Sanofi group

NCT04033367

Dupilumab - LPS15497

A randomized double-blind placebo-controlled study evaluating the effect of dupilumab on sleep in adult patients with moderate to severe atopic dermatitis

30JUL2021

Final Statistical Analysis Plan

Version 2.0

Prepared by:

[redacted]

[redacted]

[redacted]

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACQ-5	Asthma Control Questionnaire
AD	Atopic Dermatitis
AE	Adverse Event
AESI	Adverse Event of Special Interest
ANAM	Automated Neuropsychological Assessment Metrics
ANCOVA	Analysis of Covariance
AR-VAS	Allergic Rhinitis-Visual Analog Scale
ATC	Anatomical Therapeutic Chemical
BCP	Business Continuity Plan
BSA	Body Surface Area
CI	Confidence Interval
CLINRO	Clinician-Reported Outcome
DPB	Diastolic Blood Pressure
DLQI	Dermatology Life Quality Index
EASI	Eczema Area Severity Index
eCRF	Electronic Case Report Form

ESS	Epworth Sleepiness Scale
HADS	Hospital Anxiety and Depression Scale
HLT	High Level Term
HLGT	High Level Group Term
HR	Heart Rate
IcEv	Intercurrent Event
ICH	International Conference on Harmonisation
ICF	Informed Consent Form
IGA	Investigator's Global Assessment
IMP	Investigational Medicinal Product
ITT	intent-to-treat
LLT	Lower Level Term
MAR	Missing At Random
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple Imputation
MMRM	Mixed Model for Repeated Measures
NIMP	Non-Investigational Medicinal Product
NRS	Numerical Rating Scale
OLE	Open-Label Extension
PCSA	Potentially Clinically Significant Abnormalities
POEM	Patient Oriented Eczema Measure
PRO	Patient-Reported Outcome
PROMIS	Patient-Reported Outcomes Measurement Information System
PSG	Polysomnography
PT	Preferred Term
PVT	Psychomotor Vigilance Test
Q1	First Quartile
Q2W	Every 2 Weeks
Q3	Third Quartile
ROW	Rest of World
QoL	Quality of Life
SAE	Serious Adverse Event
SBP	Systolic Blood Pressure
SCORAD	SCORing Atopic Dermatitis
SD	Standard Deviation
SE	Sleep Efficiency
SOC	System Organ Class
SOL	Sleep Onset Latency
TEAE	Treatment-Emergent Adverse Events
TST	Total Sleep Time
VAS	Visual Analog Scale
WASO	Wake After Sleep Onset
WNRG	Washington Neuropsychology Research Group, LLC
WOCF	Worst Observation Carried Forward
WPAI-AD	Work Productivity and Activity Impairment-Atopic Dermatitis
WHO-DD	World Health Organization-Drug Dictionary

Document History – Changes compared to previous version of SAP

Version	Date	Changes
Final 1.0	04MAY2021	Initiation
Final 2.0	30JUL2021	• Subgroup analysis based on Disease Duration has been removed as disease onset is not collected • Baseline definition for daily assessments has been updated to include one patient who completed baseline assessment from -13 to -7 instead of -6 to Day 0. • For the responder analysis on EASI, the analysis model has been changed from GENMOD to GLIMMIX. • PROMIS missing item score imputation rule has been updated • Added some details for retest approach of Psychomotor Vigilance Test. • Sleep diary illogical data handling rules has been added as an SAP addendum.

1. Introduction

This document provides the detailed statistical methodology for the analysis of data from the Sanofi Group study LPS15497. The table, listing and figure shells supporting the statistical analysis plan (SAP) can be found in a separate SAP shell document.

The analyses described herein are based on the clinical trial protocol (Amendment 2), Version 1.0, dated on 7th April 2020. Any changes or revisions to the planned analysis described in this document will be made prior to database lock.

2. Study Objectives, Endpoints and Estimands

The study objectives and corresponding endpoints are shown in the Table 3.1.1.

Table 3.1.1- Objectives and endpoints

Objectives	Endpoints
Primary	
• To evaluate the effect of dupilumab on sleep quality in adult patients with moderate to severe atopic dermatitis (AD)	• Percentage change from baseline to week 12 in sleep quality numerical rating scale (NRS)
Secondary	
• To evaluate the effect of dupilumab on objective and subjective quantitative sleep parameters, AD-related outcomes, and the daytime	• Change from baseline to week 12 in sleep efficiency (SE) based on actigraphy data • Change from baseline to week 12 in total sleep time (TST) based on actigraphy data • Change from baseline to week 12 in wake after sleep onset (WASO) based on actigraphy data

consequences of sleep deprivation	• Change from baseline to week 12 in sleep onset latency (SOL) based on actigraphy data • Percent change from baseline to week 12 in Pruritus NRS • Change from baseline to week 12 in SCORing Atopic Dermatitis (SCORAD) total score • Change from baseline to week 12 in SCORAD sleep Visual Analog Scale (VAS) • EASI50 (reduction of Eczema Area Severity Index [EASI] score by ≥50% from baseline) at week 12 • EASI75 (reduction of EASI score by ≥75% from baseline) at week 12 • EASI90 (reduction of EASI score by ≥90% from baseline) at week 12. • Change from baseline to week 12 in Patient Oriented Eczema Measure (POEM) total score • Change from baseline to week 12 in Dermatology Life Quality Index (DLQI) total score • Change from baseline to week 12 in Patient-Reported Outcomes Measurement Information System (PROMIS) Sleep Related Impairment SF8a Total Score
• To continue to assess safety and tolerability of dupilumab throughout the study	• Incidence of adverse events through week 12 • Incidence of adverse events through week 24
Other Exploratory	
• To further evaluate the effect of dupilumab on objective and subjective sleep parameters, AD-related outcomes and daytime consequences of sleep deprivation	• Percent change from baseline in sleep quality NRS to week 1, 2, 4 and 8 • Percent change from baseline in pruritus NRS to week 1, 2, 4 and 8 • Change from baseline in SE based on actigraphy to week 1, 2, 4 and 8 • Change from baseline in WASO based on actigraphy to week 1, 2, 4 and 8 • Change from baseline in TST based on actigraphy to week 1, 2, 4 and 8 • Change from baseline in SOL based on actigraphy to week 1, 2, 4 and 8 • Change from baseline in SCORAD total score and sleep VAS to week 1, 2, 4 and 8 • EASI50, EASI75 and EASI90 at weeks 1, 2, 4 and 8 • Change from baseline to week 12 in polysomnography (PSG) measures of SE, TST, WASO and SOL • Change from baseline to weeks 1, 2, 4 and 8 in PROMIS Sleep Related Impairment SF8a total score • Change from baseline to weeks 1, 2, 4, 8 and 12 in skin pain NRS

	• Change from baseline to weeks 1, 2, 4, 8 and 12 in skin sensitivity to touch NRS • Change from baseline to weeks 1, 2, 4, 8 and 12 in skin burn NRS • Change from baseline in SOL, TST, number of awakenings, WASO, SE and how rested the participant felt when getting up at week 1, 2, 4, 8 and 12 based on sleep diary • Change from baseline to week 4 and 12 in Epworth Sleepiness Scale (ESS) • Change from baseline to week 12 in psychomotor vigilance test (PVT) • Change from baseline to week 12 in Running Memory, Mathematical Processing and Procedural Reaction Time scores based on the neurocognitive test (Automated Neuropsychological Assessment metrics [ANAM]) • Work productivity and activity impairment – Atopic Dermatitis version (WPAI-AD) associated questions change from baseline to weeks 4, 8 and 12 • Missed school/workdays from baseline to week 12
• To evaluate the relationship between sleep disturbance and other objective and subjective parameters associated with AD	• Relationship between variables relating to sleep disturbances (eg, TST, WASO) and objective and subjective parameters associated with AD (eg, EASI, SCORAD, pruritus NRS, DLQI, skin pain NRS, skin burn NRS and sensitivity to touch NRS)
• To evaluate the effect of dupilumab in comorbid conditions	• Change from baseline to week 12 in ACQ-5 among those who reported asthma at baseline • Change from baseline to week 12 Allergic Rhinitis-Visual Analog Scale (AR-VAS) among those who reported allergic rhinitis at baseline • Change from baseline to weeks 4 and 12 in Hospital Anxiety and Depression Scale (HADS)-anxiety and HADS-depression scores
• To evaluate the effect of dupilumab on individual quantitative and non-sleep related endpoints associated with AD at Week 24	• Change or percent change (as used in analyses of week 12 data) from baseline to week 24 for primary and secondary endpoints • Change or percent change (as used in analyses of week 12 data) from week 12 to week 24 for primary and secondary endpoints

Table 3.1.2 Summary of primary estimands for main objectives

Endpoint Category (estimand)	Estimands			
	Endpoint(s)[a]	**Population**	**Intercurrent event(s) handling strategy**	**Population-level summary (Analysis and missing data handling)**
Primary objective: To evaluate the effect of dupilumab on sleep quality in adult patients with moderate to severe atopic dermatitis (AD).				
Primary endpoint	Percentage change from baseline to Week 12 in sleep quality NRS	ITT	• Discontinuation or interruption of study treatment due to Covid-19 pandemic, off-treatment data will be set as missing (hypothetical strategy). The assessments after patients resume treatment will be used. • Discontinuation or interruption of study treatment not due to Covid-19 pandemic, all data collected following schedule after the treatment discontinuation will be used in the analysis (treatment policy strategy). • Taking rescue medications [b] prior to Week 12: data will be set to missing values after the medication usage (hypothetical strategy) • Taking prohibited medications [b] that affect sleep: All sleep related assessments (sleep NRS, SCORing subjective sleep loss scores, PROMIS sleep disturbance short form 8a individual question scores, ESS, actigraphy data) on that corresponding night will be set to missing (hypothetical strategy) • Taking prohibited medications [b] that affect AD, data will be set to missing values after the medication usage (hypothetical strategy)	A mixed effect model for repeated measures (MMRM) under the missing at random (MAR) framework will be used. The MMRM will include treatment, baseline value, randomization stratum, visits (up to week 12), treatment by visit interaction, and baseline sleep quality NRS by visit interaction terms all as fixed effects. Randomization stratum, treatment and visits will be considered as categorical parameters. No explicit imputation will be done for missing data.
Secondary objective: To evaluate the effect of dupilumab on objective and subjective quantitative sleep parameters, AD related outcomes, and daytime consequences of sleep deprivation.				

Endpoint Category (estimand)	Estimands			
	Endpoint(s)[a]	**Population**	**Intercurrent event(s) handling strategy**	**Population-level summary (Analysis and missing data handling)**
Secondary endpoint (treatment policy strategy/hypothetical strategy)	Change or percent change from baseline in Sleep and AD related outcomes	ITT	Same as the primary endpoint	The same MMRM with the baseline values corresponding to the outcome variables.

a Additional key secondary endpoints are not included in this table but would be handled with a similar strategy as the endpoint type (ie continuous, binary, time-to-event) at other weeks. The key secondary endpoints are listed in Section 8.2.23.

b Selected prohibited medications and/or rescue medications are listed in Section 15.11.

3. Investigational Plan

3.1. Overall Study Design and Plan

This is a Phase IV, prospective, 2:1 randomized, multinational, multicenter study, with a 12-week double-blind placebo-controlled period, followed by a 12-week open label extension to evaluate the effect of dupilumab on sleep in adult patients with moderate to severe AD.

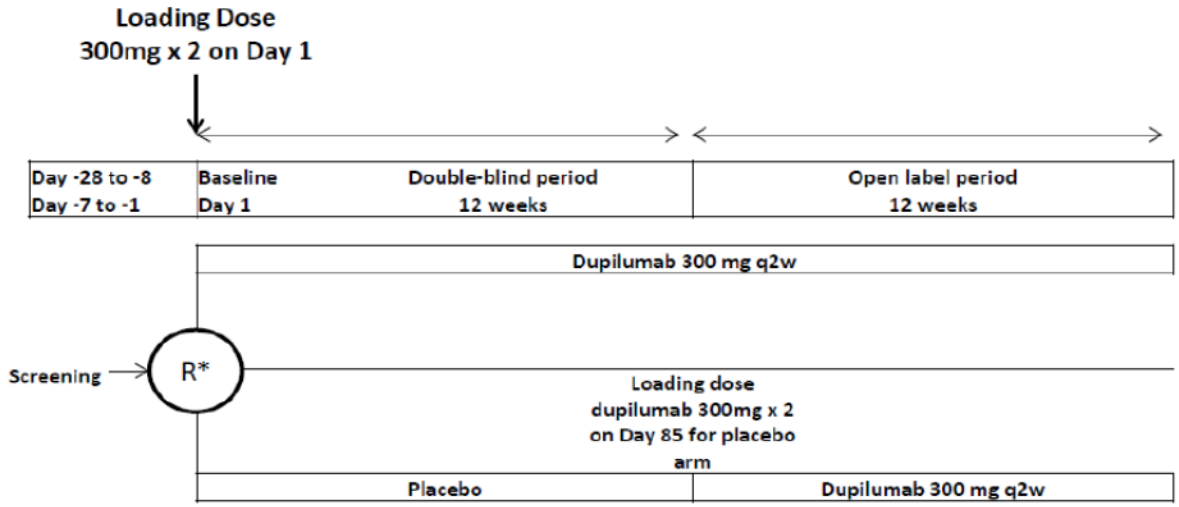
Approximately 201 participants will be randomly assigned to study intervention in a 2:1 dupilumab to placebo ratio from multiple sites globally such that at least 150 evaluable participants (100 in the active group and 50 in placebo) complete the double-blind portion (to week 12) of the study.

The study will comprise of:

- Screening Period 1 (Day -28 to -8): Patients will be evaluated according to inclusion and exclusion criteria
- Screening Period 2 (Day -7 to -1; pre-baseline assessments): eligible patients will complete the daily sleep quality, daily pruritus, skin pain, skin sensitivity, and touch skin burn, sleep diary, and will wear an actigraph at night-time during the 7 days immediately preceding the Baseline visit (Day 1). If preferred by the patient and the investigator for logistical reasons, screening can be done during a single on-site visit. In that case, all procedures planned at screening 1 and screening 2 should be done during this single screening visit. Importantly, this single screening visit will have to be done at least 7 days before the baseline visit, as patients should wear the actigraph and complete the diary for 7 days before baseline.
- Baseline (Day 1): Patients who remain eligible will be randomized
- 12-week double-blind placebo-controlled period: Patients will be randomized to receive dupilumab or placebo
- 12-week open-label extension period: following completion of double-blind period, patients randomized to dupilumab will continue to receive active treatment for open-label period, whereas patients randomized to placebo will receive active treatment with dupilumab.

The maximum study duration per participant will be 28 weeks. A patient is considered to have complete the study if he/she has completed all phases of the study including the last scheduled visit shown in the Schedules of Activities (see Appendix 15.2).

Figure 3-1 Study Plan



Abbreviations: q2w=every 2 weeks

PSG sub-study: of the 201 participants, approximately 30 patients total (approximately 20 dupilumab and 10 placebo patients) will participate in a PSG sub-study at selected US sites. Each patient who participates this sub-study will have a total of 3 overnight PSGs: at baseline timepoint, 2 (preferably) consecutive overnight PSGs to control for "first night effect" (night of Day -2) and collect baseline data (night of Day -1; which will serve as baseline assessments); then 1 overnight PSG at Week 12. Selection of patients for the PSG sub-study will be based on their proximity to participating sleep centers and availability for overnight stay.

3.2. Study Endpoints

The study objectives and corresponding endpoints are shown in the Table 3.1.1.

3.3. Treatments

The dose selected in this study is the dose approved for the treatment of moderate to severe AD for Dupixent in US, EU, and all other countries that this study will be conducted.

Patients will receive either;

- Dupilumab 600 mg (2 x 300 mg dupilumab injections) on Day 1 (ie, the loading dose), followed by 1 dupilumab injection of 300 mg every two weeks (Q2W) until week 10, then 1 dupilumab + 1 placebo injection at Week 12 (to protect the blind while patients in placebo arm receive their loading dose), followed by 1 300 mg dupilumab injection Q2W with last injection at week 22 or
- Matching placebo for dupilumab (2 x 2mL placebo injections) on Day 1 (ie, the loading dose), then 1 placebo injection Q2W until week 10, and then 2 x 300 mg dupilumab injections at week 12 (ie, the loading dose), followed by dupilumab injections Q2W with last injection at week 22.

Non-investigational medical products

Starting on Day -7, all patients are required to undergo background treatment with medium-potency topical corticosteroids (TCS) [i.e., non-investigational medical products (NIMP)].

3.4. Dose Adjustment/Modifications

Dose modification of dupilumab for an individual patient is not allowed.

4. General Statistical Considerations

4.1. Sample Size

The sample size calculations were based on the subset of patients who reported ≥5 units of baseline sleep loss and ≥3 units of baseline pruritus scores on the SCORAD instrument in a previous study (AD1424) evaluating efficacy of dupilumab for AD.

