

- **Protocol number: RD.03.SPR.114322**
- **Document title: A randomized, double-blind, multi-center, parallel-group, placebo-controlled dose-ranging study to assess the efficacy and safety of nemolizumab (CD14152) in moderate-to-severe atopic dermatitis subjects with severe pruritus receiving topical corticosteroids**
- **Version number: 2.0**
- **Date of the document: 20 August 2018**
- **NCT number: NCT03100344**

CCI

GALDERMA R&D, SNC

Protocol Number: RD.03.SPR.114322

A randomized, double-blind, multi-center, parallel-group, placebo-controlled dose-ranging study to assess the efficacy and safety of nemolizumab (CD14152) in moderate-to-severe atopic dermatitis subjects with severe pruritus receiving topical corticosteroids

Statistical Analysis Plan

CCI

CCI

SPONSOR SIGNATURE PAGE

Approved by:



Aug 20, 2018
Date

PAREXEL SIGNATURE PAGE

Signatures below confirm that the review process has been completed in accordance with **CCI**

This document has been approved and signed electronically on the final page by the following:

Signatory	
Author	PPD Project Role: Biostatistics Lead
Signatory	
Reviewer	PPD Project Role: Biostatistics Reviewer
	PPD Project Role: Primary Statistical Programmer
	PPD Project Role: GMS Reviewer
	PPD Project Role: PK Scientist

TABLE OF CONTENTS

1 INTRODUCTION 10

2 STUDY OBJECTIVES 11

 2.1 Primary Objective 11

 2.2 Secondary Objectives 11

3 INVESTIGATIONAL PLAN 11

 3.1 Overall Study Design and Plan 11

 3.2 Efficacy and Safety Variables 12

 3.2.1 Efficacy Variables 12

 3.2.1.1 Eczema Area and Severity Index (EASI) 12

 3.2.1.2 Investigator’s Global Assessment (IGA) 12

 3.2.1.3 Body Surface Area (BSA) 12

 3.2.1.4 SCORing Atopic Dermatitis (SCORAD) 13

 3.2.1.5 Pruritus Numeric Rating Scale (NRS) 13

 3.2.1.6 Pruritus Categorical Scale (PCS) 13

 3.2.1.7 Dynamic Pruritus Score (DPS) 14

 3.2.1.8 5-D itch scale 14

 3.2.1.9 Sleep Disturbance Numeric Rating Scale (NRS-Sleep) 14

 3.2.2 Efficacy endpoints 15

 3.2.2.1 Primary endpoint 15

 3.2.2.2 Secondary efficacy endpoints 15

 3.2.2.3 Other efficacy endpoints 15

 3.2.3 Safety Variables 16

 3.2.3.1 Electrocardiograms (ECG) 16

 3.2.3.2 Physical examination and vital signs 16

 3.2.3.3 Respiratory Assessments 17

 3.2.3.4 Laboratory safety tests 17

 3.2.3.5 Adverse events 18

 3.2.4 Pharmacokinetic and anti-drug antibody assessments 19

 3.2.5 Pharmacodynamic assessments 21

 3.2.6 Quality of life assessments 21

 3.2.6.1 Dermatology Life Quality Index (DLQI) 21

 3.2.6.2 Hospital Anxiety and Depression Scale (HADS) 22

 3.2.6.3 Sick leave/missed school day 23

 3.2.6.4 EuroQoL 5-Dimension (EQ5D) 23

4 STATISTICAL METHODS 23

 4.1 Data Quality Assurance 23

 4.2 General Presentation Considerations 23

 4.3 Study Subjects 25

 4.3.1 Disposition of Subjects 25

 4.3.2 Protocol Deviations 26

 4.4 Analysis Populations 27

 4.5 Demographic and Other Baseline Characteristics 28

 4.5.1 Demographics and baseline characteristics 28

 4.5.2 Medical History 29



4.5.3	Prior and concomitant therapies.....	29
4.5.3.1	Prior and concomitant medications.....	31
4.5.3.2	Prior and concomitant medical and surgical procedures.....	32
4.5.3.3	Rescue therapies.....	32
4.6	Treatment Compliance.....	33
4.7	Efficacy Evaluation	34
4.7.1	Analysis and Data Conventions.....	34
4.7.1.1	Multi-center Studies	34
4.7.1.2	Adjustments for Covariates	35
4.7.1.3	Handling of Missing Data.....	35
4.7.1.4	Multiple Comparisons/Multiplicity.....	36
4.7.1.5	Planned Analyses	36
4.7.1.6	Examination of Subgroups	37
4.7.2	Primary Efficacy Endpoint	37
4.7.3	Secondary Efficacy Endpoints.....	38
4.7.3.1	Proportion of subjects achieving IGA success (defined as IGA 0 [clear] or 1 [almost clear]) at each visit up to week 24	39
4.7.3.2	Proportion of subjects achieving IGA success (defined as IGA 0 [clear] or 1 [almost clear]) and a reduction of at least 2 points at each visit up to week 24	39
4.7.3.3	Percent change from baseline in EASI at each visit up to week 24.....	39
4.7.3.4	Absolute and percent change from baseline in weekly average of the peak and average pruritus NRS at each visit up to week 24.....	39
4.7.3.5	Proportion of subjects with an improvement of weekly average pruritus peak NRS ≥ 4 from baseline to week 24	40
4.7.3.6	Absolute and percent change from baseline in weekly average sleep disturbance NRS at week 24.....	40
4.7.3.7	Proportion of subjects achieving PCS success (defined as a weekly prorated rounded average PCS ≤ 1 [none - mild]) at week 24	40
4.7.3.8	Proportion of subjects with EASI-50, EASI-75 or EASI-90 (defined as achieving 50%, 75% or 90% reduction from baseline in EASI score) at each visit up to week 24	40
4.7.3.9	Absolute and percent change from baseline in SCORAD at week 24	41
4.7.4	Other Efficacy Endpoints	41
4.7.4.1	Disease severity scores (EASI, IGA, SCORAD, BSA, pruritus NRS, sleep disturbance NRS and PCS) at each visit including at follow up visit (week 28 and 32)	41
4.7.4.2	Change and percentage change from baseline in disease severity scores (EASI, SCORAD, pruritus NRS, and sleep disturbance NRS) at each visit	41
4.7.4.3	Change and percentage change in disease severity scores (EASI, SCORAD, pruritus NRS, and sleep disturbance NRS) from week 24/ET to follow up visits (week 28 and 32).....	41

4.7.4.4	Proportion of subjects with an improvement of weekly average pruritus NRS ≥ 4 from baseline to each visit.....	41
4.7.4.5	Proportion of subjects achieving PCS success (defined as a weekly prorated rounded average PCS ≤ 1 [none - mild]) at each visit including follow-up visits.....	42
4.7.4.6	DPS at hour 2, 4, 8, 24, 48, and 72	42
4.7.4.7	Monthly and total amount of TCS used during the treatment period (from baseline to week 24).....	42
4.7.4.8	Time to rescue therapy	42
4.7.4.9	Change from baseline in 5-D itch scale total score at week 2.....	42
4.7.4.10	Evaluation of the prurigo nodularis lesions over time	43
4.8	Safety Evaluation.....	43
4.8.1	Extent of Exposure.....	43
4.8.2	Adverse Events	43
4.8.3	Deaths, Other Serious Adverse Events, and Other Significant Adverse Events	44
4.8.4	Clinical Laboratory Evaluation.....	45
4.8.5	Vital Signs, Weight, Physical Findings and Other Observations Related to Safety	45
4.8.5.1	Vital Signs and Weight.....	45
4.8.5.2	Peak Expiratory Flow (PEF).....	46
4.8.5.3	Electrocardiogram	46
4.8.6	Safety Monitoring (IDMC).....	46
4.9	Other Analyses	46
4.9.1	Pharmacokinetic parameters and anti-drug antibody analyses	46
4.9.2	Biomarker Analyses	49
4.9.3	Quality of Life and Productivity data Analyses.....	49
4.9.3.1	Change from baseline in DLQI at each visit up to week 24.....	49
4.9.3.2	Change from baseline in HADS at each visit up to week 24.....	50
4.9.3.3	Missed work/school day in the past 4 weeks by visit	50
4.9.3.4	Change from baseline in EQ5D at week 24.....	50
4.10	Determination of Sample Size	50
4.11	Changes in the Conduct of the Study or Planned Analysis	51
5	REFERENCES	51
Appendix 1	Schedule of assessments.....	52
Appendix 2	Criteria for potentially clinically significant vital signs and weight.....	55
Appendix 3	Dermatology Life Quality Index (DLQI)	56
Appendix 4	EQ5D Index Score	59
Appendix 5	Protocol deviations.....	61
Appendix 6	Intrinsic or extrinsic AD	62
Appendix 7	Validation of the NRS responder analysis.....	63
Appendix 8	Potentially clinically significant values (PCSV) for clinical laboratory tests.....	64

