will be re-captured. Any part of the device in contact with human skin (the stand-off) will be decontaminated between uses.

Details of OCT operation and OCT data analysis will be provided in a separate study reference manual.

4.1.5 ATR FTIR spectroscopy

ATR FTIR spectroscopy is a form of molecular spectroscopy useful for the analysis of surfaces, including the lipid, water, and carboxylate content of the skin barrier and for analyzing lipid arrangement/structure.

In this study the Agilent 4300 handheld FTIR spectrometer will be used. Measurements will be performed by gently but firmly placing the probe in contact with the skin. Duplicate measurements will be collected from each target site at the visits indicated in the schedule of events. The quality of spectra will be assessed at the point of collection, and any that do not meet the established quality parameters (signal to noise ratio) will be discarded. After use in each participant the ATR (probe head) will be decontaminated with 70% alcohol wipe.

The FTIR assessments will be performed before and during the STS assessments at each designated visit (see Section 1.3, Section 1.4 and Figure 6 of Clinical Trial Protocol).

Detailed measurement and data collection from FTIR will be provided in a separate study manual.

4.1.6 Optional samples for biomarker research

The following optional serum and plasma samples for biomarker research should be collected from participants in the study, where possible:

- Venous blood samples (not more than 12 mL in total) will be collected at baseline and EoT/end of observation at Week 16 to obtain serum and plasma to be stored for possible future research related to response, disease activity, safety, the Type 2 inflammation pathway, and for assessing the effects of the study intervention on modulation of

IL-4 receptor and on atopic disease processes, as well as to study biomarkers that may have predictive utility for response to dupilumab treatment.

- The results of exploratory research testing will not be included in this Clinical Study Report (CSR).

Details on collection, storage, and shipment of these samples will be provided in the study reference manual. Plasma and serum samples may be stored for a maximum of 5 years following the last participant's last visit in the study at a facility selected by the Sponsor.

5 REFERENCES

1. Diggle P, Heagerty P, Liang KY, Zeger S. Analysis of Longitudinal Data, Oxford University Press; 2nd edition 2013:1-400.

2. Core Team. R: A language and environment for statistical computing. Foundation for Statistical Computing (2013).

Signature Page for VV-CLIN-0640804 v1.0
lps16764-16-1-9-sap-biomarker

Approve & eSign	[redacted]
Approve & eSign	[redacted]
Approve & eSign	[redacted]