Since no previous data are available on daily sleep on the NRS scale, the SCORAD sleep component (VAS) data was used as a basis for the sample size determination for the primary endpoint. The variability (standard deviation) estimate for the subset of patients mentioned above was 37% for the % change from baseline to week 12. Given the uncertainty in the variability from this variable to the NRS weekly average sleep score, the above variability was inflated by ~20%, and assumed to be 44% for the primary sleep parameter. With these assumptions, a total sample of 150 (50 in placebo and 100 in dupilumab group) patients will provide 90% power (alpha = 0.05, 2-sided) to detect a treatment difference of 25% (effect size, $\Delta/\sigma = 0.57$) for the percent change from baseline.

For secondary actigraphy endpoints, sample size estimates are provided in reference to Nemolizumab Ph2 data at week 4. For TST, to detect a 30 minutes difference between treatments: (2:1 allocation ratio; alpha=0.05, 2-sided), assuming the same variability at week 16 weeks as seen for Nemolizumab at week 4 (=56 min), then 167 subjects are needed to provide 90% power. Furthermore, the proposed sample size would have >90% power to detect a difference of 10% for change from baseline in actigraphy based sleep efficiency, assuming a variability of 11% (alpha=0.05, 2-sided). The variability estimate was based on for actigraphy based sleep efficiency at week 4, inflated by 50% to account for possible increase in variability over time (from week 4 to week 16). Of note, the data used for actigraphy parameters are limited.

Allowing for 20% dropouts, a total of approximately 201 patients will be enrolled in the study.

Considering the uncertainty of the variability of the primary endpoint, when data are available for this endpoint from about 50% of the patients completing Week 12, the variability will be assessed in a blinded manner, and the sample size will be re-estimated. If the re-estimated sample size is higher than the original estimate, the sample size may be increased accordingly, not to exceed a total of 150 additional patients.

Since this analysis will be performed in a blinded manner, no adjustment to the type I error will be made in the final analysis.

4.2. Randomization, Stratification, and Blinding

Randomization will be stratified by: PSG (regardless of geographic location), non-PSG US, non-PSG Europe and non-PSG rest of world (ROW).

A participant who has been allocated to a randomized intervention regardless whether the intervention kit was used or not (ie, participant registered by the interactive response technology (IRT) will be considered a randomized subject). A participant cannot be randomized more than once in the study.

4.3. Analysis Sets

Patients exposed to IMP before or without being randomized will not be considered randomized and will not be included in any analysis set and will be reported in listings under a treatment group named "not randomized but treated". However, if these patients experienced any safety event, they would be documented separately in the clinical study report.

Patients randomized but not treated will be included in any efficacy analyses and will be reported in listings under a treatment group named "randomized but not treated". These patients will be analyzed for efficacy analyses according to the treatment group to which they are allocated to by the IRT.

4.3.1. Screened Analysis Set

The Screened analysis set includes all participants who sign the informed consent form (ICF).

4.3.2. Intent-to-Treat (ITT) Analysis Set

The Intent-to-Treat analysis set (ITT) consists of randomized patients (patients with a treatment kit number allocated and recorded in the IRT database, regardless of whether the treatment kit was used or not).

4.3.3. Modified ITT (mITT) Analysis Set

The modified ITT analysis set (mITT) consists of ITT patients who had a baseline and at least one post-baseline measurement.

4.3.4. OLE Analysis Set

The Open-Label Extension (OLE) analysis set consists of ITT patients who completed the double-blind treatment, entered the open label extension period and took at least one dose during the OLE period.

4.3.5. PSG Analysis Set

The PSG analysis set consists of ITT patients who participated in the PSG sub-study and have both baseline and at least one post-baseline PSG assessment.

4.3.6. Safety Analysis Set

The Safety analysis set includes all randomized patients who actually received at least 1 dose or partial dose of the IMP, analyzed according to the treatment actually received.

Randomized patients for whom it is unclear whether the study medication was taken will be included in the safety analysis set as randomized.

For patients receiving more than one IMP (placebo and dupilumab) during the study, the actual treatment group for as treated analysis will be the dupilumab group.

4.3.7. Analysis set without trial impact (disruption) due to COVID-19

The analysis set without trial impact (disruption) due to COVID-19 includes any patient:

- without any critical or major protocol deviation related to COVID-19
- and who did not permanently discontinue treatment due to COVID-19
- and who did not permanently discontinue study due to COVID-19

4.3.8. Analysis set with trial impact (disruption) due to COVID-19

The analysis set with trial impact (disruption) due to COVID-19 includes any patient:

- with any critical or major protocol deviation related to COVID-19

- or who permanently discontinued treatment due to COVID-19
- or who permanently discontinued study due to COVID-19

Critical or major protocol deviations related to COVID-19 will be reviewed and finalized before database lock.

4.4. Reporting Conventions

Statistical analysis will be performed using SAS® Version 9.4 or higher.

Standardized and validated SAS macros from PPD will be used to set-up table, listing, figure (TLF) formats (headers/footers and tabulation format) and tabulate the summaries. All tables and listings will be independently validated using double programming; all figures will be independently validated manually.

4.4.1. Treatment labels

The following treatment labels will be used in the Tables, Figures and Listings. Where applicable, in the listing,

- patients who are randomized but not treated will be presented in a "Randomized, not treated" column/by-line statement, which will be presented prior any other treatment group,
- patients who are not randomized but treated will be presented in a "Not randomized, treated" column/by line statement, which will be presented prior any other treatment group.

Table 4.4.1 Treatment labels

Treatment order	Treatment group	Treatment Label
1	Not randomized, treated	Not randomized, treated
2	Randomized, not treated	Randomized, not treated
3	Placebo	Placebo
4	Dupilumab 300 mg q2w	Dupilumab 300 mg q2w
5	Total	Total

4.4.2. Visit naming Conventions

The electronic Case Report Form (eCRF) visit label will be used to classify the assessments.

4.4.3. Visit Windows

The visit windows as defined in Section 15.5 will be applied to all endpoints.

4.4.4. Unscheduled Visits

Unscheduled visit measurements will be used in the analysis on efficacy variables and will be included in the by-visit summaries for the safety variables if they are re-allocated to scheduled visits according to the visit window definitions in Section 15.5.

4.4.5. Display of Data Summary and Analysis

Continuous variables will be summarized using descriptive statistics, including the following: sample size (n), mean, standard deviation (SD), median, first quartile (Q1), third quartile (Q3), minimum and maximum for each treatment group.

All mean, Q1, Q3 and median values will be formatted to one more decimal place than the measured value. SD values will be formatted to two more decimal places than the measured value. Minimum and maximum will be formatted to the same number of decimal places as the measured value.

95% confidence intervals (CIs) will be two-sided and displayed to the same level of precision as the statistic they relate to. If an estimate or a CI is not estimable, it will be presented as 'NE'. If neither an estimate, nor its CI are estimable, it will be presented as simply 'NE', not displaying 'NE' twice.

The p-values will be two-sided and will be rounded to three decimal places. If a p-value is less than 0.001 it will be reported as '<0.001'. If a p-value is greater than 0.999 it will be reported as '>0.999'.

Categorical and ordinal data will be summarized using the counts and percentages. When count data are presented, the percent will be suppressed when the count is zero in order to draw attention to the non-zero counts. A row denoted 'Missing" will be included in count tabulations for demographics, baseline characteristics and compliance to account for missing values. No percentages will be displayed on the 'Missing' rows and the percentages on the other rows will be based on the number of non-missing observations. Unless otherwise specified, the denominator for all other percentages will be the number of subjects in that treatment within the specific analysis set of interest. All percentages will be rounded to one decimal place. The number and percentage of responses will be presented in the form xx (xx.x), where the percentage is in the parentheses. When the numerator is equal to the denominator, the percentage should be presented as (100) instead of (100.0).

All listings will be sorted for presentation in order of assigned treatment arm, study center number, patient number and date of procedure or event.

4.4.6. Baseline, study day and duration derivations

The baseline value of efficacy parameters is defined as the last available valid (non-missing) value on or prior to randomization and prior to the first dose of study medication except as noted below.

- For assessments to be completed daily during the week prior to baseline visit (e.g., sleep NRS, actigraphy parameters, sleep diary, etc.), their baseline value will be the average of 7 continuous - days of assessments (or at least 5 days if data for 7 days were not available) prior to the Visit 1 Week 0 baseline visit but after the Screening visit.
- Polysomnographic data collected at the acclimation night (Day -2) will not be used.

The baseline value of safety parameters is defined as the last available valid (non-missing) value prior to the first dose of investigational medicinal product (IMP).

If multiple valid values of a variable (efficacy or safety) exist within a same day and time is not available, the scheduled value will be used. If no scheduled value is available, then the average will be selected.

Baseline safety and efficacy results are presented in the safety and efficacy analyses.

The reference day (denoted as Day 1) for the calculation of extent of exposure and study day for AE listings will be the day of the first administration of IMP:

- Extent of exposure (see Section 7.2.1)
- Study day for AE listing: AE onset date – first administration date of IMP + 1

The reference day (denoted as Day 1) for the calculation of study day of efficacy assessments will be the randomization date.

- For visit prior to randomization, study day = assessment date – randomization date.
- For visit at or after randomization, study day = assessment date – randomization date + 1.

Intervals that are presented in weeks will be transformed from days to weeks by using (without rounding) the following conversion formula:

WEEKS = DAYS /7

Intervals that are presented in months will be transformed from days to months by using (without rounding) the following conversion formula:

MONTHS = DAYS /30.4375

The last on-treatment value of safety parameters is the final measurement assessed during the treatment epoch (double-blind treatment period and OLE period respectively), regardless of the introduction of rescue therapy, including measurements at unscheduled visits.

4.4.7. Change from baseline and Percent change from baseline

Change from baseline is defined as: Change from baseline = Value at specific time point – Baseline value

Percent change from baseline is defined as:

$$\text{Percent change from baseline}(\%) = \frac{\text{Value at specific timepoint} - \text{Baseline value}}{\text{Baseline value}} \times 100\%$$

4.5. Intercurrent Event Types

Table 4.5.1 specifies the types of Intercurrent events (IcEvs), and associated labels, used to define the estimands.

Table 4.5.1 Intercurrent Event Types

Label	Intercurrent Event Type
IcEv1 (discontinuation or interruption of study treatment due to Covid-19)	Treatment discontinuation or interruption due to Covid-19
IcEv2 (discontinuation or interruption of study treatment not due to Covid-19	Treatment discontinuation or interruption not due to Covid-19
IcEv3 (Rescue Meds)	Use of rescue medications
IcEv4 (Prohibited Meds)	Use of prohibited medications

5. Patient Disposition

5.1. Disposition

The patient disposition will be summarized for the ITT analysis set by treatment group and overall using number and percentage. This section describes patient disposition for both patient study status and the patient analysis sets.

For patient study status, the number and percentage of patients in the following categories will be presented:

- Screened patients
- Screen failure patients
- Patients treated without being randomized
- Randomized patients
- Randomized but not treated patients

- Randomized and treated patients
- Patients who have completed treatment during the double-blind treatment period
- Patients who did not complete treatment during the double-blind treatment period and the main reason for treatment discontinuation. Relationship with COVID-19 will be reported for "Adverse Event" and "Other" reasons will be summarized too.
- Patients who discontinued the study during the double-blind treatment period and the main reason for study discontinuation. Relationship with COVID-19 will be reported for "Adverse Event" and "Other" reasons will be summarized too.
- Patients who entered into open-label extension treatment period (OLE)
- Patients who did not complete treatment during the OLE treatment period and the main reason for treatment discontinuation. Relationship with COVID-19 will be reported for "Adverse Event" and "Other" reasons will be summarized too.
- Patients who completed study
- Patients who discontinued the study during the OLE treatment period and the main reason for study discontinuation. Relationship with COVID-19 will be reported for "Adverse Event" and "Other" reasons will be summarized too.
- Patients who had rescue medications (see Section 15.11)
    - Patients who have completed the double-blind treatment period and had rescue medications
    - Patients who did not complete the double-blind treatment period and had rescue medications
- Final Study Status

For screened, screen failure, and patients treated without being randomized, percentages will be calculated using the number of screened patients as the denominator for overall only. All other categories of patients will be presented by treatment group and for overall and the percentages will be calculated using the number of randomized patients within each treatment group and overall, as denominator. Reasons for treatment discontinuation will be supplied in tables giving number and percentage by treatment group.

If applicable, patients treated but not considered as randomized, and patients randomized but not treated will be identified and described in separate listings.
Listings of patients with permanent treatment discontinuation, patients with premature end of study will be provided.

Additionally, a summary of all the analysis sets for safety and efficacy (Efficacy [ITT, m-ITT, OLE, and PSG], Safety, analysis set without trial impact (disruption) due to COVID-19 and analysis set with trial impact (disruption) due to COVID 19) will be summarized for the ITT analysis set by treatment group and overall using number and percentage.

A listing of patients excluded from analysis sets will be provided. Patients with permanent treatment discontinuation, patients with premature end of study will be identified and described in separate listings.

The reasons for being in the analysis set with trial impact (disruption) due to COVID-19 will be summarized for the modified ITT analysis set by treatment group and overall using number and percentage.

A listing of patients excluded from the analysis set without trial impact (disruption) due to COVID-19 will be provided.

The disposition of screened patients, screen failure patients, randomized and not treated patients, randomized patients, randomized and treated patients by country and site will be summarized for the screened analysis set by treatment group and overall using number and percentage.

A summary table to show patient disposition by visit according to trial impact (disruption) due to COVID-19 (visit not done, visit partially done on site, visit partially done by phone and visit done but delayed) will be provided. A summary table to show the patient disposition by visit according to trial impact (disruption) due to COVID-19 by country and site will be provided too.

5.2. Protocol Deviations

All critical or major protocol deviations including randomization and drug-dispensing irregularities will be summarized according to COVID-19 impact (i.e. deviations related to COVID-19 pandemic and deviations not related to COVID-19 pandemic) for the ITT analysis set by treatment group and overall using number and percentage.

A listing of patients with at least one critical or major protocol deviations will be provided.

Randomization and drug-dispensing irregularities occur whenever:

1. A randomization is not in accordance with the protocol-defined randomization method, such as a) an ineligible patient is randomized, b) a patient is randomized based on an incorrect stratum, c) a patient is randomized twice, or d) in a dynamic randomization scheme the treatment assignment is, in fact, not random, due to a computer program error.

   OR

2. A patient is dispensed an IMP kit not allocated by the protocol-defined randomization, such as a) a patient at any time in the study is dispensed a different treatment kit than as randomized (which may or may not contain the correct-as-randomized IMP), or b) a non-randomized patient is treated with IMP reserved for randomized patients.

Randomization and drug-dispensing irregularities will be monitored throughout the study and reviewed on an ongoing basis.

All randomization and drug-dispensing irregularities will be documented in the clinical study report. If the number of irregularities is large enough to make a tabular summary useful, the critical or major irregularities will be categorized and summarized for the ITT analysis set by treatment group and overall using number and percentage. Non-randomized, treated patients will be described separately.

A listing of patients with at least one critical or major protocol deviations related to randomization and drug-dispensing irregularities will be provided.

Randomization and drug-dispensing irregularities to be prospectively identified include but are not limited to:

Table 5.2.1 Randomization and drug dispensing irregularities

Randomization and drug allocation irregularities
• Kit dispensation without IRT transaction • Erroneous kit dispensation • Kit not available • Randomization by error • Patient randomized twice • Stratification error

- Open-label kit dispensed instead of double-blind kit at Week 12

List of randomization and drug-dispensing irregularities will be provided by SANOFI.

6. Demographics and Baseline Characteristics

The demographics, patient characteristics and baseline disease characteristics will be summarized using the ITT and PSG analysis sets by treatment group and overall using descriptive statistics or using number and percentage. P-values on the treatment difference for the demographic and baseline characteristics data will not be calculated. In general, no specific description of the safety parameters will be provided at baseline. If relevant, the baseline values will be described along with each safety/efficacy analysis.