LIST OF ABBREVIATIONS

<i>Abbreviation</i>	<i>Term</i>
AD	Atopic Dermatitis
ADA	Anti-Drug Antibody
AE	Adverse Event
AESI	Adverse Event of Special Interest
AIC	Akaike's Information Criterion
ALT	Alanine Aminotransferase
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
AST	Aspartate Aminotransferase
ATC	Anatomical Therapeutic Chemical
AUC	Area under the Curve
BLQ	Below the Limit of Quantification
BMI	Body Mass Index
bpm	Beats per Minute
BSA	Body Surface Area
CI	Confidence Interval
Cmax	Maximum Concentration
CMH	Cochran-Mantel-Haenszel
CPK	Creatinine Phosphokinase
CRO	Contract Research Organization
CSR	Clinical Study Report
DLQI	Dermatology Life Quality Index
DPS	Dynamic Pruritus Scale
EASI	Eczema Area and Severity Index
ECG	Electrocardiogram
eCRF	electronic Case Report Forms
EMA	European Medicines Agency
EQ5D	EuroQoL 5-Dimension
FDA	Food and Drug Administration
HADS	Hospital Anxiety and Depression Scale

<i>Abbreviation</i>	<i>Term</i>
HSV	Health State Value
ICF	Informed Consent Form
ICH	International Conference on Harmonisation
IDMC	Independent Data Monitoring Committee
IGA	Investigator's Global Assessment
IL	Interleukin
IRR	Injection-related Reaction
ITT	Intent-to-treat; intention-to-treat
IVRS	Interactive Voice Response System
IWRS	Interactive Web Response System
KM	Kaplan-Meier
LOCF	Last Observation Carried Forward
LSmean	Least Square Mean
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple Imputation
mL	Milliliter
MMRM	Mixed-effect Model for Repeated Measures
NCA	Non compartmental analysis
NRS	Numeric Rating Scale
OC	Observe Case
PCS	Pruritus Categorical Scale
PCSV	Potentially Clinically Significant Value
PD	Pharmacodynamics
PEF	Peak Expiratory Flow
PK	Pharmacokinetics
PP	Per-Protocol
PT	Preferred Term
Q1	Lower quartile
Q3	Upper quartile
Q4W	Every 4 weeks
QoL	Quality of Life

<i>Abbreviation</i>	<i>Term</i>
REML	Restricted Maximum Likelihood
ROC	Receiver Operating Characteristic
SAE	Serious Adverse Event
SAF	Safety (analysis population)
SAP	Statistical Analysis Plan
SCORAD	SCORing Atopic Dermatitis
SD	Standard Deviation
SOC	System Organ Class
SUSAR	Suspected Unexpected Serious Adverse Reaction
TARC	thymus and activation-regulated chemokine
TB	Tuberculosis
TCI	Topical Calcineurin Inhibitor
TCS	Topical Corticosteroid
TEAE	Treatment-emergent Adverse Event
TESAE	Treatment-emergent Serious Adverse Event
TFL	Tables, Figures, and Listings
ULN	Upper Limit of Normal
UN	Unstructured
WBC	White Blood Cell
VAS	Visual Analogue Scale
WHODD	the World Health Organization Drug Dictionary

1 INTRODUCTION

This Statistical Analysis Plan (SAP) describes the planned analysis and reporting for GALDERMA R&D, SNC, protocol RD.03.SPR.114322, entitled “A randomized, double-blind, multi-center, parallel-group, placebo-controlled dose-ranging study to assess the efficacy and safety of nemolizumab (CD14152) in moderate-to-severe atopic dermatitis subjects with severe pruritus receiving topical corticosteroids”, Version 01 dated October 23, 2017.

Atopic dermatitis (AD) is a chronic inflammatory skin disease characterized by pruritus (itching), xerosis (skin dryness) and eczematous lesions. AD is currently managed with topical treatments and, when they are not sufficient to control the disease in moderate to severe AD, by systemic treatments, as well as phototherapy. Given the high response variability and the known secondary adverse effects of these drugs, there is a need for new drugs to better control the disease of moderate to severe patients while decreasing the risk of secondary adverse effects. Nemolizumab, a humanized anti-human interleukin (IL)-31 receptor A monoclonal antibody, could decrease pruritus and improve AD in these patients and therefore represent a new treatment option. Subjects with insufficient response to topical therapies and severe pruritus, which leads to extensive scratching further aggravating the disease, could particularly benefit from such a therapy.

This study is conducted to evaluate the efficacy, safety and pharmacokinetics (PK) of multiple subcutaneous doses of nemolizumab in the treatment of AD, when administered on top of background topical corticosteroid (TCS). Three doses have been selected to fully explore the dose range and enable selection of dose(s) for Phase 3 development.

The planned analyses identified in this SAP may be included in the clinical study report (CSR), regulatory submissions, or future manuscripts. Also, post hoc exploratory analyses not necessarily identified in this SAP may be performed to further examine study data. Any post hoc, or unplanned, exploratory analyses performed, if included, will be clearly identified as such in the final CSR.

This SAP is based upon the following study documents:

- Study protocol, Version 01 (23 October 2017)
- Electronic Case Report Form (eCRF), Version 5.0 (19 February 2018)

The structure and content of this SAP provides sufficient detail to meet the requirements identified by the Food and Drug Administration (FDA), European Medicines Agency (EMA), and International Conference on Harmonisation (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use: Guidance on Statistical Principles in Clinical Trials. All work planned and reported for this SAP will follow internationally accepted guidelines, e.g., Ethical Guidelines for Statistical Practice published by the American Statistical Association and Code of Conduct published by the Royal Statistical Society, for statistical practice.

The reader of this SAP is encouraged to also read the clinical protocol, and other relevant documents for details on the planned conduct of this study. Other than the schedule of assessments which is provided in Appendix 1, operational aspects related to collection and timing of planned clinical assessments are not repeated in this SAP unless relevant to the planned analyses.

2 STUDY OBJECTIVES

2.1 Primary Objective

The primary objective is to assess the efficacy of several subcutaneous doses of nemolizumab compared to placebo in moderate-to-severe AD subjects with severe pruritus receiving TCS, who were not adequately controlled with topical treatments.

2.2 Secondary Objectives

The secondary objectives are to evaluate the safety of nemolizumab and to characterize its PK profile.

3 INVESTIGATIONAL PLAN

3.1 Overall Study Design and Plan

This is a randomized, placebo-controlled, double-blind, parallel-group, dose-ranging, phase 2b study to evaluate the efficacy and safety of various doses of nemolizumab in moderate-to-severe AD subjects with severe pruritus. The duration of this clinical trial is up to 36 weeks for a given subject, including a 2 to 4-week run-in period, a 24-week treatment period and an 8 week follow-up period (12 weeks after the last study drug administration).

Eligible subjects must have moderate-to-severe AD, severe pruritus and a documented history of inadequate response to topical AD medications. Subjects meeting the inclusion/exclusion criteria at the screening visit will receive background therapy of AD (including a moisturizer, a medium potency TCS for the body, and a low potency TCS for the face, neck etc.) to be used throughout the study. Subjects who still meet the inclusion/exclusion criteria at the baseline visit will be randomized in a 1:1:1:1 ratio to either of the 3 groups of nemolizumab or placebo. Randomization will be stratified by AD severity based on the baseline Investigator's Global Assessment (IGA) scores (3 or 4). Injection of study drug will occur every 4 weeks (Q4w, at weeks 4, 8, 12, 16 & 20). A loading dose will be administered on day 1 for the groups of 10 mg and 30 mg only (20 mg and 60 mg, respectively), while for 90 mg group, the same dose will be administered at each injection visit.

An independent Data Monitoring Committee (IDMC) will be set-up to monitor safety data generated in this clinical trial on an ongoing basis and make appropriate recommendations to the sponsor.

Assessments of efficacy, safety and PK will be conducted throughout this clinical trial. Refer to Appendix 1 for the complete schedule of assessments.

3.2 Efficacy and Safety Variables

3.2.1 Efficacy Variables

3.2.1.1 Eczema Area and Severity Index (EASI)

EASI is a composite score ranging from 0 to 72. The severity of erythema, induration/papulation, excoriation, and lichenification will be assessed by the investigator or trained designee on a scale of 0 (absent) to 3 (severe) for each of the 4 body areas: head/neck, trunk, upper limbs, and lower limbs, with half points allowed. In addition, the extent of AD involvement in each of the 4 body areas 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. The EASI score will be calculated in the eCRF as follows.

Body region	EASI score
Head/Neck (H)	$(E + I + Ex + L) \times Area \times 0.1$
Upper limbs (UL)	$(E + I + Ex + L) \times Area \times 0.2$
Trunk (T)	$(E + I + Ex + L) \times Area \times 0.3$
Lower limbs (LL)	$(E + I + Ex + L) \times Area \times 0.4$
EASI =	Sum of the above 4 body region scores

The degree of severity of each sign (E=erythema, I=induration/papulation, Ex=excoriation, L=lichenification) in each of the 4 body regions is evaluated based on a scale ranging from 0 to 3 (0: none; 1: mild; 2: moderate; 3: severe), with half points allowed.

Area (the affected body area) is defined as follows: 0= 0%; 1=1-9%; 2=10-29%; 3=30-49%; 4=50-69%; 5=70-89%; 6=90-100%. Among the 4 zones, trunk includes the genital area, and lower limbs include the buttocks.