6.1. Demographics

The following demographics and patient characteristics will be summarized by treatment group and overall:

- Age (years),
- Age categories (≥18 to <40, ≥40 to <65, ≥65),
- Gender (Male, Female)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino, Not Reported, Unknown),
- Race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, Multiple, White, Not Reported, Unknown),
- Baseline body weight (kg) with grouping (<70, ≥70 to <100, ≥100),
- Baseline height (cm),
- Baseline body mass index (BMI) (kg/m^2) derived as: (Weight in kg)/(Height in meters)^2,
- Baseline BMI categories (<15, ≥15 to <25, ≥25 to <30, ≥30)
- Region (US [PSG, non-PSG], Europe and ROW)
- Alcohol habits at screening (Never, Occasional, At Least Daily, At Least Weekly, At Least Monthly)

6.2. Baseline Disease Characteristics

The following baseline disease characteristics will be summarized by treatment group and overall:

6.2.1. Sleep characteristics and sleep related impairment at baseline

- Rescaled Sleep quality NRS*
- SCORAD subjective sleep loss score (sleep VAS)
- PROMIS sleep disturbance short form 8a raw and T-scores
- Actigraphy
  - Sleep onset latency (SOL)*
  - Wake after sleep onset (WASO)*
  - Total Sleep Time (TST)*
  - Sleep efficiency (SE)*
- Sleep diary
  - SOL*
  - TST*
  - WASO*
  - SE*
  - Number of awakenings*
  - How rested the participant felt when getting up (NRS) *

- ESS score, and the categories as defined by
  - Lower Normal Daytime Sleepiness (0-5)
  - Higher Normal Daytime Sleepiness (6-10)
  - Mild Excessive Daytime Sleepiness (11-12)
  - Moderate Excessive Daytime Sleepiness (13-15)
  - Severe Excessive Daytime Sleepiness (16-24)
- PVT (Outcome metrics from data transfer)
  - Total errors
  - Number of lapses
  - Mean reciprocal response time (mean 1/RT)
  - Mean fastest response time (Fastest 10% RT)
  - Mean slowest reciprocal response time (Slowest 10% RT)
- ANAM: running memory, mathematical processing, procedural reaction time
  - Mean response time
  - Percentage of correct response
  - Throughput
- Polysomnography in the PSG analysis set (defined in 4.3.5)
  - SOL
  - WASO
  - TST
  - SE
  - Longest sleep episode
  - Stages of sleep
  - Arousal index

* average value of the assessment reported during the week prior to baseline visit

6.2.2. AD and associated disease characteristics at baseline

- Investigator's Global Assessment (IGA) score
  - Clear (0)
  - Almost clear (1)
  - Mild disease (2)
  - Moderate disease (3)
  - Severe disease (4)
- Body Surface Area (BSA) Involvement of AD
- EASI total score
- SCORAD overall score
- Peak pruritus NRS*
- Skin pain NRS*
- Skin sensitivity to touch NRS*
- Skin burn NRS*
- POEM score and categories as defined by
  - Clear or almost clear (0-2)
  - Mild eczema (3-7)
  - Moderate eczema (8-16)
  - Severe eczema (17-24)
  - Very severe eczema (25-28)
- DLQI total score
- HADS-Anxiety score, HADS-Depression score

- ACQ-5 total score
- AR-VAS score

6.2.3. Health economic endpoints at baseline

- WPAI-AD: currently employed status (yes/no); number of hours missed due to AD (only if employed); number of hours missed due to other reasons (only if employed); number of hours actually worked (only if employed); NRS of atopic dermatitis affected productivity while working (only if employed) and NRS of atopic dermatitis affected ability to do regular daily activities, other than work at job
- Day missed from school or from work due to AD : patients who were employed or enrolled in school were asked to report the number of sick leave/missed school days due to AD from each assessment since the last study visit, separately for employment/school status by full time and part time

6.3. Alcohol, Tobacco and Caffeine Usage

The following alcohol, tobacco and Caffeine habits will be summarized by treatment group and overall:

- Alcohol consumption: amount (glass) per day and per visit
- Caffeine consumption: amount (cup) per day and per visit
- Tobacco consumption: amount (cigarette) per day and per visit

The tobacco and caffeine data are only collected for PSG patients.

6.4. Medical History

6.4.1. General Medical History

Medical (or surgical) history includes all the relevant medical (or surgical) history during the lifetime of the patient.

Medical and surgical history will be coded to “lower level term (LLT)”, “preferred term (PT)”, “high level term (HLT)”, “high level group term (HLGT)”, and associated primary “system organ class (SOC)” using the version of Medical Dictionary for Regulatory Activities (MedDRA) in effect at Sanofi at the time of each database lock.

The number and percentage of subjects with any medical history will be summarized by treatment group and overall and for each SOC. SOC will be sorted in descending order of frequency based on the total of all treatment groups.

6.4.2. Disease-Specific History

Not applicable.

6.5. Inclusion and Exclusion Criteria

Inclusion and exclusion criteria will be presented for the screened analysis set for each treatment group and overall using number and percentage.

6.6. Other Background Information

Not applicable.

7. Treatments and Medications

7.1. Prior and Concomitant Medications

The prior and concomitant medications will be presented for the ITT analysis set for each treatment group and overall using number and percentage. No statistical test for the between-group difference will be performed.

All medications will be coded using the World Health Organization-Drug Dictionary (WHO-DD) using the version in effect at Sanofi at the time of each database lock.

Medications will be summarized by treatment group according to the WHO-DD, considering the first digit of the anatomic category (ATC) class (anatomic category) and the first 3 digits of the ATC class (therapeutic category). All ATC codes corresponding to a medication will be summarized, and patients will be counted once in each ATC category (anatomic or therapeutic) linked to the medication. Therefore, patients may be counted several times for the same medication.

The summaries for prior, concomitant, and post-treatment medications will be sorted by decreasing frequency of ATC followed by all other therapeutic classes based on the incidence in the dupilumab treatment group. In case of equal frequency regarding ATCs, alphabetical order will be used.

For the purpose of inclusion in prior and/or concomitant medication tables, incomplete medication start and stop dates will be imputed as described in section 15.4.3.

7.1.1. Prior Medications

Prior medications are those the patient began prior to first IMP intake. Prior medications can be discontinued before first treatment administration or can be ongoing during the treatment phase.

7.1.2. Concomitant Medications

Concomitant medications are any treatments received by the patient concomitantly to the IMP, starting from the 1st administration of IMP to the date of last administration + 14 days (boundary not included). A given medication can be classified both as a prior medication and as a concomitant medication.

7.1.3. Post-treatment medications

Post-treatment medications are those the patient took (continued or initiated) in the period running from the 14th day after the last administration of IMP up to the end of the study.

7.2. Study Treatments

The extent of IMP exposure and compliance will be assessed and summarized separately within the double-blind treatment period, the OLE period (for compliance only) and overall period (combined double-blind and OLE periods) by actual treatment group and overall using the safety analysis set and analysis set with trial impact (disruption) due to COVID-19.

7.2.1. Extent of Exposure

The extent of IMP exposure will be assessed by the duration of IMP exposure.

Duration of IMP exposure is defined as last IMP administration dose date – first IMP administration dose date + 14 days, regardless of unplanned intermittent discontinuations.

Duration of IMP exposure will be summarized descriptively as a quantitative variable (number, mean, SD, median, Q1, Q3, minimum and maximum). In addition, duration of IMP exposure will also be summarized categorically by numbers and percentages for each of the following categories and cumulatively according to these categories:

- >0 and ≤2 weeks
- >2 and ≤4 weeks
- >4 and ≤8 weeks
- >8 and ≤12 weeks
- >12 and ≤16 weeks
- >16 and ≤20 weeks
- >20 and ≤24 weeks
- >24 weeks

Additionally, the cumulative duration of IMP exposure will be provided, defined as the sum of duration of IMP exposure for all patients, and will be expressed in patient years.

7.2.2. Treatment Compliance and Modifications

A given administration will be considered non-compliant if the patient did not take the planned dose of treatment as required by protocol (i.e. a syringe not fully injected is considered as a non-compliant administration). No imputation will be made for missing or incomplete data.

Percentage of compliance for a patient will be defined as the number of administrations that the patient was compliant (planned dose fully injected) divided by the total number of administrations that the patient was planned to take during each of treatment period (from the first to the last date of IMP administration within each period).

$$\text{Percentage of compliance (\%)} = \left[\frac{\text{Total number of compliant injections during specific treatment period}}{\text{Number of planned injections during specific treatment period}}\right] \times 100\%$$

Loading doses for the same patient will be counted as 1.

Percentage of compliance and percentage of injections with compliance to the IMP administration will be summarized descriptively as quantitative variables (number, mean, SD, median, Q1, Q3, minimum and maximum) separately for each treatment period (double-blind treatment period and OLE) and overall. In addition, the percentage of compliance will be presented by the specific ranges for each treatment group:

- <80%
- ≥80 % to < 120%
- ≥120%

An overdose (accidental or intentional) with the IMP is an event suspected by the Investigator or spontaneously notified by the participant and defined as at least twice the intended dose within the intended therapeutic interval, adjusted according to the tested drug. Cases of symptomatic overdoses will constitute AESIs and will be listed as such. More generally, dosing irregularities will be listed in Section 5.2.

Dose modification of IMP for an individual patient is not allowed and therefore no summary of dose modifications will be provided. Drug-dispensing irregularities are included as part of the protocol deviations summary (Section 5.2).

8. Efficacy Analysis

Efficacy analyses will be performed on the mITT analysis set, OLE analysis set or PSG analysis set depending on the efficacy endpoints.

Patients will be analyzed for efficacy according to the treatment group to which they are allocated by the IRT according to the randomization schedule at the randomization visit (as randomized), irrespective of the treatment actually received.

For the analysis of primary objective in the double-blind treatment period,

- The null statistical hypothesis tested is that there is no treatment difference between dupilumab and placebo on sleep quality NRS in adult patients with moderate to severe AD and sleep disturbance.
- the alternative statistical hypothesis tested is that there is a treatment difference between dupilumab and placebo on sleep quality NRS in adult patients with moderate to severe AD and sleep disturbance.

Note that due to the small sample size in the PSG analysis set, no statistical testing will be conducted for PSG related endpoints. Summary statistics with 95% CI by treatment group will be provided.

Statistical testing will be only be conducted on results for the double-blind treatment period. Statistical testing will not be conducted for the assessments completed during OLE period. All related endpoints will be summarized.

All efficacy measurements collected during the study will be considered for analyses, including those obtained after IMP discontinuation.

All daily assessments (e.g., sleep quality NRS, actigraphy parameters, sleep diary, etc.) will be analyzed as weekly averages for the week immediately preceding the study visit.

Line plots for change from baseline or percent change from baseline will be presented by treatment group, displaying the mean, median, Q1 and Q3 at each visit. These plots will be provided for the primary and all secondary continuous endpoints.
Bar plots for percentage of patients with EASI50, EASI75 and EASI90 will be presented by treatment group, displaying the percentage of EASI responders at each visit.

8.1. Primary Efficacy Endpoint

The sleep quality scale was used to assess the quality of the participant's previous nights' sleep, collected on a 11-point scale (0 to 10) in which 0 indicated worst possible sleep while 10 the best possible sleep. The primary efficacy endpoint percentage change from baseline will be calculated based on the rescaled values where each assessment (of the average over the week preceding the study visit) will be rescaled as (10 minus the observed value) such that 0 indicated the best possible sleep and 10 indicated the worst possible sleep based on the rescaled value. This is done in order to have consistent interpretations on percent improvement between scales (e.g. Peak pruritus NRS).

Rescaled sleep quality NRS = 10 minus observed sleep quality NRS.

$$\text{Percent change from baseline (\%)} = \left[\frac{\text{Week 12 rescaled sleep NRS} - \text{rescaled sleep NRS baseline value}}{\text{rescaled sleep NRS baseline value}}\right] \times 100$$

Therefore, a negative percent change from baseline will indicate an improvement.

8.1.1. Primary Analysis

Percentage change from baseline in sleep quality NRS will be analyzed using a MMRM under the MAR framework. The MMRM model will include treatment (dupilumab, placebo), baseline value of sleep quality NRS, randomization stratum (region), visits (up to week 12), treatment by visit interaction, and baseline sleep quality NRS by visit interaction terms all as fixed effects. An unstructured correlation matrix will be used to model the within -patient errors. The Kenward-Roger approximation (**ddfm=kr**) will be used to estimate the denominator of degrees of freedom. Parameters will be estimated using restricted maximum likelihood method with the Newton-Raphson algorithm. LS mean difference between the dupilumab and placebo group, the corresponding 95% CI of the differences and 2-sided nominal p-values will be provided. The primary p-value is at Week 12. Any other p-values should be treated as suggestive for possible further study as the study is not powered for these tests.

If the number of patients in the PSG-sub study is less than 50% of planned, PSG-strata will be pooled with the non-PSG strata in the US and then there will be only 3 strata: US (combined non-PSG and PSG), Europe and ROW.

If the MMRM model fails to achieve convergence due to complexity of model specification, different covariance structures will be used according to the following order till convergence is achieved:

1. Unstructured correlation (UN)
2. Heterogeneous Toeplitz (TOEPH)
3. Homogeneous Toeplitz (TOEP)
4. First-Order Autoregressive [AR(1)]
5. Compound Symmetry (CS)

MMRM model will be built with the following sample SAS code to evaluate the treatment effect through week 12:

```
proc mixed data = adqs method = reml;
        where paramcd eq 'sleep quality NRS' and avisitn ge 2 and avisitn le 6 and ITTFL eq
        "Y";
        class usubjid stratum trt01pn avisitn;
        model pctchg = base trt01pn stratum avisitn trt01pn*avisitn base*avisitn / solution ddfm
        = kr;
        repeated avisitn / type = un subject = usubjid;
        lsmeans trt01pn*avisitn / cl;
        /*contrast "Dupilumab vs Placebo at Week 1 Visit 2"
                trt01pn -1 1 trt01pn*avisitn -1 0 0 0 0 1 0 0 0 0;
        contrast "Dupilumab vs Placebo at Week 2 Visit 3"
                trt01pn -1 1 trt01pn*avisitn 0 -1 0 0 0 0 1 0 0 0;
        contrast "Dupilumab vs Placebo at Week 4 Visit 4"
                trt01pn -1 1 trt01pn*avisitn 0 0 -1 0 0 0 0 1 0 0;
        contrast "Dupilumab vs Placebo at Week 8 Visit 5"
                trt01pn -1 1 trt01pn*avisitn 0 0 0 -1 0 0 0 0 1 0; */
        contrast "Dupilumab vs Placebo at Week 12 Visit 6"
                trt01pn -1 1 trt01pn*avisitn 0 0 0 0 -1 0 0 0 0 1;
        ods output diffs=diffs lsmeans=lsmeans;
run;
```

Note that sample code assumes the avisitn for visit 2 is 2 and avisitn for visit 6 is 6, if that is not the case, the code should be adjusted. The contrast statement assumes trt01pn for placebo is lower than that of

dupilumab. The sample contrast statement for other timepoint except week 12 is provided as well in the comment, which could be used for comparisons at other visits.

8.1.2. Assumption Testing

No assumption testing will be performed.

8.1.3. Missing Data Handling

No imputation will be made for the missing values for the primary approach.

8.1.4. Subgroup Analysis

The primary endpoint will be summarized using the following subgroups:

- Age categories (≥18 to <40, ≥40 to <65, ≥65),
- Gender (Male, Female)
- Baseline Body weight at screening (kg) with grouping (<70, ≥70 to <100, ≥100),
- Region (US [non-PSG and PSG], Europe and ROW).
- Baseline IGA (moderate [IGA=3] and severe [IGA=4])
- Baseline moderate to severe EASI (≤21, >21)
- Baseline severe EASI (<25, ≥25)
- Baseline SCORAD (≤50, >50)
- Baseline peak pruritus NRS (<7, ≥7)
- Previous use of systemic immunosuppressants for AD (Yes, No)

List of coded terms for systemic immunosuppressants (medications) used for AD will be provided by SANOFI.

Results will also be illustrated with forest plots.

8.1.5. Sensitivity Analysis

If a specific weekly average for the sleep quality NRS was derived using 3 or less assessments, such assessments will be set to missing in any sensitivity analysis.

8.1.6. Supplementary Analyses

Not applicable.

8.1.7. Impact of Covid-19 Pandemic

During the Covid-19 pandemic, a business continuity plan (BCP) was put in place to minimize the impact of the clinic visit interruption, ensure IMP treatment continuity and data collection. Direct to patient delivery of IMP from the site(s) was made available, home injection(s) were done as planned per protocol. IMP interruption due to Covid-19 pandemic is therefore expected to be minimum. Data collection for ¨PROs are not expected to be missing as they can be completed at patient's home on his/her iPad. Site(s) can update/correct the schedule of completion if visit is rescheduled/delayed – unless site staff cannot access vendor platform to update scheduled dates.

Missing data due to Covid-19 pandemic is likely to be missing at random. Therefore, the primary analysis using MMRM can adequately address this type of missing data. Patient disposition by visit according to trial impact (disruption) due to COVID-19 and by country and site are already mentioned in section 5.1 of this SAP.

8.2. Secondary and Exploratory Efficacy Endpoint

The secondary efficacy endpoints associated with "Change from baseline" and "Percent change from baseline" during the double-blind treatment period will be analyzed using the same approach as used for the analysis of the primary endpoint.
Definitions of endpoints are included in Sections 15.8, 15.9 and 15.10.