The EASI will be assessed at screening, baseline, weeks 1, 2, 4, 8, 12, 16, 20, 24, 28, 32, or early termination and/or unscheduled visits when applicable.

3.2.1.2 Investigator's Global Assessment (IGA)

IGA is a 5-point scale ranging from 0 (clear) to 4 (severe) used by the investigator or trained designee to evaluate the global severity of AD.

The IGA will be assessed at screening, baseline, weeks 1, 2, 4, 8, 12, 16, 20, 24, 28, 32, or early termination and/or unscheduled visits when applicable.

3.2.1.3 Body Surface Area (BSA)

The BSA involvement of AD will be assessed by the investigator or trained designee for each part of the body (the possible highest score for each region is: head and neck [9%], anterior trunk [18%], back [18%], upper limbs [18%], lower limbs [36%], and genitals [1%]), and will be reported as a percentage of all major body sections combined.



The BSA will be assessed at screening, baseline, weeks 1, 2, 4, 8, 12, 16, 20, 24, 28, 32, or early termination and/or unscheduled visits when applicable.

3.2.1.4 SCORing Atopic Dermatitis (SCORAD)

SCORAD ranges from 0 to 103 and has 3 components: extent (BSA, as described in Section 3.2.1.3), signs and symptoms of AD, and subjective symptoms (European task force on atopic dermatitis 1993). Investigator or designee will assess the severity of 6 signs of AD (erythema/darkening, edema/papulation, oozing/crusting, excoriation, lichenification/prurigo and dryness), each on a scale ranging from 0 (none) to 3 (severe). Investigator or designee will also ask the subjects to evaluate their symptoms of pruritus and sleep loss (average for the last 3 days/nights), each evaluated on a Visual Analogue Scale (VAS) from 0 to 10. The SCORAD score will be calculated in the eCRF as follows.

SCORAD score $\frac{A + B + C}{3}$

where: A = Extent of AD, B = Intensity of AD, C = Subjective symptoms (itch and sleeplessness).

The maximum SCORAD score is 103. The SCORAD will be assessed at baseline, weeks 1, 2, 4, 8, 12, 16, 20, 24, 28, 32, or early termination and/or unscheduled visits when applicable.

3.2.1.5 Pruritus Numeric Rating Scale (NRS)

Pruritus NRS is a scale to be used by the subjects to report the intensity of their pruritus (itch) during the last 24 hours on a scale of 0 to 10. Subjects will record their pruritus NRS scores (average itch intensity and maximum itch intensity during the previous 24 hours) on an electronic device, and will complete the assessment once daily in the evening throughout the clinical trial (including the run-in and the follow-up period).

Weekly prorated average score will be calculated and used in the analysis. If a subject has less than 4 diary entries in a week, the weekly average scores will be set to missing. Refer to Section 4.2 for classification of diary data into analysis week/visit.

3.2.1.6 Pruritus Categorical Scale (PCS)

The 4-point PCS is used by the subjects to report the intensity of their pruritus. The scale ranges from 0 (absence of pruritus) to 3 (severe pruritus). Subjects will record their PCS scores on an electronic device, and will complete the assessment once daily in the evening throughout the clinical trial (including the run-in and the follow-up period).

Weekly prorated average score will be calculated. When used in analysis on proportion of subjects achieving PCS success (e.g., PCS ≤ 1), the calculated scores will be rounded to 1 before the determination of responder status. If a subject has less than 4 diary entries in a week, the

weekly average scores will be set to missing. Refer to Section 4.2 for classification of diary data into analysis week/visit.

3.2.1.7 Dynamic Pruritus Score (DPS)

The 9-point DPS is a dynamic scale to be used by subjects to evaluate the change of their pruritus compared with an earlier time point (i.e., before injection on day 1). The scale ranges from 0 (strongly worsened pruritus) to 8 ([almost] no pruritus anymore), including intermediate marks for slightly improved/worsened, moderately improved/worsened, and rather improved/worsened. Subjects will record their DPS score on an electronic device, and will complete the assessment 2, 4, 8, 24, 48 and 72 hours after study drug injection on day 1.

3.2.1.8 5-D itch scale

The 5-D itch scale is a multidimensional measure of itching that has been validated in subjects with chronic pruritus. The 5 dimensions included in the scale are duration, degree, direction, disability and distribution. The scores for each of the 5 domains are achieved separately and then summed together to obtain a total 5-D score, which ranges between 5 (no pruritus) to 25 (most severe pruritus). At a given visit, there must be responses to at least 3 out of 5 domains. If fewer than 3 responses, then the total 5-D score will be set to missing.

The duration, degree and direction domains are single-item domain.

The score for the disability domain is achieved by taking the highest score on any of the 4 items. The disability score is missing when scores for all 4 items are missing.

For the distribution domain, the number of affected body parts is tallied (potential sum 0-16) and the sum is sorted into 5 scoring bins as follows:

- sum of 0-2 = score of 1
- sum of 3-5 = score of 2
- sum of 6-10 = score of 3
- sum of 11-13 = score of 4
- sum of 14-16 = score of 5

The 5-D itch scale will be administered at baseline and week 2.

3.2.1.9 Sleep Disturbance Numeric Rating Scale (NRS-Sleep)

The sleep disturbance NRS is a scale to be used by the subjects to report the degree of their sleep loss related to AD. Subjects will record their sleep disturbance NRS scores on an electronic device, and will complete the assessment once daily in the morning throughout the clinical trial (including the run-in and the follow-up period).

Weekly prorated average score will be calculated and used in the analysis. If a subject has less than 4 diary entries in a week, the weekly average scores will be set to missing. Refer to Section 4.2 for classification of diary data into analysis week/visit.

3.2.2 Efficacy endpoints

3.2.2.1 Primary endpoint

Percent change from baseline in EASI at week 24

3.2.2.2 Secondary efficacy endpoints

- Proportion of subjects achieving IGA success (defined as IGA 0 [clear] or 1 [almost clear]) at each visit up to week 24
- Proportion of subjects achieving IGA success (defined as IGA 0 [clear] or 1 [almost clear]) and a reduction of ≥ 2 points at each visit up to week 24
- Percent change from baseline in EASI at each visit up to week 24
- Absolute and percent change from baseline in weekly average of the peak and average pruritus NRS at each visit up to week 24
- Proportion of subjects with an improvement of weekly average pruritus peak NRS ≥ 4 from baseline to week 24
- Absolute and percent change from baseline in weekly average sleep disturbance NRS at week 24
- Proportion of subjects achieving PCS success (defined as a weekly prorated rounded average PCS ≤ 1 [none - mild]) at week 24
- Proportion of subjects with EASI-50, EASI-75 or EASI-90 (defined as achieving 50%, 75% or 90% reduction from baseline in EASI score) at each visit up to week 24
- Absolute and percent change from baseline in SCORAD at week 24

The endpoints of IGA, EASI, and peak pruritus NRS are considered as selected secondary efficacy endpoints hereafter.

3.2.2.3 Other efficacy endpoints

- Disease severity scores (EASI, IGA, SCORAD, BSA, pruritus NRS, sleep disturbance NRS and PCS) at each visit including follow up visits (week 28 and 32)
- Change and percentage change from baseline in disease severity scores (EASI, SCORAD, pruritus NRS and sleep disturbance NRS) at each visit
- Change and percentage change in disease severity scores (EASI, SCORAD, pruritus NRS, and sleep disturbance NRS) from week 24/ET to follow up visits (week 28 and 32)
- Proportion of subjects with an improvement of weekly average pruritus NRS ≥ 4 from baseline to each visit
- Proportion of subjects achieving PCS success (defined as a weekly prorated rounded average PCS ≤ 1 [none - mild]) at each visit including follow-up visits
- DPS at hour 2, 4, 8, 24, 48, and 72

- Monthly and total amount of TCS (medium and low potency TCS only) used during the treatment period (baseline to week 24)
- Time to rescue therapy
- Change from baseline in 5-D itch scale total score at week 2
- Prurigo nodularis lesions over time, if applicable

3.2.3 Safety Variables

Safety assessments will be conducted for all subjects at the screening visit (upon the signature of the informed consent form [ICF]) and at every subsequent visit.

Safety assessments include electrocardiogram (ECG), physical examination, vital signs, body weight, respiratory assessments (peak expiratory flow [PEF] measurement only for subjects with a medical history of asthma), laboratory safety tests, and AE recording (including serious AEs [SAE], AEs of Special Interest [AESI] and selected AEs).

3.2.3.1 Electrocardiograms (ECG)

A 12-lead ECG will be performed at screening, baseline, weeks 1, 12, 24, or early termination and/or unscheduled visits when applicable.

All abnormal ECG findings considered to be clinically significant by the investigator at the screening visit will be recorded in the Medical History form. Any clinically significant changes from the screening visit will be reported as AEs in the eCRF.

3.2.3.2 Physical examination and vital signs

Physical examination: The physical examination will be performed at screening, baseline, weeks 1, 2, 4, 8, 12, 16, 20, 24, 28, 32, or early termination and/or unscheduled visits when applicable.

All clinically significant abnormal findings at the screening visit will be recorded in the Medical History form. Any clinically significant changes from the screening visit will be recorded as an AE.