8.2.1. Actigraphy data

SE, TST, WASO and SOL, all based on actigraphy data, will be summarized at each post-baseline visit (Week 1, Week 2, Week 3, Week 4, Week 8, Week 12, Week 16, Week 20, and Week 24).

In addition, at each post-baseline visit until Week 12, the change from baseline will be modelled using the same MMRM model used for the analysis of the primary endpoint. LS mean differences between the dupilumab and placebo group, the corresponding 95% CI of the differences and two-sided p-values will be provided. All p-values except for Week 12 will be considered nominal.

Change from baseline to Week 24 and Change from Week 12 to Week 24 will be summarized on the OLE analysis set using descriptive statistics in each treatment group (those originally randomized to in the double-blind study treatment period). No comparative statistical analyses will be done.

8.2.2. Sleep diary

SE, TST, WASO, SOL, number of awakenings and how rested the participant felt when getting up, all based on sleep diary data, will be summarized at each post-baseline visit (Week 1, Week 2, Week 3, Week 4, Week 8, and Week 12).

In addition, at each post-baseline visit, the change from baseline will be modelled using the same MMRM model used for the analysis of the primary endpoint. LS mean differences between the dupilumab and placebo group, the corresponding 95% CI of the differences and two-sided p-values will be provided. All p-values except for Week 12 will be considered nominal.

8.2.3. Sleep Quality NRS

Rescaled sleep Quality NRS (see section 8.1) will be summarized at each post-baseline visit (Week 1, Week 2, Week 3, Week 4, Week 8, Week 16, Week 20, and Week 24)

In addition, at each post-baseline visit until Week 12, the percent change from baseline will be modelled using the same MMRM model used for the analysis of the primary endpoint. LS mean differences between the dupilumab and placebo group, the corresponding 95% CI of the differences and two-sided p-values will be provided. All p-values except for Week 12 will be considered nominal.

Percent change from baseline to Week 24 and percent change from Week 12 to Week 24 will be summarized on the OLE analysis set using descriptive statistics in each treatment group (those originally randomized to in the double-blind study treatment period). No comparative statistical analyses will be done.

8.2.4. Peak Pruritus NRS

Peak pruritus NRS will be summarized at each post-baseline visit (Week 1, Week 2, Week 3, Week 4, Week 8, Week 12, Week 16, Week 20, and Week 24). Daily results for each day of Week 1 will also be provided.

In addition, at each post-baseline visit until Week 12, the percent change from baseline will be modelled using the same MMRM model used for the analysis of the primary endpoint. LS mean differences between

the dupilumab and placebo group, the corresponding 95% CI of the differences and two-sided p-values will be provided. All p-values except for Week 12 will be considered nominal.

Percent change from baseline to Week 24 and percent change from Week 12 to Week 24 will be summarized on the OLE analysis set using descriptive statistics in each treatment group (those originally randomized to in the double-blind study treatment period). No comparative statistical analyses will be done.

Times to pruritus improvement are defined as:

1. Time to achieving ≥3 points improvement in NRS [Time frame: Baseline till Week 4]
2. Time to achieving ≥4 points improvement in NRS [Time frame: Baseline till Week 4]

The time variable will be the first day of achieving ≥x points improvement, and if that was not achieved during the first 4 weeks (where daily assessments were done), it will be censored at 28 days. Although the time variable is counting the number of days, the baseline will still be the weekly average at Visit 1 since that is the inclusion criteria and better represents the patient pruritus at baseline. The baseline date will be the date of Visit 1. Patients who discontinued for any reason within the first 4 weeks will be censored at the time of discontinuation.

The median time to pruritus improvement and proportion of patients achieving ≥x points improvement at Week 1, Week 2, Week 3, and at Week 4 along with the corresponding 95% CI will be calculated using Kaplan Meier method.

A summary with the number of patients achieving the pruritus improvement, the number of censored patients, the median time, rates at Week 1, Week 2, Week 3, and Week 4 along with the corresponding 95% CI will be provided. Kaplan-Meier curves will be provided too.

8.2.5. Skin Pain NRS

Skin Pain NRS will be summarized at each post-baseline visit (Week 1, Week 2, Week 3, Week 4, Week 8, and Week 12). Daily results for each day of Week 1 will also be provided.

In addition, at each post-baseline visit, the change from baseline will be modelled using the same MMRM model used for the analysis of the primary endpoint. LS mean differences between the dupilumab and placebo group, the corresponding 95% CI of the differences and two-sided p-values will be provided. All p-values except for Week 12 will be considered nominal.

8.2.6. Skin Sensitivity to Touch NRS

Skin Sensitivity to Touch NRS will be summarized at each post-baseline visit (Week 1, Week 2, Week 3, Week 4, Week 8, and Week 12).

In addition, at each post-baseline visit, the change from baseline will be modelled using the same MMRM model used for the analysis of the primary endpoint. LS mean differences between the dupilumab and placebo group, the corresponding 95% CI of the differences and two-sided p-values will be provided. All p-values except for Week 12 will be considered nominal.

8.2.7. Skin Burn NRS

Skin Burn NRS will be summarized at each post-baseline visit (Week 1, Week 2, Week 3, Week 4, Week 8, and Week 12).

In addition, at each post-baseline visit, the change from baseline will be modelled using the same MMRM model used for the analysis of the primary endpoint. LS mean differences between the dupilumab and

placebo group, the corresponding 95% CI of the differences and two-sided p-values will be provided. All p-values except for Week 12 will be considered nominal.

8.2.8. Eczema Area Severity Index (EASI)

The EASI total score and percent change from baseline in EASI total score will be summarized at each post-baseline visit. At each post-baseline visit until Week 12, the percent change from baseline in EASI total score will be modelled using the same MMRM model used for the analysis of the primary endpoint. LS mean differences between the dupilumab and placebo group, the corresponding 95% CI of the differences and two-sided p-values will be provided. All p-values except for Week 12 will be considered nominal.

In addition, the responder binary variables EASI50, EASI75 and EASI90 indicating whether the patient achieved at least 50%, 75% and 90% improvement (reduction), respectively, will be summarized and modelled. The number and percentage of patients with EASI50, EASI75 and EASI90 will be presented for each post-baseline visit (Week 1, Week 2, Week 4, Week 8, Week 12, Week 16, Week 20 and Week 24) alongside the number of patients evaluated at that visits. The percentage for each responder variable will be derived based on the number of patients who are evaluable at the visit.

Moreover, at each post-baseline visit until Week 12, the responder variables (EASI50, EASI75 and EASI90) will be modelled using Generalized Linear Mixed Model (GLMM). The terms included will be the same as those in the MMRM model used for the analysis of the primary endpoint. The model will be fitted using PROC GLIMMIX assuming a binomial distribution and logit link. The estimated incidence rate for each treatment group and the corresponding 95% CI of the estimated incidence rates will be provided. The odds ratios and the corresponding 95% CI of the odds ratio with two-sided p-value will be provided. All p-values except for Week 12 will be considered nominal.

The GLIMMIX procedure for modelling the binomial responder variables will be based on the following sample SAS code to evaluate the treatment effect through week 12:

```
proc glimmix data = adcc;
        where paramcd eq "EASI50" and avisitn ge 2 and avisitn le 6 and ITTFL="Y";
        class usubjid stratum avisitn trt01pn(ref='x');
        ** 'x' should be code for the placebo group **
        model resp (event='1') = base trt01pn stratum avisitn avisitn*trt01pn base*avisitn /
        dist=binomial link=logit solution ddfm=kr;
        random _residual_ / subject=usubjid type=UN; ** type can also change here as in mixed
        to TOEPH, TOEP, AR(1) or CS to achieve convergence.**
        lsmeans trt01pn*avisitn / pdiff ilink OR CL;
run;
```

Note that sample code assumes that the relevant post-baseline visits until Week 12 are coded between 2 and 6 inclusive, if that is not the case, the code should be adjusted.

If the GLMM model fails to achieve convergence (even with simpler covariance structures), the responder variables will be analyzed using a logistic regression model, separately at each assessment time point through Week 12. The logistic regression model will include treatment (dupilumab, placebo) as fixed effect, and will include baseline value of EASI, randomization stratum (region) as fixed effect covariates. A summary table will present the estimated incidence for each treatment group and corresponding 95%

confidence interval (CI); also, the odds ratio, corresponding 95% CI and p-value for the treatment difference between the dupilumab and placebo group. An estimate of the odds ratio and a CI above 1 is an indication that you are more likely to be a responder on the active treatment group rather than placebo.

If the logistic regression model fails to achieve convergence at a post-baseline visit, the Clopper Pearson 95% CI of the incidence rate (expressed in percentage) and the 95% CI of the difference in incidence rate between treatment group and placebo (i.e. treatment – placebo) and p-value will be presented. The 95% CI for the difference between dupilumab and placebo will be based on the Cochran-Mantel-Haenszel (CMH) test adjusted by the randomization stratum, at the post-baseline time point.

8.2.9. Patient oriented eczema measure (POEM)

The POEM total score will be summarized at each post-baseline visit (Week 4, Week 12, and Week 24).

In addition, at each post-baseline visit until Week 12, the change from baseline will be modelled using the same MMRM model used for the analysis of the primary endpoint. LS mean differences between the dupilumab and placebo group, the corresponding 95% CI of the differences and two-sided p-values will be provided. All p-values except for Week 12 will be considered nominal.

Change from baseline to Week 24 and Change from Week 12 to Week 24 will be summarized on the OLE analysis set using descriptive statistics in each treatment group (those originally randomized to in the double-blind study treatment period). No comparative statistical analyses will be done.

8.2.10. PROMIS Sleep Related Impairment Short Form 8a

The PROMIS Sleep Related Impairment SF8a total raw and T-scores will be summarized at each post-baseline visit (Week 1, Week 2, Week 4, Week 8, Week 12, Week 16, Week 20, and Week 24). T-score calculation is defined in Table 15.8.1.

In addition, at each post-baseline visit until Week 12, the change from baseline for raw score and T-score will be modelled using the same MMRM model used for the analysis of the primary endpoint. LS mean differences between the dupilumab and placebo group, the corresponding 95% CI of the differences and two-sided p-values will be provided. All p-values except for Week 12 will be considered nominal.

Change from baseline to Week 24 and Change from Week 12 to Week 24 will be summarized on the OLE analysis set using descriptive statistics in each treatment group (those originally randomized to in the double-blind study treatment period). No comparative statistical analyses will be done.

8.2.11. Epworth Sleepiness Scale (ESS)

The ESS will be summarized at each post-baseline visit (Week 4, Week 12, and Week 24).

In addition, at each post-baseline visit, the change from baseline will be modelled using the same MMRM model used for the analysis of the primary endpoint. LS mean differences between the dupilumab and placebo group, the corresponding 95% CI of the differences and two-sided p-values will be provided. All p-values except for Week 12 will be considered nominal.

8.2.12. Dermatology Life Quality Index (DLQI)

The DLQI will be summarized at each post-baseline visit (Week 4, Week 12, and Week 24).

In addition, at each post-baseline visit until Week 12, the change from baseline will be modelled using the same MMRM model used for the analysis of the primary endpoint. LS mean differences between the

dupilumab and placebo group, the corresponding 95% CI of the differences and two-sided p-values will be provided. All p-values except for Week 12 will be considered nominal.

Change from baseline to Week 24 and Change from Week 12 to Week 24 will be summarized on the OLE analysis set using descriptive statistics in each treatment group (those originally randomized to in the double-blind study treatment period). No comparative statistical analyses will be done.

8.2.13. SCORing of Atopic Dermatitis (SCORAD)

The SCORAD total score and sleep VAS will be summarized at each post baseline visit (Week 1, Week 2, Week 4, Week 8, Week 12, Week 16, Week 20, and Week 24).

In addition, at each post-baseline visit until Week 12, the change from baseline will be modelled using the same MMRM model used for the analysis of the primary endpoint. LS mean differences between the dupilumab and placebo group, the corresponding 95% CI of the differences and two-sided p-values will be provided. All p-values except for Week 12 will be considered nominal.

Change from baseline to Week 24 and Change from Week 12 to Week 24 will be summarized on the OLE analysis set using descriptive statistics in each treatment group (those originally randomized to in the double-blind study treatment period). No comparative statistical analyses will be done.

8.2.14. Juniper Asthma Control Questionnaire (ACQ-5)

The ACQ-5 global score will be summarized at each post-baseline visit (Week 12 and Week 24).

In addition, the change from baseline at Week 12 will be modelled using the same MMRM model used for the analysis of the primary endpoint. LS mean difference between the dupilumab and placebo group, the corresponding 95% CI of the difference and two-sided p-value will be provided. All p-values except for Week 12 will be considered nominal.

8.2.15. Allergic Rhinitis Visual Analog Scale (AR-VAS)

The AR-VAS score will be summarized at each post-baseline visit (Week 12 and Week 24).

In addition, at each post-baseline visit until Week 12, the change from baseline will be modelled using the same MMRM model used for the analysis of the primary endpoint. LS mean difference between the dupilumab and placebo group, the corresponding 95% CI of the difference and two-sided p-values will be provided. All p-values except for Week 12 will be considered nominal.

8.2.16. Psychomotor Vigilance Test (PVT)

The PVT variables, total errors, number of lapses, mean reciprocal response time, mean fastest response time, and mean slowest reciprocal response time will be summarized at each post-baseline visit (Week 12 and Week 24).

In addition, at each post-baseline visit until Week 12, the change from baseline will be modelled using the same MMRM model used for the analysis of the primary endpoint. LS mean difference between the dupilumab and placebo group, the corresponding 95% CI of the difference and two-sided p-values will be provided. All p-values except for Week 12 will be considered nominal.

For tests that indicate “Failed QC: clear non-compliance”, the tests and the associated metrics will be excluded from the analyses since the results are judged to be procedurally non-compliant rather than due to psychomotor impairment. If the difference in the proportion of “Failed QC: clear non-compliance” tests

between the treatments is more than 20 %, an exploratory analysis of the summaries mentioned above will be repeated by including these tests.

8.2.17. Neurocognitive Test (Automated Neuropsychological Assessment Metrics [ANAM])

The ANAM variables, mean response time for all responses, percentage of correct response and throughput will be summarized at each post-baseline visit (Week 12 and Week 24) and for each of the three subtests.

In addition, at each post-baseline visit until Week 12, the change from baseline will be modelled using the same MMRM model used for the analysis of the primary endpoint. LS mean difference between the dupilumab and placebo group, the corresponding 95% CI of the difference and two-sided p-values will be provided. All p-values except for Week 12 will be considered nominal.

For tests with validity flag as "Invalid", the tests and the associated metrics will be excluded from the ANAM data analysis. If the difference in proportion of "invalid" tests between the treatments is more than 20%, an exploratory analysis of the summaries mentioned above will be repeated by including these tests.

8.2.18. Polysomnography (PSG)

SE, TST, WASO, and SOL, all based on PSG measures, will be summarized at post-baseline visit (Week 12) on the PSG analysis set.

8.2.19. Work Productivity and Activity Impairment Questionnaire: Atopic Dermatitis (WPAI-AD)

The number and percentage of patients currently employed (yes/no) will be presented for each post-baseline visit (Week 4, Week 8, Week 12, and Week 24) alongside the number of patients evaluated at that visits

If employed, the number of hours missed due to AD in the past week, number of hours missed due to other reasons in the past week, number of hours actually worked in the past week, NRS of atopic dermatitis affecting productivity while working in the past week and NRS of atopic dermatitis affecting the ability to do regular daily activities other than work at a job in the past week will be summarized at each post-baseline visit (Week 4, Week 8, Week 12, and Week 24).

In addition, at each post-baseline visit until Week 12, the change from baseline will be modelled using the same MMRM model used for the analysis of the primary endpoint. LS mean difference between the dupilumab and placebo group, the corresponding 95% CI of the difference and two-sided p-values will be provided. All p-values except for Week 12 will be considered nominal.

8.2.20. Days missed from school or from work due to AD

The Days missed from school or from work will be summarized at each post-baseline visit (Week 4, Week 8, Week 12, Week 16, Week 20, and Week 24).

The total number of days of sick leave over 12 weeks will be analyzed using a negative binomial model, where days of sick leave through week 12 will be the response variable, the independent variables will be the treatment group and region, and the log transformation of duration of employment through week 12 will be used as the offset variable. No missing data will be imputed.

Sample code for negative binomial model:

```
proc genmod data = days_missed;
        class trt01pn region;
```

```
 model aval = trt01pn region/ dist=negbin link=log offset=logdur;
lsmeans trt01pn / pdiff cl exp;
ods output lsmeandiffs=diffs lsmeans=lsmeans;
run;
```

The duration of employment through week 12, days of sick leave through week 12 and employment duration adjusted sick leave days per 12 weeks will be summarized for each treatment group using descriptive statistics.

8.2.21. Hospital Anxiety and Depression Scale (HADS)

The HADS-anxiety and HADS-depression scores will be summarized at each post-baseline visit (Week 4, Week 12, and Week 24).

In addition, at each post-baseline visit until Week 12, the change from baseline will be modelled using the same MMRM model used for the analysis of the primary endpoint. LS mean difference between the dupilumab and placebo group, the corresponding 95% CI of the difference and two-sided p-values will be provided. All p-values except for Week 12 will be considered nominal.

8.2.22. Relationship between variables relating to sleep disturbances and objective and subjective parameters associated with AD.

Relationship between variables (change from baseline and percent change from baseline) relating to sleep disturbances and objective and subjective parameters associated with AD will be explored. Pearson correlation coefficient and p-value will be provided by visit. For the PSG subset, Spearman correlation coefficient and p-value will be provided by visit.

Variables relating to sleep disturbances include:

- Sleep quality NRS
- SE, TST, WASO, and SOL based on actigraphy data
- SE, TST, WASO, and SOL based on sleep diary data
- SCORAD sleep VAS
- PROMIS Sleep Related Impairment SF8a total score
- SE, TST, WASO, and SOL based on the PSG data
- PVT variables: total errors, number of lapses, mean reciprocal response time, mean fastest response time, and mean slowest reciprocal response time
- ANAM variables: mean response time, percentage of correct response and throughput for running memory, mathematical processing and procedural reaction time.

Variables relating to objective and subjective parameters associated with AD include:

- SCORAD total score
- EASI total score
- POEM total score
- DLQI total score
- Peak pruritus NRS

The correlations at baseline, Week 12 and change from baseline to Week 12 will be analyzed by group as defined below:

- Group 1: Sleep quality NRS versus EASI total score, SCORAD total score, POEM total score, DLQI total score, and Peak Pruritus NRS

- Group 2: Sleep diary versus Actigraphy (corresponding measure, SE, TST, WASO, and SOL), POEM total score, SCORAD total score, PROMIS Sleep Related Impairment SF8a total score.
- Group 3: PSG versus Actigraphy (corresponding measure), sleep diary (corresponding measure) on the subset of PSG patients.
- Group 4: Sleep quality NRS versus PVT variables, ANAM variables, WPAI-AD variables and days missed from school or from work.
- Group 5: Sleep quality NRS versus PVT variables, ANAM variables, WPAI-AD variables and days missed from school or from work for Week 24, and Change from baseline to Week 24

For actigraphy and sleep diary data, the weekly averages will be used for correlations. Additionally, for Group 3, the same day assessment (not the average) will be used.

Additionally, relationship between PSG and actigraphy data for the PSG analysis set will be analyzed by providing scatter plots for each of the PSG parameters versus the corresponding actigraphy parameters. For actigraphy data, the same day assessment (not the average) as those of the PSG data will be used.

8.2.23. Multiplicity issues

The overall type-I error rate will be controlled at the two-sided 0.05 level using a sequential testing procedure. In order for any secondary endpoints to be eligible for being declared significant, the primary endpoint must be significant at 0.05 significance level. The secondary endpoints will be tested following the hierarchical testing procedure with a pre-specified order, that is, inferential conclusions about successive secondary endpoints require statistical significance at 0.05 significance level of the previous one.

The following key secondary endpoints will be included in the multiplicity adjustment scheme and tested in the following order:

1. Percent change from baseline to Week 12 in peak pruritus NRS
2. Change from baseline to Week 12 in SCORAD total score
3. Change from baseline to Week 12 in SCORAD sleep VAS
4. Change from baseline to Week 12 in PROMIS Sleep Impairment SF8a score
5. Change from baseline to Week 12 in TST based on actigraphy data
6. Change from baseline to Week 12 in SE based on actigraphy data
7. Change from baseline to Week 12 in WASO based on actigraphy data
8. Change from baseline to Week 12 in SOL based on actigraphy data

Each endpoint will be tested at 0.05 (two-sided) level of significance. If at any step the null statistical hypothesis of no treatment difference is not rejected, the endpoints listed after that step will be reported at nominal level, otherwise will be technically eligible for being declared significant. Regardless of eligibility for being declared significant, all endpoints (non-key secondary and exploratory) will be tested at 0.05 (two-sided) level of significance and reported at nominal level.

9. Safety Analysis

All safety results will be summarized separately within the double-blind treatment period and the overall period (combined double-blind and OLE periods), using the Safety analysis set by actual treatment group and overall.

The observation periods include:

- The **pre-treatment** period is defined as the time from the signed informed consent date up to first administration of the IMP.
- The **randomized on-treatment** period is defined as the time from the first administration of the IMP (on Day 1) to the last administration of the IMP prior to Week 12 + 14 days if the patient does

not enter the OLE period, or until prior to the loading dose at Week 12 if the patient enters the OLE period.

- The **open-label extension on-treatment** period is defined as the time from the loading dose at Week 12 to the last administration of the IMP + 14 days.
- The **treatment-emergent adverse event (TEAE)** period is defined as the time from the first administration of the IMP (on Day 1) to the last administration of the IMP + 14 days (both randomized on-treatment and open-label extension on-treatment periods).

9.1. Adverse Events

The adverse event observation periods include:

- **Pre-treatment adverse events** are adverse events that developed or worsened or became serious from the signed informed consent date up to the first administration of the IMP.
- **Treatment-emergent adverse events** are adverse events that developed or worsened or became serious during the TEAE period.
- **Post-treatment adverse events** are adverse events that developed or worsened or became serious after TEAE period.

The primary focus of adverse event reporting will be on treatment-emergent adverse events. Pre-treatment and post-treatment adverse events will be summarized separately.

All adverse events will be coded to a LLT, PT, HLT, HLGT, and associated primary SOC using the version of Medical Dictionary for Regulatory Activities (MedDRA) currently in effect at Sanofi at the time of each database lock.

Adverse events will be recorded from the time of signed informed consent until the end of the study or the resolution/stabilization of all SAE and AESI.

If an adverse event date/time of onset (occurrence, worsening, or becoming serious) is incomplete, an imputation algorithm will be used to classify the adverse event as pre-treatment, treatment-emergent, or post-treatment. The algorithm for imputing date/time of onset will be conservative and will classify an adverse event as treatment-emergent unless there is definitive information to determine it is pre-treatment or post-treatment. Details on classification of adverse events with missing or partial onset dates are provided in Section 15.4.4.

Adverse event will be summarized by primary SOC, HLGT, HLT, and PT, sorted in alphabetical order for each treatment group using number and percentage of patients experiencing an adverse event. Multiple occurrences of the same event in the same patient will be counted only once in the tables within a treatment phase. The denominator for computation of percentages is the safety analysis set within each treatment group.

Sorting within tables ensures the same presentation for the set of all adverse events within the observation period (pre-treatment, treatment-emergent, and post-treatment). For that purpose, the table of all treatment-emergent adverse events presented by primary SOC and PT (sorted by the internationally agreed SOC order and decreasing frequency of PTs within SOCs in the dupilumab group) will define the presentation order for all other summaries unless otherwise specified. In case of equal frequency regarding PTs, alphabetical order will be used.

On top of the analysis planned below, all TEAEs, all treatment-emergent SAEs, TEAEs leading to permanent treatment discontinuation and TEAEs leading to deaths will be summarized by treatment group and by trial impact (disruption) due to COVID-19.

Overview summaries of the number and percentages of patients within the following categories will be provided by treatment group and overall, for pre-treatment AEs, TEAEs during double-blind treatment period, TEAEs during the study, and post-treatment AEs.

- Any TEAE
- Any study drug related TEAE
- Any Severe TEAE
- Any Serious TEAE
- Any TEAE leading to permanent treatment discontinuation
- Any TEAE of special interest
- Any Serious TEAE of special interest
- Any TEAE leading to death

The overview summaries for TEAEs during the double-blind treatment period and during the overall period will be provided on the safety analysis set and the analysis set with trial impact (disruption) due to COVID-19.

Individual listings (AEs, SAEs, AEs leading to permanent treatment discontinuation, AEs leading to death, AESI and all AEs in treated patients but not considered as randomized) will be provided to support the summary tables based on the safety analysis set.

9.1.1. Incidence of Adverse Events

The following TEAE summaries will be generated for the safety analysis set, for each treatment group and overall.

- All TEAEs presented by primary SOC and PT, showing the number and percentage of patients with at least 1 TEAE, sorted by internationally agreed order of primary SOC and by decreasing incidence of PTs within each SOC in the dupilumab group. This sorting will be applied to all other summaries, unless otherwise specified.
- All TEAEs presented by primary SOC, HLGT, HLT, and PT, showing the number and percentage of patients with at least 1 TEAE, sorted by internationally agreed order of primary SOC. The other levels (HLGT, HLT, PT) will be presented in alphabetical order
- All pre-treatment AE presented by primary SOC and PT, showing number and percentage of patients with at least 1 TEAE, sorted by internationally agreed order of primary SOC and by decreasing incidence of PTs within each SOC in the dupilumab group.
- All treatment-emergent COVID-19 related adverse events presented by primary SOC and PT, showing number and percentage of patients with at least 1 treatment-emergent COVID-19 related AE, sorted by internationally agreed order of primary SOC and by decreasing incidence of PTs within each SOC in the dupilumab group.
- All post-treatment AE presented by primary SOC and PT, showing number and percentage of patients with at least 1 TEAE, sorted by internationally agreed order of primary SOC and by decreasing incidence of PTs within each SOC in the dupilumab group.
- Common TEAEs (PTs with an incidence $\geq$ 5% in any treatment group) by primary SOC, HLGT, HLT, and PT, sorted by internationally agreed order of SOCs. The other levels (HLGT, HLT, PT) will be presented in alphabetic order.

Search criteria for COVID-19 related AE will be provided by SANOFI.

9.1.2. Relationship of Adverse Events to Study Drug

The following TEAE summaries will be generated for the safety analysis set, for each treatment group and overall.

- All TEAEs by relationship, presented by primary SOC, HLGT, HLT and PT, showing the number and percentage of patients with at least 1 TEAE, sorted by internationally agreed order of primary SOC. The other levels (HLGT, HLT and PT) will be presented in alphabetical order

9.1.3. Severity of Adverse Event

The following TEAE summaries will be generated for the safety analysis set, for each treatment group and overall.

- All TEAEs by maximal severity, presented by primary SOC and PT, showing the number and percentage of patients with at least 1 TEAE by severity (i.e., mild, moderate, severe).

9.1.4. Serious Adverse Events

The following TEAE summaries will be generated for the safety analysis set, for each treatment group and overall.

- All serious TEAEs presented by primary SOC, HLGT, HLT and PT, showing the number and percentage of patients with at least 1 serious TEAE, sorted by internationally agreed order of primary SOC. The other levels (HLGT, HLT and PT) will be presented in alphabetical order.
- All serious TEAEs by relationship, presented by primary SOC, HLGT, HLT and PT, showing the number and percentage of patients with at least 1 TEAE, sorted by internationally agreed order of primary SOC. The other levels (HLGT, HLT and PT) will be presented in alphabetical order

9.1.5. Adverse Events Leading to Permanent Treatment Discontinuation

The following TEAE summaries will be generated for the safety analysis set, for each treatment group and overall.

- All TEAEs leading to permanent treatment discontinuation, presented by primary SOC, HLGT, HLT and PT, showing the number and percentage of patients with at least 1 serious TEAE, sorted by internationally agreed order of primary SOC. The other levels (HLGT, HLT and PT) will be presented in alphabetical order

9.1.6. Adverse Events Leading to Study Discontinuation

Not applicable

9.1.7. Adverse Events of Special Interest (AESI)

As defined in Appendix 15.6, Adverse Events of Special Interest (AESI) include:

- Anaphylaxis
- Systemic or severe hypersensitivity reactions
- Malignancy (except in situ carcinoma of the cervix, non-metastatic squamous or basal cell carcinoma of the skin)
- Helminthic infections
- Suicide-related events
- Any type of conjunctivitis or blepharitis (severe or serious)
- Keratitis
- Pregnancy occurring in a female patient or in a female partner of male patient administered IMP/NIMP
- Symptomatic overdose of IMP/NIMP

The following TEAE summaries will be generated for the safety analysis set, for each treatment group and overall.

- All TEAEs of Special Interest, presented by primary SOC, HLGT, HLT and PT, showing the number and percentage of patients with at least 1 serious TEAE, sorted by internationally agreed order of primary SOC. The other levels (HLGT, HLT and PT) will be presented in alphabetical order
- All serious TEAEs of Special Interest, presented by primary SOC, HLGT, HLT and PT, showing the number and percentage of patients with at least 1 serious TEAE, sorted by internationally agreed order of primary SOC. The other levels (HLGT, HLT and PT) will be presented in alphabetical order

9.1.8. Adverse Events Leading to Death

The following TEAE summaries will be generated for the safety analysis set, for each treatment group and overall.

- All TEAEs leading to death (death as an outcome on the eCRF form "Adverse Event" as reported by Investigator), presented by primary SOC, HLGT, HLT and PT, showing the number and percentage of patients with at least 1 serious TEAE, sorted by internationally agreed order of primary SOC. The other levels (HLGT, HLT and PT) will be presented in alphabetical order

9.1.9. Death

The following summaries of deaths will be generated for the safety analysis set, for each treatment group and overall.

- Number and percentage of patients who died during the overall period and cause of death by study period.
- Number and percentage of non-randomized patients or randomized but not treated patients who died.

A listing of all patients who died during the study will be provided.

9.2. Clinical Laboratory Evaluations

Urine human chorionic gonadotropin pregnancy test (as needed for women of childbearing potential) will be performed by the local laboratory, at designated visits. No other laboratory tests are planned in protocol.

Additional tests may be performed at any time during the study as determined necessary by the Investigator or required by local regulations; any clinically significant abnormal lab values will be included in the adverse event analyses.

No laboratory variables will be summarized for this study.

9.3. Vital Sign Measurements

Vital signs including height, weight, heart rate (HR), sitting systolic (SBP) and diastolic blood pressure (DBP), temperature and respiration rate, will be collected at screening visit and all subsequent visits before injection and at 30 minutes (± 10 minutes) post-injection.

The vital sign results will be summarized for the safety analysis set by treatment group using descriptive statistics and for each visit or study assessment (baseline, each-post-baseline time point, last on-treatment value).

The Potentially Clinically Significant Abnormalities (PCSA) at any time during the TEAE period will be summarized for the safety analysis set by treatment group and overall, using number and percentage. The PCSAs are provided in Section 15.7.

All measurements collected during the TEAE period, including values from unscheduled visits, will be considered for the PCSA summaries. The summaries will include patients in the safety analysis set who have at least 1 assessment performed during the TEAE period. When a PCSA definition involves a change from baseline value, patients must also have a baseline value to be included in the summaries.

A listing of patients with at least 1 post-baseline PCSA will be provided and will display the patient's profile over time of all vital sign parameters.

Individual data listings will include the following flags:

- Baseline values will be flagged "B",
- Parameter values reaching a PCSA limit will be flagged (+, or - depending of the direction).

9.4. Physical Examination

There will be a physical examination at baseline visit to examine and assess any abnormalities that may be present, as indicated by the patient's medical history. They will include an assessment of head, heart, lung, abdomen, musculoskeletal, skin and neurological. Results (normal, abnormal or not done) from the assessment will be reported. All deviations from normal will be recorded, including those attributable to the patient's disease.
Any new or worsening finding will be reported as a new adverse event.

The number and percentage of patients with abnormal physical exams at baseline will be summarized for the safety analysis set by treatment group and overall.

9.5. Electrocardiogram

Not applicable.

9.6. Other Safety Data

Not applicable.

10. Pharmacokinetics

Pharmacokinetic parameters are not evaluated in this study.

11. Pharmacodynamics

Pharmacodynamic parameters are not evaluated in this study.

12. Interim Analysis

After all patients have completed the 12 weeks of double-blind, randomized treatment period, the database for this portion of the study will be locked, and the study will be unblinded. Comparative efficacy data will be analyzed following this database lock.

Since the unblinding will be performed after the randomized treatment period, no adjustments to the type I error is needed.

Once all patients complete the last visit post randomization (after OLE period), the final database lock will be performed, and the remaining data will be analyzed.

13. Changes in the Planned Analysis

The changes from the protocol have are listed below:

- PSG analysis set to be used for the analysis of the endpoint based on PSG measures
- OLE analysis set to be used for the exploratory endpoints listed below:
    - Change or percent change (as used in analyses of week 12 data) from baseline to week 24 for primary and secondary endpoints
    - Change or percent change (as used in analyses of week 12 data) from week 12 to week 24 for primary and secondary endpoints
- EASI90 added to the analyses of EASI score.
- Intercurrent events have been added for the primary and key secondary endpoints in order to take into consideration the patients who discontinued due to Covid-19 pandemic, not due to covid-19 pandemic and those patients taking selected prohibited medications and/or rescue medications.
- For the responder analysis on EASI, CMH method have been changed to MMRM model as the analysis first, then if the MMRM model fails to achieve convergence, the logistic regression at each visit will be used in order to add baseline value as one of the covariates. If the logistic regression model fails to achieve convergence, CMH method will be used. Change from baseline of EASI total score will be modelled with a MMRM model.
- Subgroup analysis based on Disease Duration has been removed as disease onset is not collected
- Baseline definition for daily assessments has been updated to include one patient who completed baseline assessment from -13 to -7 instead of -6 to Day 0.
- For the responder analysis on EASI, the analysis model has been changed from GENMOD to GLIMMIX.
- Added some details for retest approach of Psychomotor Vigilance Test.
- Sleep diary illogical data handling rules has been added as an SAP addendum.