At baseline, investigator or qualified personnel should assess on each subject whether lesions of prurigo nodularis are present or absent. If the lesions of prurigo nodularis are present, investigator needs to evaluate the lesions at each subsequent visit using a 5-point scale (2: very much improved; 1: improved; 0: no change; -1: worsened; -2: very much worsened).

Vital signs: The vital signs will be evaluated at screening, baseline, weeks 1, 2, 4, 8, 12, 16, 20, 24, 28, 32, or early termination and/or unscheduled visits when applicable.

Vital signs will include pulse rate, systolic and diastolic blood pressure (after the subject has been sitting for at least 5 minutes), and body temperature. All abnormal values at the screening

visit identified as clinically significant by the investigator will be recorded in the Medical History form. Any clinically significant changes from the screening visit will be recorded as an AE.

Height and Weight: Height will be measured at the screening visit only, and weight will be measured at screening, baseline, weeks 4, 8, 12, 16, 20 and 24.

Any clinically significant weight changes from the screening visit will be recorded as an AE.

Blood pressure (systolic and diastolic), pulse rate, and weight values meeting criteria for potentially clinically significant values (PCSV) will be reported.

Criteria for PCSV for vital signs and weight are listed in Appendix 2.

3.2.3.3 Respiratory Assessments

At each visit, investigator or designee will perform a respiratory physical examination and ask all subjects whether they have experienced any signs/symptoms of asthma. In addition, for subjects reporting a medical history of asthma, PEF will be performed.

Newly diagnosed asthma or worsening of asthma during the study will be reported as AESI.

3.2.3.4 Laboratory safety tests

The following laboratory safety tests will be performed:

- Hematology: white blood cell (WBC) count with differential count (including eosinophils), red blood cell (RBC) count, hemoglobin (Hb), hematocrit (hct), mean cell volume (MCV), and platelet count (Plt)
- Blood chemistry: sodium, potassium, calcium, chloride, glucose, urea, creatinine, aspartate aminotransferase (AST)/ alanine aminotransferase (ALT), alkaline phosphatase (ALP), total and direct bilirubin, creatinine phosphokinase (CPK), high sensitivity C-reactive protein (hsCRP), fibrinogen, gamma-glutamyl transpeptidase (GGT), lactate dehydrogenase (LDH), total protein, albumin, uric acid, total cholesterol, low-density lipoprotein (LDL), high-density lipoprotein (HDL) and triglycerides
CPK isoenzyme test will be performed only if CPK is elevated to >2.5X Upper Limit of Normal (ULN).
- Urinalysis: blood, proteins, leukocytes, glucose, ketones, nitrites, bilirubin, urobilinogen, pH, and specific gravity

All clinically significant out-of-range laboratory values at the screening visit will be recorded in the medical history form. All clinically significant out-of-range laboratory values after the screening visit are to be reported as an AE if this abnormality was not present at the screening visit or is assessed as having worsened since the screening visit (i.e., changed significantly from the screening visit).

3.2.3.5 Adverse events

Adverse events (AE) will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 19.0.

An AE will be considered a treatment-emergent AE (TEAE) if it started after the date of the first dose of study drug or it started before the date of the first dose of study drug and worsened after the date of the first dose. AEs occurring on the date of the first dose of study drug will be considered a TEAE if the onset of the AE is on or after the time of injection on that day. Where dates are missing or partially missing, AEs will be assumed to be treatment-emergent, unless there is clear evidence (through comparison of partial dates) to suggest that the AE started prior to the first dose of study treatment.

The investigator will have to assess if there is a reasonable possibility of a causal relationship between the study drug and/or a study procedure (e.g. injection, TCS, blood sample collection) and an AE. Related AEs will be defined as study drug related AEs or study procedure related AEs.

The AESIs for this clinical trial have been defined as follows:

- Elevated ALT or AST (> 3 ULN) in combination with elevated bilirubin > 2 ULN, whether or not considered as related to the study drug by the investigator
- Newly diagnosed asthma or worsening of asthma
- Elevated CPK (≥ 2.5 ULN) if considered to be related to study drug by the investigator

The following AEs will be considered as selected AEs for this clinical trial:

- Exacerbation of AD: defined as a clinically significant worsening of AD signs and/or symptoms that requires therapeutic intervention (rescue therapy) and that is not considered by the investigator to be a part of the natural course of AD (e.g. change in severity or the nature of the disease). When recording worsening of AD on the AE form in the eCRF, the AE reported term should include an appropriate descriptor (e.g. “worsening of AD” or “AD flare”).
- Injection-related reaction (IRR): defined as any local or systemic reactions (including hypersensitivity) related to the injection, regardless of the time of the AE onset.
- Peripheral edema: defined as swelling of any skin or subcutaneous body tissues (for example legs, hands, etc.) but excluding edema linked to urticaria.
- Skin infection
- Systemic infection
- Headache

Missing date information for AEs

The following imputation rules only apply to cases in which the start date for AEs is incomplete (i.e., partial missing).

Missing Month and Day

- If the year of the incomplete start date is the same as the year of the first administration of the study drug, then the month and day of the first administration of the study drug will be assigned to the missing fields.
- If the year of the incomplete start date is before the year of the first administration of the study drug, *December 31* will be assigned to the missing fields.
- If the year of the incomplete start date is after the year of the first dose of double-blind IP, *January 1* will be assigned to the missing fields.

Missing Month Only

- If only the month is missing, then the day will be treated as missing and both the month and the day will be replaced according to the procedure above.

Missing Day Only

- If the month and year of the incomplete start date are the same as the month and year of the first administration of the study drug, then the day of the first administration of the study drug will be assigned to the missing day.
- If either the year of the incomplete start date is before the year of the date of the first administration of the study drug or if both years are the same but the month of the incomplete start date is before the month of the date of the first administration of the study drug, then the last day of the month will be assigned to the missing day.
- If either the year of the incomplete start date is after the year of the date of the first administration of the study drug or if both years are the same but the month of the incomplete start date is after the month of the date of the first administration of the study drug, then the first day of the month will be assigned to the missing day.

If the stop date is complete and the imputed start date as above is after the stop date, then the start date will be imputed by the stop date.

If the start date is completely missing and the stop date is complete, then the following algorithm will be used to impute the start date:

- If the stop date is after the date of the first administration of the study drug, then the date of the first administration of the study drug will be assigned to the missing start date.
- If the stop date is before the date of the first administration of the study drug, then the stop date will be assigned to the missing start date.

3.2.4 Pharmacokinetic and anti-drug antibody assessments

Blood samples will be collected to determine the PK profile of nemolizumab and to assess anti-drug antibodies (ADA). The serum concentration of nemolizumab will be assessed at baseline, weeks 1, 2, 4, 8, 12, 16, 20, 24, 28, 32 and at any unscheduled visit for safety reasons. ADA will be assessed at baseline, weeks 4, 8, 16, 24, 32 and at any unscheduled visit for safety reasons. ADA assessment will include screening, confirmatory, titer, neutralizing antibodies (Nab) and IgE-based assays.

At the injection visits (baseline, weeks 4, 8, 12, 16 and 20), PK samples will be collected within 30 minutes before drug injection (pre-dose samples); At weeks 1, 2, 24, 28 and 32, blood samples will be collected at the usual time of drug injection with an allowed time window of ± 1 h.

Results will be described in the bioanalytical report, which will be included as an appendix in the final clinical study report.

PK parameters will be determined by a model independent approach (non-compartmental analysis [NCA]) using individual serum concentration values. The details surrounding NCA analysis will be covered under a separate PK Analysis Plan. When appropriate, the following PK parameters will be determined for each subject:

From baseline to week 4:

- C_{\max} : The observed peak drug concentration
- T_{\max} : The time at which C_{\max} occurs
- AUC_{0-28d} : Area under the concentration time curve from pre dose through approximately 28 days post dosing. AUC_{0-28d} will be calculated by mixed linear logarithmic trapezoidal method. AUC_{0-28d} includes the serum drug concentration up to week 4 (i.e. the serum concentration collected at baseline, week 1, 2 and 4).

From baseline to week 12:

- AUC_{0-12w} : Area under the concentration time curve from pre dose through 12 weeks post dosing. AUC_{0-12w} will be calculated by mixed linear logarithmic trapezoidal method. AUC_{0-12w} includes the serum drug concentration up to week 12 (i.e. the serum concentration collected at baseline, week 1, 2, 4, 8 and 12).

From baseline to week 24:

- AUC_{0-24w} : Area under the concentration time curve from pre dose through 24 weeks post dosing. AUC_{0-24w} will be calculated by mixed linear logarithmic trapezoidal method. AUC_{0-24w} includes the serum drug concentration up to week 24 (i.e. the serum concentration collected at baseline, week 1, 2, 4, 8, 12, 16, 20 and 24).

From week 20 to week 32:

- AUC_{0-t} : Area under the concentration time curve calculated by the mixed linear logarithmic trapezoidal method from the last administration (week 20) up to the sampling time corresponding to the last quantifiable concentration (C_{last}).

- AUC_{0-inf} : Area under the plasma concentration time curve calculated by the mixed linear logarithmic trapezoidal method from the last administration (week 20) and extrapolated to time infinity
- $t_{1/2}$: the terminal half-life value ($t_{1/2}$) will be calculated using the equation $\ln 2/k$ after the last drug injection (week 20).