14. References

Gittler JK, Shemer A, Suárez-Fariñas M, Fuentes-Duculan J, Gulewicz KJ, Wang CQ, et al. Progressive activation of T(H)2/T(H)22 cytokines and selective epidermal proteins characterizes acute and chronic atopic dermatitis. J Allergy Clin Immunol. 2012 Dec;130(6):1344-54.

Woodruff PG, Modrek B, Choy DF, Jia G, Abbas AR, Ellwanger A, Koth LL, Arron JR, Fahy JV. T-helper type 2-driven inflammation defines major subphoenotypes of aspthma. Am J Respir Crit Care Med 2009 Sep 1; 180(5):388-95

Bieber T. Mechanisms of disease: Atopic Dermatitis. N Engl J Med 2008; 358:1483-94.

Murray CJ, Richards MA, Newton JN, Fenton KA, Anderson HR, Atkinson C et al. UK health performance: findings of the Global Burden of Disease Study 2010. Lancet 2013 Mar 23;381(9871):997-1020.

Zuberbier T, Orlow SJ, Paller AS, Taieb A, Allen R, Hernanz-Hermosa JM, et al. Patient perspectives on the management of atopic dermatitis. J Allergy Clin Immunol 2006 Jul;118(1):226-32.

Dinges DF et al. Cumulative sleepiness, mood disturbance and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night. Sleep 1997, 20 (4): 267 -77

Jeon C, Yan D, Nakamura M, Sekhon S, Bhutani T, Liao W. Frequency and management of sleep disturbance in adults with atopic dermatitis. A systematic review. Dermatol Ther (Heidelb) 2017; Sep; 7(3):349-364

Silverberg JI. Associations between atopic dermatitis and other disorders. 2018; 7 (F1000 Faculty Rev): 303

Ruzicka T, Mihara R. Anti-interleukin-31 receptor A antibody for atopic dermatitis. N Engl J Med 2017; 376:826

Basner M, Dinges, DF. Maximizing sensitivity of the Psychomotor Vigilance Test (PVT) to sleep loss. Sleep. 2011; 34: 581-591.

Wilken JA, Kane R, Sullivan CL, Wallin M, Usiskin JB, Quig ME, et al. The utility of computerized neuropsychological assessment of cognitive dysfunction in patients with relapsing-remitting multiple sclerosis. Mult Scler. 2003;9:119-27.

Reeves DL, Winter KP, Bleiberg J, Kane RL. ANAM® genogram: Historical perspectives, description, and current endeavors. Arch Clin Neuropsychol. 2007;22S:S15-S38.

Hanifin JM, Thurston M, Omoto M, et al. The eczema area and severity index (EASI): assessment of reliability in atopic dermatitis. Exp Dermatol 2001; 10(1): 11-18.

Charman CR, Venn AJ, and Williams, HC. The patient-oriented eczema measure. Arch Dermatol 2004; 140:1513-1519.

Badia X, Mascaro JM, Lozano R, et al. Measuring health-related quality of life in patients with mild to moderate eczema and psoriasis: clinical validity, reliability and sensitivity to change of the DLQI. Br J Dermatol 1999; 141:698-702.

Zigmond AS and Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983;67:361-370.

Herrmann C. International experiences with the hospital anxiety and depression scale – a review of validation data and clinical results. J Psychosom Res 1997;42(1):17-41.

Sanofi Data Validity Verification Plan - LPS15497_Dupistad. Vista Life Sciences, Inc. 2019

Michael J. Thorpy, et al. A randomized study of solriamfetol for excessive sleepiness in narcolepsy. Ann Neurol. 2019 Mar; 85(3): 359–370.

Yu, Lan & Buysse, Daniel & Germain, Anne & Moul, Douglas & Stover, Angela & Dodds, Nathan & Johnston, Kelly & Pilkonis, Paul. (2011). Development of Short Forms from the

PROMIS Sleep Disturbance and Sleep-Related Impairment Item Banks. Behavioral sleep medicine. 10. 6-24. 10.1080/15402002.2012.636266.

Fogelberg, Donald & Vitiello, Michael & Hoffman, Jeanne & Bamer, Alyssa & Amtmann, Dagmar. (2014). Comparison of Self-Report Sleep Measures for Individuals With Multiple Sclerosis and Spinal Cord Injury. Archives of physical medicine and rehabilitation. 96. 10.1016/j.apmr.2014.10.001.

Basner, M. and D. Dinges, Maximizing Sensitivity of the Psychomotor Vigilance Test (PVT) to Sleep Loss. Sleep, 2011 May 1. 34(5): p. 581-591.

Lim, J.L. and D.F. Dinges, Sleep deprivation and vigilant attention, in Molecular and Biophysical Mechanisms of Arousal, Alertness, and Attention. 2008. p. 305-322.

Dorrian, J., N.L. Rogers, and D.F. Dinges, Psychomotor vigilance performance: Neurocognitive assay sensitive to sleep loss, in Sleep deprivation Clinical issues, pharmacology, and sleep loss effects, C.A. Kushida, Editor. 2005, Marcel Dekker: New York. p. 39-70.

Doran, S.M., H.P.A.v. Dongen, and D.F. Dinges, Sustained attention performance during sleep deprivation: Evidence of state instability. Archives Italiennes de Biologie, 2001. 139: p. 253-267.

Balkin, T.J., et al., Comparative utility of instruments for monitoring sleepiness-related performance decrements in the operational environment. Journal of Sleep Research, 2004. 13: p. 219-227.

15. Appendices

15.1. Summary of Statistical Analyses

Endpoint	***Analysis Set***	***Primary analysis***	***Supportive analysis***	***Subgroup analysis***	***Estimands***
Primary endpoint					
Sleep Quality NRS: the percent change from baseline to week 12 (dupilumab, placebo)	*mITT*	*MMRM, without imputation of missing values*	*Sensitivity analysis: Set missing to those weeks with 3 or less available sleep NRS data. No imputation will be made for the missing data. Using the same MMRM model*	*Yes*	*Yes*
Secondary endpoints					
Continuous variables	*mITT*	*Same approach as primary endpoints*	*No*	*No*	*Yes (for key secondary endpoints only)*
Binary variables	*mITT*	*Cochran-Mantel-Haenszel test, define missing as non-responder*	*No*	*No*	*No*
Exploratory endpoints					
Continuous variable	*mITT*	*Same approach as primary endpoints*	*No*	*No*	*No*
Days missed from school or work to week 12	*mITT*	*Negative Binomial Model*	*No*	*No*	*No*
PVT & ANAM	*mITT*	*Same approach as primary endpoint*	*No*	*No*	*No*

Endpoint	***Analysis Set***	***Primary analysis***	***Supportive analysis***	***Subgroup analysis***	***Estimands***
Binary variables	*mITT*	*Same approach as secondary endpoints*	*No*	*No*	*No*
Sleep disturbances (eg, TST, WASO) and parameters associated with AD (eg, EASI, SCORAD, pruritus NRS, DLQI, EASI, skin pain NRS, and skin burn NRS, sensitivity to touch NRS)	*mITT*	*Correlation*	*No*	*No*	*No*
All efficacy and Patient-Report Outcome (PRO) measures: from week 12 to week 24 and from baseline to week 24: *will be analyzed separately for those originally randomized to dupilumab, and those originally randomized to placebo (and switched to dupilumab during the open label extension)*	*mITT*	*Summary stats only*	*No*	*No*	*No*

15.2. Schedule of Study Procedures

Study Procedure													
Visit (V)	Screen-ing 1[g]	Screen-ing 2[g]	Baseline V1	V2	V3	V4	V5	Primary End point Visit V6	V7	V8	End of Study/ Early discontin-uation V9[h]	UnSch Visit	Notes
Week (W)			W0	W1	W2	W4	W8	W12	W16	W20	W24		
Day (D) Visit window (±d)	D (-28 to -8)	D (-7 to -1)	D1	D8 ±3d	D15 ±3d	D29 ±3d	D57 ±3d	D85 ±3d	D113 ±3d	D141 ±3d	D169 ±3d		
Enrollment Screening/Baseline:													
Inclusion and exclusion criteria	X		X										
Informed consent	X												
Demography	X												
Concurrent illness	X		X										
IVRS call	X		X		X	X	X	X	X	X	X		
Study Treatment/Concomitant Medication:													
Study drug dispensation			X[a]		X[b]	X	X	X[c]	X	X			
Concomitant medications and procedures	X	X	X	X	X	X	X	X	X	X	X	X	
Efficacy:													
Physician Assessments:													

Study Procedure													
Visit (V)	Screen-ing 1[g]	Screen-ing 2[g]	Baseline V1	V2	V3	V4	V5	Primary End point Visit V6	V7	V8	End of Study/ Early discontin-uation V9[h]	UnSch Visit	Notes
Week (W)			W0	W1	W2	W4	W8	W12	W16	W20	W24		
Day (D) Visit window (±d)	D (-28 to -8)	D (-7 to -1)	D1	D8 ±3d	D15 ±3d	D29 ±3d	D57 ±3d	D85 ±3d	D113 ±3d	D141 ±3d	D169 ±3d		
IGA and BSA (for eligibility purpose only)	X		X										
EASI	X		X	X	X	X	X	X	X	X	X	X	
SCORAD	X		X	X	X	X	X	X	X	X	X	X	
Patient Assessments (Daily including week preceding baseline visit):													
Sleep Quality NRS[d]		X	X	X	X	X	X	X	X	X	X	X	
Sleep Diary[d]		X	X	X	X	X	X	X	X	X	X	X	
Peak pruritus NRS, Skin pain NRS, Skin sensitivity to touch, Skin burn NRS[d]		X	X	X	X	X	X	X	X	X	X	X	
Actigraphy[d]		X	X	X	X	X	X	X	X	X	X	X	
Patient Assessment (at clinic):													
PROMIS Sleep Related Impairment SF 8a			X	X	X	X	X	X	X	X	X	X	
Epworth Sleepiness Scale (ESS)			X			X		X			X	X	

Study Procedure													
Visit (V)	Screening 1[g]	Screening 2[g]	Baseline V1	V2	V3	V4	V5	Primary End point Visit V6	V7	V8	End of Study/ Early discontinuation V9[h]	UnSch Visit	Notes
Week (W)			W0	W1	W2	W4	W8	W12	W16	W20	W24		
Day (D) Visit window (±d)	D (-28 to -8)	D (-7 to -1)	D1	D8 ±3d	D15 ±3d	D29 ±3d	D57 ±3d	D85 ±3d	D113 ±3d	D141 ±3d	D169 ±3d		
POEM			X			X		X			X	X	
DLQI			X			X		X			X	X	
HADS			X			X		X			X	X	
ACQ-5, AR-VAS			X					X			X	X	
WPAI-AD			X			X	X	X			X	X	
Assess missed school/work days			X			X	X	X	X	X	X	X	
Center assisted Assessment:													
Psychomotor Vigilance Test (PVT)	X		X					X			X	X	
Neurocognitive test (Automated Neuropsychological Assessment Metrics-ANAM)	X		X					X			X	X	
Polysomnography (sub-study, select sites)[e]			X[e]					X					
Skin Photography[f]			X	X	X	X	X	X	X	X	X	X	

Study Procedure													
Visit (V)	Screening 1[g]	Screening 2[g]	Baseline V1	V2	V3	V4	V5	Primary End point Visit V6	V7	V8	End of Study/ Early discontinuation V9[h]	UnSch Visit	Notes
Week (W)			W0	W1	W2	W4	W8	W12	W16	W20	W24		
Day (D) Visit window (±d)	D (-28 to -8)	D (-7 to -1)	D1	D8 ±3d	D15 ±3d	D29 ±3d	D57 ±3d	D85 ±3d	D113 ±3d	D141 ±3d	D169 ±3d		
Sleep diary uploading			X	X	X	X	X	X	X	X	X	X	
Safety:													
Weight	X										X		
Height	X												
Vital signs	X		X	X	X	X	X	X	X	X	X	X	
Physical examination			X										
Adverse events	X	X	X	X	X	X	X	X	X	X	X	X	
Laboratory Testing:													
Urine Pregnancy Test (WOCBP only)	X		X			X	X	X	X	X	X	X	

Abbreviations: ACQ-5 = asthma control questionnaire; AR-VAS = Allergic Rhinitis-visual analog scale; BSA = body surface area; D = day; DC = discontinuation; DLQI = dermatology life quality index; EASI = eczema area and severity index; HADS = hospital anxiety and depression scale; IGA = Investigator's global assessment; NRS = numerical rating scale; POEM = patient oriented eczema measure; PROMIS = patient-reported outcomes measurement information system; SCORAD = SCORing atopic dermatitis; UnSch = Unscheduled; V = visit; WOCBP = women of childbearing potential; WPAI-AD = work productivity and activity impairment – Atopic Dermatitis version; w = week

a Loading dose dupilumab arm 600 mg (300 mg x 2 ie, 2 syringes); 2 placebo syringes for the placebo arm.

b Regular dose dupilumab arm 300 mg (1 syringe) and 1 syringe for placebo arm, for q2w injections at sites and for home injections.

c In order to protect the blind prior to this point to reduce bias in the open label period, at the beginning of open label (week 12), for loading dose, dupilumab arm will have 1 x 300 mg syringe and 1 placebo syringe; placebo arm will need to remain 2 syringes, 300 mg each; then subsequent dose will be 1 syringe of 300 mg for all patients in both arms.

d For patients who do not participate to the PSG sub-study: Patients eligible to continue, except for baseline severity criteria, will complete baseline assessments from Day -7 to Day -1 (right before Day 1); daily assessment thereafter until week 4; then only the week before a clinical visit until last study visit. For patients who participate to the PSG sub-study: same procedures as for the other patients, except that the sleep quality NRS and the peak pruritus NRS will be completed for 5 days, from Day -7 to Day -3 prior to baseline.

e Each patient who participates in this sub-study will have a total of 3 overnight PSGs: at baseline time point, 2 (preferably) consecutive overnight PSGs to control for "first night effect" and collect baseline data, and 1 more at week 12.

f Not mandatory, site professional will be taking the photographs with patient's consent; Site staff will take the photos of the same lesioned areas for each visit, starting from the baseline visit. Patients will be asked to sign a separate photography informed consent at Day 1.

g If preferred by the patient and the investigator for logistical reasons, screening can be done during a single on-site visit. In that case, all procedures planned at screening 1 and screening 2 should be done during this single screening visit. Importantly, this single screening visit will have to be done at least 7 days before the baseline visit, as patients should wear the actigraph and complete the diary for 7 days before baseline.

h V9 assessments are applied to patients who prematurely discontinue the study, complete the treatment (end of treatment), and complete the study (end of study).

15.3. Missing Efficacy Data

15.3.1. Handling of missing data and data whose interpretations have changed

The following will be applied to all efficacy endpoints:

- Patients will be encouraged to stay in the study if they discontinue the IMP, even due to Covid-19 pandemic (IcEv1and IcEv2).
    - If discontinuation or interruption of study treatment due to Covid-19, off-treatment data will be set to missing. The assessments after patients resume treatment will be used.
    - If discontinuation or interruption of study treatment not due to Covid-19, all data collected following schedule after the treatment discontinuation will be used in the analysis.
- Taking selected prohibited medications that affect sleep (Protocol Section 6.5.1), all sleep related assessments (sleep quality NRS, SCORAD sleep VAS and total score, PROMIS sleep impairment SF8a total score, ESS, and actigraphy data) on that corresponding night will be set to missing. AD assessments will be included in the analyses.
- Taking selected prohibited medications that affect AD (Protocol Section 6.5.1), all AD assessments will be set to missing after the medication usage.
- Taking rescue medications for AD (ie. to control intolerable AD symptoms) is allowed. For the purpose of efficacy analysis, all AD assessments will be set to missing after the medication usage.
- If a patient discontinues from treatment or from the study, or is taking rescue or selected prohibited medications, the patient will be counted as a non-responder for the time points after discontinuation or rescue/prohibited medication start for all binary endpoints.

15.3.2. Handling of missing data for Patient-Reported Outcomes (PROS)

<u>Patient-Oriented Eczema Measure (POEM)</u>

- If only one question is left unanswered, this is scored 0 and the scores are summed and expressed as usual out of a maximum of 28
- If more than one question is left unanswered, the questionnaire is not scored
- If 2 or more response options are selected for a single question, the response option with the highest score will be recorded.

<u>PROMIS Sleep Related Impairment SF8a Total Score</u>

The total score for PROMIS is the sum of 8 item scores.

- For any missing item score at a specific time point, the mean of the remaining non-missing item scores for that patient at this specific time point is used to impute the missing item score(s). In these cases, the total score is calculated as the sum of the observed and the imputed item scores

<u>Dermatology Life Quality index (DLQI)</u>

- If only one question is left unanswered, this is scored 0 and the scores are summed and expressed as usual out of a maximum of 30
- If 2 or more questions are left unanswered, the questionnaire is not scored
- If 2 or more response options are selected for a single question, the response option with the highest score should be recorded
- Related to question 7:
    - If question 7 is answered "yes", this is scored 3
    - If question 7 is answered "no" but then either "a lot" or "a little" is ticked, this is then scored 2 or 1 (respectively).
    - If "not relevant" is ticked, this is scored 0
    - If question 7 is answered "no", but the second half is left incomplete, the score will remain 0.

Juniper Asthma Control Questionnaire (ACQ-5)

- If only one question is left unanswered at a specific time point,
    - If the questionnaire from the previous time point is complete, it is interpolated (pro-rated) using the completed questions from the previous time point
    - If the questionnaire from the previous time point is not complete, the missing value will be imputed as the average of the completed questions within the current visit.
- If 2 or more questions are left unanswered, the global score is invalid and is considered as missing

Hospital Anxiety and Depression Scale (HADS)

- If only one question is left unanswered in the Anxiety subscale (sum of all the A questions) or Depression subscale (sum of all D questions) at a specific time point, a score can be computed using the person specific mean of the corresponding subscale at this specific time point
- If 2 or more questions are left unanswered in the Anxiety subscale or Depression subscale, the score is invalid and is considered as missing in the corresponding subscale.

Other Patient-Reported Outcomes

For all other questionnaires, if one or more questions are left unanswered, the questionnaire is not scored.

15.3.3. Handling of missing data for Clinician Reported Outcomes (CLINROS)

No imputation will be made for missing or incomplete data.

15.4. Missing Safety Data

For categorical variables, patients with missing data are not included in calculations of percentages unless otherwise specified. When relevant, the number of patients with missing data is presented.

15.4.1. Handling of missing age

If age is missing but year of birth is collected, then age will be derived as date of informed consent signed minus year of birth and June 30th, keeping the integer part of the result.

15.4.2. Handling of computation of treatment duration if investigational medicinal product end of treatment date is missing

For the calculation of the treatment duration, the date of the last dose of IMP is equal to the date of last administration reported on the e-CRF "Treatment status" page. If this date is missing, the last available administration date in the "Exposure" form will be used.

The last dose intake should be clearly identified in the eCRF and should not be approximated by the last returned package date

15.4.3. Handling of medication missing/partial dates

No imputation of medication start/end dates or times will be performed. If a medication date or time is missing or partially missing and it cannot be determined whether it was taken prior or concomitantly, it will be considered a prior, concomitant and post-treatment medication.

15.4.4. Handling of adverse events with missing or partial date/time of onset

Missing or partial adverse event onset dates and times will be imputed so that if the partial adverse event onset date/time information does not indicate that the adverse event started prior to treatment or after the treatment-emergent adverse event period, the adverse event will be classified as treatment-emergent. No imputation of adverse event end dates/times will be performed. These data imputations are for categorization purpose only and will not be used in listings. No imputation is planned for date/time of adverse event resolution.

15.4.5. Handling of adverse events when date and time of first investigational medicinal product administration is missing

When the date and time of the first IMP administration is missing, all adverse events that occurred on or after the day of randomization should be considered as treatment-emergent adverse events. The exposure duration should be kept as missing.
The last dose intake should be clearly identified in the eCRF and should not be approximated by the last returned package date.

15.4.6. Handling of missing assessment of relationship of adverse events to investigational medicinal product

If the assessment of the relationship to IMP is missing, then the relationship to IMP has to be assumed and the adverse event considered as such in the frequency tables of possibly related adverse events, but no imputation should be done at the data level.

15.4.7. Handling of missing severity/grades of adverse events

If the severity/grade is missing for 1 of the treatment-emergent occurrences of an adverse event, the maximal severity on the remaining occurrences will be considered. If the severity is missing for all the occurrences, a "missing" category will be added in the summary table.

15.4.8. Handling of potentially clinically significant abnormalities

If a patient has a missing baseline he will be grouped in the category "normal/missing at baseline."

For PCSAs with 2 conditions, one based on a change from baseline value or a normal range and the other on a threshold value, with the first condition being missing, the PCSA will be based only on the second condition.

15.5. Visit Windows

The following steps will decide how the scheduled and/or unscheduled visits will be used in the analyses on efficacy variables and the by-visit summaries for safety variables.

Step 1: A scheduled measurement (described in Appendix 15.2) will be used if it is available; otherwise, an unscheduled measurement (including the end of treatment/study visit for those prematurely discontinued) will be used if it happens to be on the same date as the date of the scheduled visit.

Step 2: After Step 1, if there is still no measurement for a given parameter at a scheduled visit, the analysis window below "[in brackets]" will be applied to re-allocate a post-baseline unscheduled measurement to a scheduled measurement.

Table 15.5.1 Analyses window definition

Scheduled visit post baseline	Targeted study day	Analysis window			
		PSG	Weight	PVT, ANAM, ACQ-5 and AR-VAS	All other endpoints/parameters
Week 1 (Visit 2)	7				[1, 11]
Week 2 (Visit 3)	14				[12, 21]
Week 4 (Visit 4)	28				[22, 42]
Week 8 (Visit 5)	56				[43, 70]
Week 12 (Visit 6)	84	[1,~]		[1, 126]	>=71 for patients who did not enter OLE. [71, 1st dose date of OLE[for patients entered OLE.
Week 16 (Visit 7)	112				[First dose of OLE, 126]
Week 20 (Visit 8)	140				[127, 154]

Scheduled visit post baseline	Targeted study day	Analysis window			
		PSG	Weight	PVT, ANAM, ACQ-5 and AR-VAS	All other endpoints/parameters
Week 24 (Visit 9)	168		[1,~]	[127,~]	[155, ~]

Study days are calculated from the day of first administration of IMP; the day of first administration of IMP (or the day of randomization if not exposed) is Day 1.

For endpoints with daily assessments until Week 4, Week 3 will be summarized. Week 3 visit is assigned to records which lie on the 7 days following the final day of Week 2. If this 7 day window overlaps with already assigned Week 4 records then Week 3 will have a duration of less than 7 days.

After applying the above time windows, if multiple assessments are associated to the same time point, the closest from the targeted study day will be used. In case of equality, the last measurement will be used. Re-allocated scheduled visits (i.e., visit numbers) should be sequential if ordered by the date of measurement.

After Step 2, if there is still no measurement for a given parameter at a scheduled visit, data is considered missing.

For endpoints where measurements were only scheduled once for post-baseline (e.g., weight at week 24 and PSG at week 12), the measurement that is closest to the target day will be used.

15.6. Search Criteria for AESI

Table 15.6.1 Search Criteria for AESI

AESI	Search Criteria
Anaphylaxis	SMQ: Anaphylaxis [Narrow]
Systemic or severe hypersensitivity reactions	SMQ: Hypersensitivity [Narrow]
Malignancy (except in situ carcinoma of the cervix, non-metastatic squamous or basal cell carcinoma of the skin)	Sub-SMQ (20000091) – Malignant or unspecified tumors
Helminthic infections	HLTs of "Helminthic infections NEC"
Suicide-related events	PTs of • Completed suicide • Suicidal ideation • Depression suicidal • Suicidal behavior • Suicide attempt
Any type of conjunctivitis or blepharitis (severe or serious)	1. Narrow conjunctivitis is defined as the following PTs • Conjunctivitis • Conjunctivitis allergic • Conjunctivitis bacterial • Conjunctivitis viral • Atopic Keratoconjunctivitis 2. Broad conjunctivitis is defined as the following PTs • Conjunctivitis, Conjunctivitis allergic • Conjunctivitis bacterial • Conjunctivitis viral • Atopic Keratoconjunctivitis • Blepharitis • Dry eye • Eye irritation • Eye pruritus • Lacrimation increased • Eye discharge • Foreign body sensation in eyes • Photophobia, Xerophthalmia • Ocular hyperaemia • Conjunctival hyperaemia 3. Conjunctivitis is defined as the following PTs • Conjunctivitis • Non-infective conjunctivitis

	• Conjunctivitis allergic • Conjunctivitis bacterial • Conjunctivitis viral • Eye irritation • Eye inflammation
Keratitis	PTs of • Keratitis, Allergic keratitis • Ulcerative keratitis • Atopic keratoconjunctivitis • Herpes ophthalmic • Ophthalmic herpes simplex
Pregnancy occurring in a female patient or in a female partner of male patient administered IMP/NIMP	"Pregnancy" or "Partner Pregnancy" checked in eCRF form "Pregnancy"
Symptomatic overdose of IMP/NIMP	"Overdose of Study Treatment" or "Overdose of NIMP" checked and "Symptomatic overdose" checked in eCRF form "Overdose"

15.7. Potentially Clinically Significant Abnormalities (PCSA) Criteria

Table 15.7.1 Potentially Clinically Significant Abnormalities Criteria

CRITERIA for POTENTIALLY CLINICALLY SIGNIFICANT ABNORMALITIES for phase 2/3 studies (oncology excepted)

(From BTD-009536 May 21, 2014)

Parameter	PCSA	Comments
HR	≤50 bpm and decrease from baseline ≥20 bpm ≥120 bpm and increase from baseline≥20 bpm	To be applied for all positions (including missing) except STANDING.
SBP	≤95 mmHg and decrease from baseline ≥20mmHg ≥160 mmHg and increase from baseline ≥20 mmHg	To be applied for all positions (including missing) except STANDING.
DBP	≤45 mmHg and decrease from baseline ≥10 mmHg ≥110 mmHg and increase from baseline ≥10 mmHg	To be applied for all positions (including missing) except STANDING.
Weight	≥5% increase from baseline ≥5% decrease from baseline	FDA Feb 2007

15.8. Patient-Reported Outcomes (PROS)

Patient oriented eczema measure (POEM)

The POEM is a 7-item, validated questionnaire used in clinical practice and clinical trials to assess frequency of disease symptoms in children and adults (9). The format is a response to 7 items (dryness, itching, flaking, cracking, sleep loss, bleeding, and weeping) based on frequency during the past week (ie, 0 = no days, 1 = 1 to 2 days, 2 = 3 to 4 days, 3 = 5 to 6 days, and 4 = all days) with a scoring system of 0 to 28; the total score reflects disease-related severity.

POEM score will be directly provided by KONEKSA.

PROMIS Sleep Related Impairment Short Form 8a

The PROMIS is a DSM-5, Level 2, sleep disturbance measure. This study uses the 8-item PROMIS Sleep Related Impairment Short Form that assesses the domain of sleep related impairment in the past 7 days in individuals age 18 and older.

Each item asks the patient to rate the severity of the patient's sleep related impairment during the past 7 days. Each item on the measure is rated on a 5-point scale (1 = not at all; 2 = a little bit; 3 = somewhat; 4 = quite a bit; and 5 = very much) with a range in score from 8 to 40 with higher scores indicating greater severity of sleep impairment.

Locate the applicable score conversion in Table *15.8.1*, and use this table to translate the total raw score or imputed total score into a PROMIS T-score for each participant. The T-score rescales the raw score into a standardized score with a mean of 50 and a standard deviation of 10. Therefore, a person with a T-score of 40 is one SD below the mean. The standardized T-score is reported as the final score for each participant.

Table 15.8.1 PROMIS Sleep-Related Impairment conversion table

Raw Score	T-score	Raw Score	T-score
8	30	25	60.3
9	35.1	26	61.3
10	38.7	27	62.3
11	41.4	28	63.3
12	43.6	29	64.3
13	45.5	30	65.3
14	47.3	31	66.3
15	48.9	32	67.3
16	50.3	33	68.4
17	51.6	34	68.5
18	52.9	35	70.7
19	54	36	71.9
20	55.1	37	73.3
21	56.1	38	75
22	57.2	39	76.9
23	58.2	40	80
24	59.3		

PROMIS score will be directly provided by KONEKSA.

Epworth Sleepiness Scale (ESS)

The ESS is a self-administered questionnaire with 8 questions. Respondents are asked to rate, on a 4-point scale (0-3), their usual chances of dozing off or falling asleep while engaged in eight different activities. Most people engage in those activities at least occasionally, although not necessarily every

day. The ESS score (the sum of 8 item scores, 0-3) can range from 0 to 24. The higher the ESS score, the higher that person's average sleep propensity in daily life, or their daytime sleepiness.

ESS score will be directly provided by KONEKSA.

Sleep Quality NRS

The Sleep Quality NRS is a simple assessment tool that patients will use to report the quality of patient's last night sleep. Patients will be asked the following question:

"On a scale of 0 to 10, with 0 being "worst possible sleep" and 10 being the "best possible sleep", select the number that best describe the quality of your sleep last night"

Patients will complete the rating scale daily including 7 days immediately preceding the baseline visit (or daily for 5 days (from Day -7 to Day -3 prior to baseline) for patients participating in the PSG sub-study), and daily for 4 weeks (through visit 4), then daily during the week before each planned visit.

Instructions: Please complete the following question upon awakening for the day.

Select the number that best describe the quality of your sleep last night.

0	1	2	3	4	5	6	7	8	9	10
Worst possible sleep										Best possible sleep

Patient chooses a value from 0 to 10 in the only question of the scale and that value is the global score of the scale.

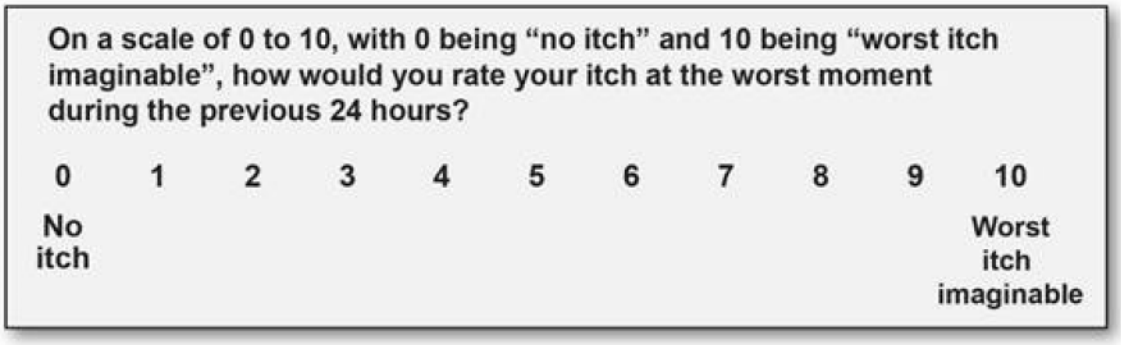
Sleep Quality NRS score will be directly provided by KONEKSA.

Peak Pruritus NRS

The Peak Pruritus NRS is a simple assessment tool that patients will use to report the intensity of their pruritus (itch) during a daily recall period. Patients will be asked the following question:

"On a scale of 0 to 10, with 0 being 'no itch' and 10 being the 'worst itch imaginable', how would you rate your itch at the worst moment during the previous 24 hours?"

Patients will complete the rating scale daily for 7 days prior to baseline (or daily for 5 days (from Day -7 to Day -3 prior to baseline) for patients participating in the PSG sub-study), and daily for 4 weeks (through visit 4), and then daily during the week before each planned visit.



Patient chooses a value from 0 to 10 in the only question of the scale and that value is the global score of the scale.

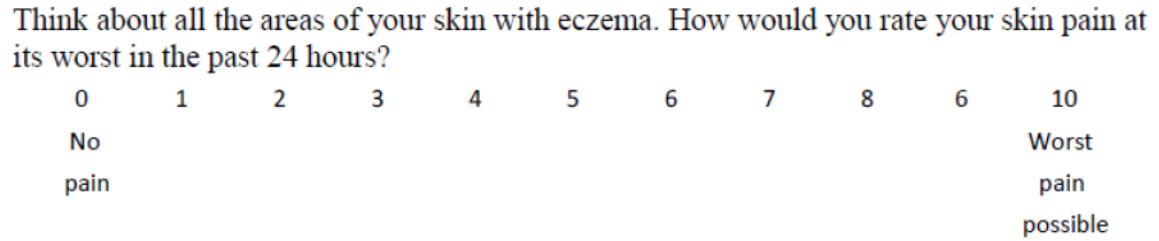
Peak Pruritus NRS score will be directly provided by KONEKSA.