At weeks 4, 8, 12, 16, 20, 24, 28, 32:

- C_{trough} : The residual drug concentration (pre-dose level).
- Accumulation index: calculated with the serum nemolizumab trough concentration obtained 4 weeks after the first dose (week 4) and 4 weeks after last drug injection (week 24).

ADA will be evaluated using a validated ELISA screening assay. If serum circulating ADA is detected, they will be characterized using a validated assay. ADA assessment will include screening, confirmatory, titer, neutralizing antibodies (Nab) and IgE-based assays.

3.2.5 Pharmacodynamic assessments

Blood and stratum corneum samples will be collected according to schedule of assessments to investigate the effect of nemolizumab on selected biomarkers, including but not limited to thymus and activation-regulated chemokine (TARC), IgE, IL6, IL8, and IL18.

3.2.6 Quality of life assessments

3.2.6.1 Dermatology Life Quality Index (DLQI)

DLQI is a validated 10-item questionnaire, covering domains including symptoms/feelings, daily activities, leisure, work/school, personal relationships and treatment. Subject will rate each question ranging from 0 (not at all) to 3 (very much), and the total score ranges from 0 to 30, with a higher score indicating a poorer quality of life (QoL). Please note that “Not relevant” is scored as 0, and Question 7, “prevented work or studying” is scored as 3. DLQI will be administered only to the subset of subjects who fluently speak a language in which the questionnaire is presented.

DLQI will be assessed at baseline, weeks 2, 12, 24, or early termination visit when applicable.

The DLQI score will be categorized as follows:

- 0-1 = no effect at all on subject’s life
- 2-5 = small effect on subject’s life
- 6-10 = moderate effect on subject’s life
- 11-20 = very large effect on subject’s life
- 21-30 = extremely large effect on subject’s life

The DLQI can be analyzed under 6 sub-scales, namely, symptoms and feelings (question 1 and 2), daily activities (question 3 and 4), leisure (question 5 and 6), work and school (question 7), personal relationships (question 8 and 9), and treatment (question 10).

For general inflammatory skin conditions a change in DLQI score of at least 4 points is considered clinically important.

Refer to Appendix 3 for details on scoring, data handling rules, and analysis on DLQI.

3.2.6.2 Hospital Anxiety and Depression Scale (HADS)

HADS is a validated questionnaire containing 14 items, 7 each for anxiety and depression. Subject will rate each question ranging from 0 to 3, and the total score ranges from 0 to 21 for each sub-scale. HADS will be administered only to the subset of subjects who fluently speak a language in which the questionnaire is presented.

For each sub-scale, the score will be categorized as follows:

- 0-7 = Normal
- 8-10 = Borderline abnormal (borderline case)
- 11-21 = Abnormal (case):

HADS will be assessed at baseline, weeks 12, 24, or early termination visit when applicable.

At a given visit, there must be responses to at least 6 out of 7 questions within each sub-scale of anxiety and depression. If fewer than 6 responses, then the sub-scale will be set to missing.

3.2.6.3 Sick leave/missed school day

Subjects who are employed or enrolled in school/university at baseline will be asked to report the number of missed work/school days due to AD since the last visit (or since the last 4 weeks at week 4) not including missed work/school days due to participation in this clinical trial.

3.2.6.4 EuroQoL 5-Dimension (EQ5D)

EQ5D is a validated questionnaire for the assessment of the general health state. It contains 2 parts: a descriptive system and a VAS. The descriptive system is made up of 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. The VAS consists of a vertical line where the subject can assess his or her own health status. EQ5D will be administered according to the schedule of assessments only to the subset of subjects who fluently speak a language in which the questionnaire is presented.

Refer to Appendix 4 for details on calculation of EQ5D index score.

4 STATISTICAL METHODS

4.1 Data Quality Assurance

All tables, figures and data listings (TLFs) to be included in the report will be independently checked for consistency, integrity and in accordance with standard PAREXEL procedures.

4.2 General Presentation Considerations

Unless otherwise stated, descriptive summaries will be presented by treatment group. In all tables the treatment groups will be presented in the order of Placebo, Nemolizumab 10 mg, Nemolizumab 30 mg, and Nemolizumab 90 mg. Comparisons on the difference between each nemolizumab group and placebo will be presented.

Continuous data will be summarized in terms of the mean, standard deviation (SD), median, minimum, maximum, lower quartile (Q1), upper quartile (Q3), and number of observations, unless otherwise stated. Where data are collected over time, both the observed data and change from baseline will be summarized at each time point. The minimum and maximum will be reported to the same number of decimal places as the raw data recorded in the database. The mean, median, Q1, and Q3 will be reported to one more decimal place than the raw data recorded in the database. The SD will be reported to 2 more decimal places than the raw data recorded in the database. In general, the maximum number of decimal places reported shall be four for any summary statistic.

Categorical data will be summarized in terms of the number of subjects providing data at the relevant time point (n), frequency counts and percentages. Percentages will be presented to one decimal place. A percentage of 100% will be reported as 100%. Percentages will not be presented for zero counts. Percentages will be calculated using n as the denominator. If sample sizes are small, the data displays will show the percentages, but any textual report will describe

frequencies only. Changes from baseline in categorical data will be summarized using shift tables where appropriate.

P-values greater than or equal to 0.001, in general, will be presented to three decimal places. However, if a p-value is only presented to four decimal places (by SAS) it will not be rounded again, but will be presented to four decimal places. P-values less than 0.001 will be presented as “<0.001”, and p-values greater than 0.999 will be presented as “>0.999”.

Confidence intervals (CIs) will be presented to one more decimal place than the raw data.

All report outputs will be produced using SAS® version 9.3 or a later version in a secure and validated environment. The outputs will be provided as individual TLFs in RTF format and all TLFs in one PDF bundle.

All statistical tests will be two-sided and will be performed at the 5% level of significance, unless otherwise stated.

Baseline

‘Baseline’ is defined as the last available assessment prior to the first administration of study drug, if it exists; otherwise, ‘baseline’ is defined as the last available assessment prior to date of randomization.

For diary data (pruritus NRS, PCS, Sleep disturbance NRS), the baseline values will be derived from data collected during the 7 days prior to the first administration of study drug. Baseline score will be the weekly prorated average of non-missing subject diary scores reported during the 7 days. A minimum of 4 daily scores out of the 7 days is required to calculate the weekly prorated average score.

Analysis Visit

Study day is relative to date of randomization. Day -1 is the day before randomization, and day 1 is date of randomization.

Assessments at early termination and unscheduled visits will be slotted based on the following visit window.

Period	Analysis Visit	Target day (relative to randomization)	Day ranges
Run-in	Screen	<1	<1
Treatment	Baseline	1	≤1
	Week 1	8	2-11
	Week 2	15	12-22
	Week 4	29	23-43
	Week 8	57	44-71
	Week 12	85	72-99
	Week 16	113	100-127

	Week 20	141	128-155
	Week 24	169	156-183
Follow-up	Week 28	197	184-211
	Week 32	225	>211

The diary reported during study will be classified into analysis visits as follows (i.e., during the 7 days prior to the target day of analysis visit. In this case, the target day of analysis visit is relative to date of the first administration of study drug.):

- Day -7 to -1 = Baseline
- Day 1 to 7 = Week 1
- Day 8 to 14 = Week 2
- Day 22 to 28 = Week 4
- Day 50 to 56 = Week 8
- Day 78 to 84 = Week 12
- Day 106 to 112 = Week 16
- Day 134 to 140 = Week 20
- Day 162 to 168 = Week 24
- Day 190 to 196 = Week 28
- Day 218 to 224 = Week 32

These analysis visits will be used in the calculations for all week-based parameters collected on subject's diary (e.g., pruritus NRS, PCS, Sleep disturbance NRS). Other non-diary data, such as EASI, IGA, BSA, SCORAD, DPS, DLQI, HADS, EQ5D, 5-D itch, missed work/school day, vital signs, body weight, clinical laboratory assessments, respiratory assessment/PEF, will be analyzed according to actual scheduled visits.

For efficacy and safety data (except clinical laboratory assessments), where two or more assessments (include both scheduled and unscheduled assessments) are available for the same visit interval, the one closest to the target visit date will be used for the summary and analyses. For clinical laboratory assessments, if repeated measurements are taken for either time point (scheduled visit), then the last measurement will be used for the value for that time point. Unscheduled labs will be included in listings, but not summaries of the data. All post baseline assessments, including repeated or unscheduled visits, will be used for potentially clinically significant (PCS) value determinations.

4.3 Study Subjects

4.3.1 Disposition of Subjects

A clear accounting of the disposition of all subjects who enter the study will be provided, from screening to study completion.

The following subject disposition summaries will be produced:



- The number of subjects screened (Analysis population: All Subjects)
- Subjects Randomized by clinical site and overall (Analysis population: Screened Subjects)
- Screen failures (Analysis population: Screened Subjects)

The following summary will be produced for ITT population.