Skin Pain NRS

Patients will be asked to rate their skin pain using a 0 to 10 NRS. Patients will be asked the following question:

"Think about all the areas of your skin with eczema. How would you rate your skin pain at its worst in the past 24 hours?"

Patients will be asked to complete this assessment scale daily from 7 days prior to baseline through week 4 (visit 4), and then daily during the week before each planned visit.



Patient chooses a value from 0 to 10 in the only question of the scale and that value is the global score of the scale.

Skin Pain NRS score will be directly provided by KONEKSA.

Skin Sensitivity to touch NRS

Patients will answer the question: "Think about all the areas of your skin with eczema. How sensitive was your skin at its worst in the past 24 hours?"

Patients will be asked to complete this assessment scale daily from 7 days prior to baseline through week 4 (visit 4), and then daily during the week before each planned visit.

Think about all the areas of your skin with eczema. How sensitive was your skin at its worst in the past 24 hours?

0	1	2	3	4	5	6	7	8	9	10
Normal										Extremely sensitive

Patient chooses a value from 0 to 10 in the only question of the scale and that value is the global score of the scale.

Skin Sensitivity to Touch NRS score will be directly provided by KONEKSA.

Skin Burn NRS

Patients will answer the question: "Think about all the areas of your skin with eczema. How much did your skin burn at its worst in the past 24 hours?"

Patients will be asked to complete this assessment scale daily from 7 days prior to baseline through week 4 (visit 4), and then daily during the week before each planned visit.

Think about all the areas of your skin with eczema. How much did your skin burn at its worst in the past 24 hours?

0	1	2	3	4	5	6	7	8	9	10
Not at all										Very much

Patient chooses a value from 0 to 10 in the only question of the scale and that value is the global score of the scale.

Skin Burn NRS score will be directly provided by KONEKSA.

Sleep diary

Collected daily including 7 days immediately preceding the baseline visit, and daily for 4 weeks (through visit 4), then daily during the week before each planned visit.

A sleep diary is designed to gather information about patient's daily sleep pattern. Sleep diary will measure night-time sleep assessments. If possible, it should be completed within 1 hour of getting out of bed in the morning. The following variables will be derived from sleep diary:

WASO = Time awake after initial sleep onset but before the final awakening

This is data collected from Question 3 of the Sleep Diary: "Considering all the times you woke up last night, how much time were you awake in total?"

TST = Time of waking up for the day minus Time of falling sleep minus WASO

SOL = Time of falling sleep – Time of trying to fall sleep

$$SE = \frac{TST}{\text{Time of waking up for the day} - \text{Time of trying to fall sleep}} \times 100\%$$

Sleep diary data will be directly provided by KONEKSA.

Due to illogical data responses to the Sleep diary questionnaires, a SAP addendum with data handling rules to be applied will be provided in a separate document.

Dermatology Life Quality Index (DLQI)

The DLQI is a 10-item, validated questionnaire used in clinical practice and clinical trials to assess the impact of AD disease symptoms and treatment on Quality of Life (QoL). The format is a simple response (0 to 3 where 0 is "not at all' and 3 is "very much") to 10 questions, which assess QoL over the past week, with an overall scoring system of 0 to 30; a high score is indicative of a poor QoL.

The DLQI score will be directly provided by KONEKSA.

Juniper Asthma Control Questionnaire (ACQ-5) [among those reporting asthmas at baseline]

The 5-question version of the Juniper asthma control questionnaire (ACQ) is a validated questionnaire to evaluate asthma control. The questionnaire will be administered only to the subset of patients with a medical history of asthma, at Baseline (Week 0), Week 12 and Week 24.

Each question is a 7-point scale (0 = no impairment to 6 = maximum impairment). A global score is calculated: the questions are equally weighted, and the ACQ-5 score is the mean of the 5 questions and, therefore, between 0 (totally controlled) and 6 (severely uncontrolled).
On the 7-point scale of the ACQ-5, a change or difference in score of 0.5 is the smallest that can be considered clinically important, corresponding to the Minimal Clinically Important Difference (MCID) defined by the developer.

ACQ-5 score will be directly provided by KONEKSA.

Allergic Rhinitis Visual Analog Scale (AR-VAS) [among those reporting allergic rhinitis at baseline]

The AR-VAS is a validated instrument for the documentation of symptoms and therapy monitoring in allergic rhinitis, Patients with comorbid AR will be provided with an ungraded VAS and will be asked to place a mark on the scale to indicate the severity of AR symptoms. Patients will be asked:

"Overall how much are your allergic symptoms bothering you today?".

The VAS extremities will be noted as 'Not at all bothersome" to the left and 'Extremely bothersome' to the right. This questionnaire will be administered at Baseline, Week 12 and Week 24 to the patients with AR at baseline and to patients who develop AR during the study.

The AR-VAS score will be directly provided by KONEKSA.

Work Productivity and Activity Impairment Questionnaire: Atopic Dermatitis (WPAI-AD)

The WPAI-AD is designed to assess the impact of AD on the patient's productivity. The WPAI is a 6-item, validated questionnaire to measure impairments in work and activities over a 7-day recall period. The WPAI-AD outcomes are expressed as impairment percentages, with higher numbers indicating greater impairment and less productivity.

The WPAI-AD questionnaire will be directly provided by KONEKSA.

Days missed from school or from work due to AD

Patients will complete questions about their current employment status and days missed from work/school due to their AD for all patients who are working/going to school.

Other assessments will be performed:

- Duration of employment through week 12 = sum of all employment intervals (specific assessment date – last assessment date) from the first visit till the last assessment of employment status on or prior to week 12 + 1 day. For the first post-baseline assessment, if no baseline assessment is recorded then randomization date should be used as last assessment date. Duration of employment will be calculated disregarding full or partial employment status.
- Days of sick leave through week 12 is the sum of all missed work/school days up to the last assessment date on or prior to week 12. For the first post-baseline assessment, if no baseline assessment is recorded and the number of missed days is higher than imputed employment duration then the number of missed days will be equal to employment duration.
- Employment duration adjusted sick leave days per 12 weeks will be used for treatment comparison. It is only applied for patients who are fully or partially employed. If days of sick leave through week 12 value is missing for a specific assessment, the length of this assessment interval will be deducted from the duration of employment through week 12.

$$\text{Employment duration adjusted sick leave days per 12 weeks} = \frac{\text{Days of sick leave through week 12}}{\text{Duration (in days) of employment through week 12}} * 12 \text{ weeks (in days)}$$

Days missed from school or from work due to AD will be directly provided by KONEKSA.

Hospital Anxiety and Depression Scale (HADS)

The HADS is an instrument for assessing symptoms of anxiety and depression in non-^psychiatric populations; repeated administration also provides information about changes to a patient's emotional state. The HADS consists of 14 items, 7 each for anxiety and depression symptoms; possible scores range from 0 to 21 for each subscale.

The HADS scores will be directly provided by KONEKSA.

15.9. Clinician-Reported Outcomes (CLINROS)

15.9.1. Clinician Assessments

SCORing of Atopic Dermatitis (SCORAD)

The SCORAD is a validated tool used in clinical research and clinical practice that was developed to standardize the evaluation of the extent and severity of AD. There are 3 components to the assessment:

- A = extent or affected BSA
- B = severity
- C = subjective symptoms

The extent of AD is assessed as a percentage of each defined body area and reported as the sum of all areas, with a maximum score of 100% (assigned as 'A' in the overall SCORAD calculation). The severity of 6 specific symptoms of AD (redness, swelling, oozing/crusting, excoriation, skin thickening/lichenification, dryness) is assessed using the following scale: none (0), mild (1), moderate (2), or severe (3) (for a maximum of 18 total points, assigned as 'B' in the overall SCORAD calculation). Subjective assessment of itch and sleeplessness is recorded for each symptom by the patient or relative on a visual analog scale (VAS, where 0 is no itch (or sleeplessness) and 10 is the worst imaginable itch (or sleeplessness), with a maximum possible score of 20 (assigned as 'C' in the overall SCORAD calculation). The SCORAD is calculated as: A/5 + 7B/2 + C where the maximum is 103.

The SCORAD total score, itch and sleeplessness VAS will be directly provided by KONEKSA.

Eczema Area and Severity Index (EASI)

The EASI is a validated measure used in clinical practice and clinical trials to assess the severity and extent of AD (18). 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 each be assessed for severity by the investigator or designee on a 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, trunk, upper limbs, and lower limbs, and converted to a score of 0 to 6. In each body region, the area 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%).
For each body region (head, trunk, upper limbs, and lower limbs), EASI score = (sum of the four AD disease characteristics scores) multiply (area of AD involvement score). The total EASI sore is the weighted total of the body region EASI scores using the weights 10% = head, 20% = upper

limbs, 30% = trunk, 40% = lower limbs.

The EASI is a composite index with scores ranging from 0 to 72. EASI score will be directly provided by KONEKSA.

- EASI50 is achieved when percent reduction from baseline in EASI score >=50%.
- EASI75 is achieved when percent reduction from baseline in EASI score >= 75%.
- EASI90 is achieved when percent reduction from baseline in EASI score >= 90%.

15.9.2. Center Assisted Assessments

Psychomotor Vigilance Test (PVT)

The PVT is a widely used measure of behavioral alertness, having a high sensitivity to sleep deprivation. There are several versions of the PVT, but the standard form lasts 10 minutes and assesses sustained attention/vigilance by recording response times to stimuli that occur at varying inter-stimulus intervals. The PVT assesses vigilance by sampling many responses to stimuli appearing at random inter-stimulus intervals (eg. Between 2 seconds and 10 seconds) occurring over a period of time (i.e. 10 minutes). The patient is instructed to monitor an area on a computer/iPAD/electronic screen and then to press a response button as soon as the stimulus appears in order to keep the response time as low as possible without pressing the button prematurely.

A quality control check will be performed for the PVT tests based on pre-defined QC thresholds. If the PVT test exceeds the QC thresholds, a retest maybe performed after at least a 10-minute break (up to no more than 2 hours). Only one retest is needed, though there are a small number of instances where multiple retests were administered. The table below summarizes recommended usage of results in the analyses for all combinations. Per these rules, if a patient has both a "1" (first test) and a "0" (retest) on the same visit, only the "0" (retest) should be included in the analysis. A test that failed QC may indicate a lack of effort, major distractions, or not understanding the test instructions. Upon review by Joggle Research of the PVT tests that failed QC and also discussion with test facilitators from the study site, a flag will be created to indicate "Passed QC", "Failed QC: ambiguous compliance" and "Failed QC: clear non-compliance". For those tests that indicate "Failed QC: clear non-compliance", the tests and the associated metrics will be excluded from the analyses since the results are judged to be procedurally non-compliant rather than due to psychomotor impairment. It is believed that the test facilitators play important roles on the patient's training and procedure compliance. If the difference in proportion of "invalid" tests between the treatments is more than 20%, an exploratory analysis of the summaries in section 8.2.16 will be repeated by including these tests.

The table below summarizes recommended usage of results in the analyses for all combinations. Per these rules, if a patient has both a "1" (first test) and a "0" (retest) on the same visit, only the "0" (retest) should be included in the analysis.

Table 15.9.1 Recommended usage of results in the analyses for all combinations

First Test	Retest Test	Primary Analysis Approach	Sensitivity Analysis Approach

NO RETEST PROCESS IMPLEMENTED			
(0) No Flag	-	Include	(same as primary analysis)
(1) Ambiguous Compliance Flag	-	Include	(same as primary analysis)
(2) Clear Non-Compliance Flag	-	Exclude	Include
AFTER RETEST PROCESS IMPLEMENTED			
(0) No Flag	-	Include first	(same as primary analysis)
(1) Ambiguous Compliance Flag	(0) No Flag	Exclude first, include retest	(same as primary analysis)
	(1) Ambiguous Compliance Flag	Exclude first, include retest	(same as primary analysis)
	(2) Clear Non-Compliance Flag	Include first, exclude retest	(same as primary analysis)
(2) Clear Non-Compliance Flag	(0) No Flag	Exclude first, include retest	(same as primary analysis)
	(1) Ambiguous Compliance Flag	Exclude first, include retest	(same as primary analysis)
	(2) Clear Non-Compliance Flag	Exclude first, exclude retest	Exclude first, include retest
If a subject has multiple retests performed, the above process will be performed for all the retests in chronological order i.e. first retest will be compared to second retest, etc…			

Neurocognitive Test

The ANAM tests are administered by a computer/iPad/electronic device and patients need only to use the 2 mouse buttons to respond. For each of the tests, patients will be provided instructions that describe how to take the test and indicate a correct response. The ANAM comprises of three subsets:

- Running memory assesses sustained attention, working memory, and resistance to interference. The test requires sustained attention in response to a stimulus on the computer screen during a forced-pace, rapid task. Numbers are presented on the screen and the user must press a specified key indicating whether the number is the same as or different than the previous number. The subtest takes approximately 5 minutes to complete.
- Mathematical processing assesses mathematical computation and working memory. The test involves Math problems being presented on the screen. The answer must be figured out then the user must decide if the answer is > or < the number 4. The subtest takes approximately 5 minutes to complete
- Procedural reaction time is an assessment of choice reaction time. The user must tap the left button if the number shown on the screen is 2 or 3 and the right mouse button if the number shown on the screen is 4 or 5. The subtest takes approximately 3 minutes to complete.

ANAM test scores will be reviewed periodically by Vista Life Sciences and Washington Neuropsychology Research Group, LLC (WNRG) to identify atypical observations or those that are indicative of invalid performance. Accuracy scores ≤ 56% will be identified and flagged as "Invalid".

The tests included in the Dupistad Study battery are not overly difficult. In addition, they all involve binary response options, thus 50% accuracy would reflect chance-level responding. Tests with validity flag as "Invalid" will be excluded from the ANAM data analysis. If the difference in proportion of "invalid" tests between the treatments is more than 20%, an exploratory analysis of the summaries in section 8.2.17 will be repeated by including these tests.

15.10. Objective Assessments

Actigraphy data

Wrist actigraphy is a technique for measuring movement of a limb over an extended recording period (days to weeks). The signals generated by wrist movement are sensed by a tiny microcomputer contained within the watch and translated into activity counts.
Actigraphy is being worn on the wrist of the non-dominant hand in this study to provide estimates of the duration, timing and patterns of sleep in study participants (eg, TST, SOL, WASO, and SE). Eligible patients will wear an Actiwatch during the 7 days immediately preceding the baseline visit, daily through week 4 (visit 4), and then daily during the week before each planned visit.

Actigraphy data score will be directly provided by KONEKSA.

Polysomnography (PSG)

Polysomnography is a multi-parametric test used in the study of sleep (eg. Biophysiological changes that occur during sleep) and as a diagnostic tool in sleep medicine. Polysomnography is typically performed at night and monitors many body functions. In this study, PSG is used to record SOL, WASO, TST, longest sleep episode, stages of sleep, SE and arousal index.

For this PSG-sub study: 30 patients total (approximately 20 dupilumab and 10 placebo patients) will participate in the PSG sub-study. Each patient who participates in this sub-study will have a total of 3 overnight PSGs: at baseline timepoint, 2 'preferably) consecutive overnight PSGs to control for "first night effect" and collect baseline data; then 1 overnight PSG at week 12.

PSG data will be directly provided by STANFORD.

15.11. Rescue and Prohibited Medications.

Rescue medications will be identified by selecting "rescue therapy" for the reason for treatment in "Medications" CRF page.

Treatment with following concomitant medications is prohibited until the end of the study

- Sedative anxiolytic or hypnotic treatments other than melatonin
- Patients taking systemic sedative antihistamines more than 5 days per week
- Lipophilic beta blockers, opioids, theophylline, clonidine, antidepressants or other medications known to interfere with sleep and AD as determined by the Investigators
- Live (attenuated) vaccine
- Immunomodulating biologics
- Investigational drugs

- High-potency and super-potent TCS, except when used for rescue
- Systemic corticosteroids or non-steroidal systemic immunosuppressive drugs (eg, CsA, AZA, MTX, MMF, JAK inhibitors, etc.), except when used for rescue

Study drug will be immediately discontinued if any of the following are used through week 24:

- Treatment with a live (attenuated) vaccine
- Treatment with immunomodulating biologics
- Treatment with an investigational drug
- Treatment with systemic corticosteroids or non-steroidal systemic immunosuppressive drugs (eg, CsA, AZA, MTX, MMF, JAK inhibitors, etc.)

Note: If a patient receives rescue treatment with systemic corticosteroids or other systemic immunosuppressive drugs (CsA, AZA, MTX, MMF, JAK inhibitors, etc.), study treatment may be resumed if deemed appropriate by the Investigator and the medical monitor, but not sooner than 5 half-lives after the last dose of systemic rescue medication.

The following procedures are prohibited during study participation:

- Phototherapy
- Tanning in a bed/booth

List of coded terms for prohibited medications will be provided by SANOFI.