- The number of subjects randomized
- The number and percentage of subjects treated (with at least one dose of study drug)
- The number and percentage of subjects randomized but not treated (not taken any dose of study drug)
- The number and percentage of subjects who temporarily discontinued the study treatment and reasons for discontinuation
- The number and percentage of subjects who permanently discontinued the study treatment and reason of discontinuation
- The number of subjects completing study treatment (defined as completing week 24 visit)
- The number of subjects completing the study
- The number of subjects who discontinued from the study and reason of discontinuation
- A Kaplan-Meier plot of the time to permanent study treatment discontinuation will be provided. Subjects who did not discontinue from the study drug will be censored at their last dose of study drug (Analysis population: ITT).
- A Kaplan-Meier plot of the time to study discontinuation will be provided. Subjects who did not discontinue from the study will be censored at the date of study completion or the last visit date on the study (Analysis population: ITT).
- Time (days) to permanent discontinuation of study drug by reasons for discontinuation will be displayed graphically in subjects having permanently discontinued study drug.
- Time (days) to study discontinuation by reasons for discontinuation from study will be displayed graphically in subjects having discontinued study.

By-subject listings of eligibility details, randomization details, visit dates and withdrawal/study completion details (including reason for discontinuation and duration of treatment prior to discontinuation) will be provided.

4.3.2 Protocol Deviations

Deviations from the protocol including violations of inclusion/exclusion criteria will be assessed as ‘minor’ or ‘major’. Major PDs are defined as those deviations from the protocol that are likely to have an impact on the subject’s right, safety, well-being, and/or the validity of the data for analysis. Minor deviations include all deviations from the protocol excluding those considered as major. The impact of major protocol deviations on the efficacy and/or safety results will be investigated by assessing the robustness of the study results and conclusions to the choice of analysis population (see Section 4.4), both including and excluding data potentially affected by major protocol deviations. Major PDs that will lead to the exclusion of a subject from the per-protocol (PP) population will be identified. The final determination of major protocol deviations and the exclusion of subjects from each of the analysis populations will be made prior to database lock.

Major PDs that could potentially impact efficacy analysis and exclude a subject from PP population are listed in Appendix 5. Refer to the study protocol deviation specification for more information.

The following protocol deviation summaries will be provided:

- Number and percentage of subjects with a major protocol deviation by type of deviation (Analysis population: ITT)
- Number and percentage of subjects with a major protocol deviation resulting in exclusion of subjects from PP analysis by type of deviation (Analysis population: ITT)

A by-subject listing of protocol deviations will be provided.

4.4 Analysis Populations

The following analysis populations will be used to analyze the data.

Intent-to-Treat Population (ITT):

The ITT population will consist of all randomized subjects.

Per Protocol Population (PP):

The PP population will comprise all subjects in the ITT population who have no major protocol deviations that would have a significant effect on the efficacy of the study treatment. The PP population will be finalized prior to the final database lock.

Safety Population (SAF):

The safety (SAF) population will comprise all subjects in ITT population who receive at least one dose of study drug.

PK Analysis Population (PKAP):

The PK analysis population will include all subjects in the SAF population who provide at least one post-baseline evaluable drug concentration value.

The efficacy summaries and analyses will be based on the ITT population, which is based upon the Intention-to-Treat (ITT) principle. In the event that a subject is allocated the incorrect study treatment as per the study randomization list, subjects will be summarized and analyzed ‘as randomized’ (i.e., by randomized treatment group). In the event that a subject is stratified incorrectly, ‘randomized stratum’ will be used rather than ‘actual stratum’.

For the primary and selected secondary efficacy endpoints, a sensitivity analysis will be performed on the PP population to assess the robustness of the study conclusions to the choice of analysis population. Subjects will be included in the analysis according to the treatment ‘as randomized’.

The safety summaries and analyses will be based on the SAF population. Randomized subjects will only be excluded if there is clear, documented evidence that the subject did not receive any study drug injection. In the event that a subject is allocated the incorrect study treatment as per the study randomization list, subjects will be summarized and analyzed ‘as treated’ (i.e., by allocated treatment group).

Upon database release, protocol deviation and analysis population outputs will be produced and will be reviewed by Sponsor. An analysis population classification meeting will be arranged to discuss the outputs and to decide which subjects and/or subject data will be excluded from certain analyses. Decisions made regarding the exclusion of subjects and/or subject data from analyses will be made prior to database lock and unblinding and will be documented and approved by Sponsor.

The following analysis population summaries will be provided:

- The number and percentage of subjects in each study population will be presented by study treatment group including total (Analysis population: ITT).

A by-subject listing of analysis population details will be provided. This listing will be presented by treatment group and will include: subject identifier, inclusion/exclusion flag for each population and reason for exclusion from each population. All subjects screened will appear on this listing.

The following derived and computed variables will be produced:

Variable Description	Computation Methods, Notes, or Equation(s)
Flag of ITT subjects	'Y' if subjects randomized into the study; otherwise 'N'
Flag of SAF subjects	'Y' if subjects in ITT and date and/or time of the first dose injection on Study Drug Administration form is non-missing; otherwise 'N'
Flag of PP subjects	'Y' if subjects in ITT who complete 24-week Treatment Period with the exception of major protocol violators who have significant effect on the efficacy; otherwise 'N'

4.5 Demographic and Other Baseline Characteristics

Descriptive summaries of demographic and other baseline conditions will include the following:

- Demographics - age, gender, race, ethnicity, height, weight, and body mass index (BMI)
- Medical history
- Previous and concomitant medications

4.5.1 Demographics and baseline characteristics

The following demographic and baseline characteristics will be summarized by descriptive statistics (Analysis population: ITT):



- Age at time of informed consent (in years, as a continuous variable)
- Age categorization (≤ 65 , > 65 years)
- Gender (Male, Female)
- Race (White, Black or African American, Asian, American Indian or Alaska Native, Hawaiian Native or Other Pacific Islander, Other)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- Weight at baseline (in kilograms)
- Height at screening (in centimeters)
- BMI (kg/m^2) calculated using weight and height at screening
 $\text{BMI} (\text{kg}/\text{m}^2) = \text{weight} (\text{kg}) / \text{height} (\text{m})^2$
- Baseline EASI
- Baseline IGA
- Baseline AD involvement of BSA (as a continuous variable)
- Baseline average and peak pruritus NRS (weekly average during the 7 days prior to day 1, as a continuous variable)
- Number of days with severe pruritus (based on PCS) during the 7 days prior to day 1 (≥ 3 , 3, 4, 5, 6, 7)
- Intrinsic or extrinsic AD (refer to Appendix 6)
- Total SCORAD

Stratification factor includes IGA (3 or 4). A summary table for the stratification factor will be provided to show if any discrepancies between what was reported through Interactive Voice Response System (IVRS)/Interactive Web Response System (IWRS) versus eCRF data (at baseline visit).

By-subject listings of demographic data and other baseline characteristics (as summarized above) will be provided.

4.5.2 Medical History

Medical history from the screening visit will be summarized by treatment group and overall. Medical history will be coded using the MedDRA version 19.0. The number and percentage of subjects experiencing at least one such diagnosis will be summarized by the MedDRA system organ class (SOC) and preferred term (PT) (Analysis population: ITT).

Summary (n, %) may be presented according to MedDRA groups for specific MH analyses (such as asthma, food allergy, allergic conjunctivitis, etc).

By-subject listings of medical history will be provided.

4.5.3 Prior and concomitant therapies

Start and stop dates of medications or medical and surgical procedures will be compared to the date of first dose of study drug to allow them to be classified as either Prior or Concomitant.



Medications or medical and surgical procedures that start and stop prior to the date of first dose of study drug will be classified as Prior. Prior medications that are taken from the date of screening to the date prior to randomization, regardless of start or stop, are also classified as Run-In Medications.

If start and/or stop dates of medications or medical and surgical procedures are missing or partial, the dates will be compared as far as possible with the date of first dose of study drug. Medications or medical and surgical procedures will be assumed to be Concomitant, unless there is clear evidence (through comparison of partial dates) to suggest that they started and stopped prior to the first dose of study treatment. If there is clear evidence to suggest that the medication or medical and surgical procedure started and stopped prior to the first dose of study drug, they will be assumed to be Prior.

Missing date information for prior and concomitant medications

For prior and concomitant medications, incomplete (i.e., partial missing) start dates and/or stop dates will be imputed. When the start date and the stop date are both incomplete for a subject, the start date will be imputed first.

Incomplete start date

The following rules will be applied to impute the missing numeric fields for an incomplete prior or concomitant medication start date. If the stop date is complete (or imputed) and the imputed start date is after the stop date, then the start date will be imputed using the stop date.

Missing Month and Day

- If the year of the incomplete start date is the same as the year of the first administration of the study drug, then the month and day of the first administration of the study drug will be assigned to the missing fields.
- If the year of the incomplete start date is before the year of the first administration of the study drug, *December 31* will be assigned to the missing fields.
- If the year of the incomplete start date is after the year of the first administration of the study drug, *January 1* will be assigned to the missing fields.

Missing Month Only

- If only the month is missing, then the day will be treated as missing and both the month and the day will be replaced according to the procedure above.

Missing Day Only

- If the month and year of the incomplete start date are the same as the month and year of the first administration of the study drug, then the day of the first administration of the study drug will be assigned to the missing day.
- If either the year of the incomplete start date is before the year of the date of the first administration of the study drug or if both years are the same but the month of the incomplete start date is before the month of the date of the first administration of the study drug, then the last day of the month will be assigned to the missing day.

- If either the year of the incomplete start date is after the year of the date of the first administration of the study drug or if both years are the same but the month of the incomplete start date is after the month of the date of the first administration of the study drug, then the first day of the month will be assigned to the missing day.

If the stop date is complete and the imputed start date as above is after the stop date, then the start date will be imputed by the stop date.

Incomplete stop date

The following rules will be applied to impute the missing numeric fields for an incomplete prior or concomitant medication stop date. If the date of the last administration of the study drug is missing, then replace it with the last visit date in the imputations described below. If the imputed stop date is before the start date (imputed or non-imputed start date), then the imputed stop date will be equal to the start date.

Missing Month and Day

- If the year of the incomplete stop date is the same as the year of the last administration of the study drug, then the month and day of the last administration of the study drug will be assigned to the missing fields.
- If the year of the incomplete stop date is before the year of the last administration of the study drug, *December 31* will be assigned to the missing fields.
- If the year of the incomplete stop date is after the year of the last administration of the study drug, *January 1* will be assigned to the missing fields.

Missing Month Only

- If only the month is missing, then the day will be treated as missing and both the month and the day will be replaced according to the above procedure.

Missing Day Only

- If the month and year of the incomplete stop date are the same as the month and year of the last administration of the study drug, then the day of the last administration of the study drug will be assigned to the missing day.
- If either the year of the incomplete stop date is before the year of the date of the last administration of the study drug or if both years are the same but the month of the incomplete stop date is before the month of the date of the last administration of the study drug, then the last day of the month will be assigned to the missing day.
- If either the year of the incomplete stop date is after the year of the date of the last administration of the study drug or if both years are the same but the month of the incomplete stop date is after the month of the date of the last administration of the study drug, then the first day of the month will be assigned to the missing day.

4.5.3.1 Prior and concomitant medications

Prior and concomitant medications will be coded using the World Health Organization Drug Dictionary (WHODD), version March 2016.

Prior and Concomitant medications will be summarized separately by Anatomical Therapeutic Chemical (ATC) level 2, level 3, and preferred name in frequency tables (Analysis population: ITT). Subjects with more than one medication in a given ATC level and preferred name will be counted only once in that category. The following summaries will be produced:

- Number and percentage of subjects who had Prior medications by ATC level 2, level 3, and preferred name
- Number and percentage of subjects who had Concomitant medications by ATC level 2, level 3, and preferred name
- Number and percentage of subjects who had medications during run-in period by ATC level 2, level 3, and preferred name

Number and percentage of subjects who had medications during treatment period by ATC level 2, level 3, and preferred name

A by-subject listing of all prior and concomitant medications will be provided.

4.5.3.2 Prior and concomitant medical and surgical procedures

Medical and surgical procedures will be coded using the MedDRA version 19.0.

Prior and Concomitant medical and surgical procedures will be summarized separately by SOC and PT in frequency tables (Analysis population: ITT). Subjects with more than one procedure in a given SOC and PT will be counted only once in that category. The following summaries will be produced:

- Number and percentage of subjects who had Prior medical/surgical procedures by SOC and PT
- Number and percentage of subjects who had Concomitant medical/surgical procedures by SOC and PT

A by-subject listing of all prior and concomitant medical and surgical procedures will be provided.

4.5.3.3 Rescue therapies

The number and percentage of subjects initiating use of rescue therapies will be summarized overall and by preferred name (for rescue medications) or preferred term (for rescue procedures) for all subjects in the ITT population. Rescue medication will be summarized by topical and systemic groups. Subjects with more than one rescue therapy in a given preferred name or preferred term will be counted only once in that category.

By-subject listings of all rescue therapies will be provided.

4.6 Treatment Compliance

Subject compliance with study treatment will be assessed via relative dose (%) of study drug. Refer to Section 4.8.1 for more analyses.

Study drug usage: Actual Dose (mg) of study drug at each planned dosing day is defined as

CCI

where: start volume in syringe = 1.0 mL.

Actual Total Dose (mg) is defined as the sum of the actual doses (mg) administered to a subject during a specified period.

Prescribed Total Dose (mg) is defined as the sum of the prescribed doses (mg) to a subject during a specified period.

Relative Dose (%) is defined as

CCI

The study drug compliance will be evaluated according to number of injections. Subjects with a compliance rate (number of actual injection / number of planned injection x 100) between 80% and 120% is considered as compliant with study drug during their participation in the study. Proportion of subjects considered as compliant with study drug will be summarized.

TCS usage: TCS are dispensed at screening, baseline, weeks 4, 8, 12, 16, 20, 24, 28, and returned at subsequent visit and weighted. Additional dispensation and weight may also occur at weeks 1 & 2, and/or unscheduled visits.

Actual Dose (g) at each planned dosing interval (one dispensation-return cycle) is defined as

CCI

In general, this will be considered as the monthly TCS usage. TCS dispensed at baseline, week 1 and 2 (returned at weeks 1, 2 and 4, respectively) will be summed up to get the TCS usage for the first month.

Actual Total Dose (g) of background TCS is defined as the sum of the actual doses (g) administered to a subject during a specific period (e.g., during the study treatment period). This information during the follow-up period should also be included. TCS usage will be presented by potency level and overall.

CCI

For missing TCS, the following rules will be followed:

- Dispensed weight: missing; returned weight: not missing
Impute the missing dispensed weight by the weight of another intact tube (entered by site).

- Dispensed weight: missing; returned weight: missing
The worst case scenario is that the subject has used the entire tube. Therefore, consider to use the content of the entire tube (45, 50, 15 or 30g depending on the actual TCS) instead of imputing the missing data. So the result of subtracting returned weight from dispensed weight is the content of the tube i.e. return weight imputed as 0.

- Dispensed weight: not missing; returned weight: missing
To apply the same way of imputation as above, using the total content of the tube as the weight of TCS used.

By-subject listings of study drug administration and compliance and TCS usage will be provided.

4.7 Efficacy Evaluation

Primary inference for efficacy analysis will be based on the ITT population at week 24.

The following efficacy endpoints will be analyzed at each visit up to week 24:

- Percent change from baseline in EASI
- Proportion of subjects achieving IGA success (defined as IGA 0 [clear] or 1 [almost clear])
- Percent change from baseline in weekly average of the peak pruritus NRS
- Proportion of subjects with an improvement of weekly average pruritus peak NRS ≥ 4
- Proportion of subjects with EASI-50, EASI-75 or EASI-90 (defined as achieving 50%, 75% or 90% reduction from baseline in EASI score)

All other efficacy endpoints (except above) will be analyzed at week 24 and summarized descriptively for the rest of visits.

The follow-up data for all efficacy endpoint will be summarized descriptively only.

Sensitivity analysis and subgroup analysis (where applicable) will be carried out for Week 24 only.

4.7.1 Analysis and Data Conventions

This study is designed to test for superiority. The null hypothesis is that there is no dose response effect for nemolizumab (10 mg, 30 mg, 90 mg) in percent change from baseline in EASI at week 24. The alternative hypothesis will be that there is a dose response.

4.7.1.1 Multi-center Studies

For the purpose of the summaries and analyses, the term ‘Center’ will be used to define each investigator site.

If at end the impact of the center(s) justified, the ad hoc analyses will be performed by including ‘Center’ as a main effect term in the model. In that case, the centers may be grouped by geographic regions (Europe, North America, Asia-Pacific). Evaluation of the consistency of treatment effects across geographic regions will be performed via subgroup analysis (see Section 4.7.1.6).

4.7.1.2 Adjustments for Covariates

The primary efficacy analysis will be adjusted for the following baseline covariates:

- Baseline IGA severity (the stratification factor)
- Baseline EASI

4.7.1.3 Handling of Missing Data

Use of rescue therapy: For the purpose of efficacy analysis, subjects receiving any rescue therapies will be considered as treatment failures. All efficacy data, except Observed Case (OC), will be set to missing after rescue medication is used. In OC analysis, observed data after subject has received rescue treatment will be included.

Diary data: If a subject has less than 4 diary entries in a week, the weekly prorated average scores will be set to missing.

The following methods will be used to impute the missing values:

Continuous endpoints: To impute the missing values for the primary analysis of continuous endpoints, Mixed-effect Model for Repeated Measures (MMRM) approach will be used for the primary and secondary endpoints (where applicable).

In addition, the following imputation approaches will be carried out as sensitivity analyses for primary and selected secondary endpoints to assess the robustness of findings. For sensitivity analyses, continuous endpoints will be analysed using analysis of covariance (ANCOVA) including terms of treatment group, baseline IGA severity (randomization strata), and the corresponding baseline value.

1. **Multiple imputation (MI) method under assumption missing at random (MAR):**

- a. Imputation phase: The missing data are filled in m times to generate m complete data sets. The selection of m depends on the required computing time. A reasonable range is 20 to 100 times, or may follow a rule of thumb that the number of imputations should be similar to the percentage of cases that are incomplete.
- b. Analysis phase: The m complete data sets are analyzed by using the same method as the corresponding primary analysis.

- c. Pooling phase: The results from the m complete data sets are combined for the inference.
2. **Last Observation Carried Forward (LOCF):** The last observed post-baseline value will be used to replace missing values for continuous variables. For questionnaire-based variables, LOCF will be applied to missing individual questions first, and the value of the total (or sub-scale) score will be derived based on the imputed individual questions.
3. **Worst Observation Carried Forward (WOCF):** The worst observed post-baseline value will be used to replace missing values for continuous variables. For questionnaire-based variables, WOCF will be applied to missing individual questions first, and the value of the total (or sub-scale) score will be derived based on the imputed individual questions.
4. **Observed Case (OC):** No data will be imputed. The observed values are used in analysis, including assessments post rescue medication.
5. **Pattern-mixture model under assumption missing not at random (MNAR), controlled-based pattern imputation:** refer to the 3 phases for MI-based imputation above. With controlled-based pattern imputation, only observations in placebo treatment group are used to derive the imputation model.

Binary endpoints:

All missing values will be treated as a Non-Responder for the binary endpoints. If a subject withdraws from the study, all assessment after withdrawal will be considered as Non-Responder.

LOCF, OC, MI and Pattern-Mixture Model under missing not at random assumption approaches will be used as sensitivity analysis to impute the missing values for the selected secondary endpoints as appropriate. For sensitivity analyses, binary endpoints will be analysed using **Cochran-Mantel-Haenszel (CMH)** test stratified by baseline IGA severity.

There will be no imputations for missing laboratory and vital sign data.

4.7.1.4 Multiple Comparisons/Multiplicity

Multiple comparison procedures-modeling (MCP-Mod, Bretz 2005) approach handles the multiplicity using a set of candidate dose-response models to test for a dose-response relationship via model-associated statistics.

4.7.1.5 Planned Analyses

Interim Analysis: No interim analyses are planned.

Week 24 Analysis: Primary analysis will be carried out once all subjects have completed week 24 visit (excluding Follow-up) or have withdrawn from the study. No one directly involved with the

conduct of the study will see the unblinded data before the completion of the trial, in order to avoid biasing the remaining data of the study.

Final Analysis: The final analysis will be performed when the last subject has completed the study (i.e. completed the safety follow-up or has withdrawn from the study).

4.7.1.6 Examination of Subgroups

To evaluate the consistency of treatment effects, subgroup analyses will be explored for the primary and selected secondary efficacy endpoints based on:

- Region (Europe, North America, Asia-Pacific)
- Age (18-65 or >65)
- Gender (Male, Female)
- Race (White, Black, Asian, Other)
- Baseline IGA (3 or 4)
- Baseline EASI (\leq median or $>$ median)
- Intrinsic or extrinsic AD (based on IgE and eosinophil counts)

Subgroup analyses will be explored for safety endpoints (e.g., AEs) based on:

- Age (18-65 or >65)
- Gender (Male, Female)
- Race (White, Black, Asian, Other)

If the number of subjects in a subgroup is too small, subgroups may be pooled for analyses.

4.7.2 Primary Efficacy Endpoint

The primary variable for the assessment of efficacy is the percent change from baseline in EASI at week 24.

The primary endpoint will be analyzed using a MMRM approach, including terms of treatment group, and baseline IGA severity for ITT and PP populations. Visit will be fitted as a categorical variable, with the effect of treatment group and baseline IGA varying at each visit (i.e., interaction between treatment group and visit and interaction between baseline IGA and visit). The corresponding baseline values of the response variable will be included as a covariate in the model. An unstructured (UN) covariance matrix for repeated measures within a subject will be used to model the within-subject errors in the analysis, unless there are issues related to convergence. If the UN covariance matrix leads to non-convergence, then Akaike's information criterion (AIC) will be used to select the best covariance matrix among autoregressive, compound symmetric, and Toeplitz. The Kenward-Rogers adjustment will be used with restricted maximum likelihood (REML) to make proper statistical inferences. Denominator degrees of freedom will be estimated using approximation of Kenward-Rogers. A linear contrast will be used, within the MMRM framework, to estimate difference between each nemolizumab

dose regimen and placebo. LSmeans with 95% CI and the corresponding statistical p-value will be presented for the treatment difference between each nemolizumab dosing regimen and placebo at a given visit.

Other covariates may be explored and will be added into the model, if appropriate.

Dose Response Analysis: MCP-Mod approach will be applied using a set of candidate dose-response models (see table below) to test for a dose-response relationship via model-associated statistics. The analysis will be performed using the available R software package “MCPMod” or any other available software.

Linear	E _{max}	Sigmoid E _{max}
Linear	1. ED50=10mg 2. ED50=30mg	1. ED50=10mg, H (Hill)=1 2. ED50=10mg, H (Hill)=2 3. ED50=10mg, H (Hill)=3 4. ED50=15mg, H (Hill)=2 5. ED50=20mg, H (Hill)=4 6. ED50=30mg, H (Hill)=2

ED50=dose produces 50% of maximum effect

Sensitivity analysis: The robustness of the primary analysis will be investigated by performing the sensitivity analyses as described in 4.7.1.3.

Subgroup analysis: The homogeneity of the treatment effect for a number of important subgroups will be investigated as described in 4.7.1.6.

Least square mean (LSmean) of treatment difference and the respective 95% CIs will be used for all subgroup comparisons at week 24. In addition, the p-values for the interaction of the treatment groups and the subgroups will also be provided, wherever appropriate. The LSmean of treatment difference, respective 95% CI and interaction p-values will be based on ANCOVA including terms of treatment group, the corresponding baseline value, subgroup factor and interaction between treatment group and subgroup.

A by-subject listing of the primary efficacy data will be provided. In addition, line plots showing the mean percent change from baseline in EASI over time within each treatment group will be produced for ITT and PP populations

4.7.3 Secondary Efficacy Endpoints

For analyses up to week 24 using the MMRM approach for continuous endpoints, only data collected up to week 24 will be included in the model.



4.7.3.1 Proportion of subjects achieving IGA success (defined as IGA 0 [clear] or 1 [almost clear]) at each visit up to week 24

A CMH test stratified by baseline IGA severity (3, 4) will be used to test the difference between each nemolizumab regimen and the placebo for the proportion of subjects achieving certain “success” at a given visit for ITT and PP populations. A 95% CI for the treatment difference will be calculated.

A bar chart and line plots will be produced over time for each treatment group to summarize the data (“responders”).

Sensitivity analysis: Sensitivity analyses will be performed as described in 4.7.1.3.

Subgroup analysis: Subgroup analyses will be investigated as described in 4.7.1.6.

Logistic regression will be used with response status as the dependent variable and treatment group, subgroup factor and interaction between treatment group and subgroup as independent variables. Odds ratios and the respective 95% CIs will be used for all subgroup comparisons. In addition, the p-values for the interaction of the treatment groups and the subgroups will also be provided, wherever appropriate.

4.7.3.2 Proportion of subjects achieving IGA success (defined as IGA 0 [clear] or 1 [almost clear]) and a reduction of at least 2 points at each visit up to week 24

This endpoint will be analyzed as described in section 4.7.3.1.

4.7.3.3 Percent change from baseline in EASI at each visit up to week 24

This endpoint will be analyzed as the primary efficacy endpoint using a MMRM approach for ITT and PP populations.

4.7.3.4 Absolute and percent change from baseline in weekly average of the peak and average pruritus NRS at each visit up to week 24

The following endpoints will be analyzed as the primary efficacy endpoint using a MMRM approach for ITT and PP populations.

- Absolute change from baseline in weekly average of the peak pruritus NRS at each visit up to week 24
- Percent change from baseline in weekly average of the peak pruritus NRS at each visit up to week 24
- Absolute change from baseline in weekly average of the average pruritus NRS at each visit up to week 24
- Percent change from baseline in weekly average of the average pruritus NRS at each visit up to week 24

The plots showing the changes over time within each treatment group will be provided.

Sensitivity analysis: Sensitivity analyses will be performed as described in 4.7.1.3 for percent change from baseline in weekly average of the peak pruritus NRS for Week 24 only.

Subgroup analysis: Subgroup analyses will be performed as described in 4.7.1.6 for percent change from baseline in weekly average of the peak pruritus NRS for Week 24 only using the same method as subgroup analysis for the primary efficacy endpoint.

4.7.3.5 Proportion of subjects with an improvement of weekly average pruritus peak NRS ≥ 4 from baseline to week 24

The response status will be analyzed using a CMH test as described in Section 4.7.3.1 for ITT and PP populations.

A detailed validation plan of the peak NRS responder analysis using anchor-based approach is included in Appendix 7.

4.7.3.6 Absolute and percent change from baseline in weekly average sleep disturbance NRS at week 24

The following endpoints will be analyzed as the primary efficacy endpoint using a MMRM approach for ITT and PP populations.

- Absolute change from baseline in weekly average sleep disturbance NRS at week 24
- Percent change from baseline in weekly average sleep disturbance NRS at week 24

The plots showing the changes over time within each treatment group will be provided.

4.7.3.7 Proportion of subjects achieving PCS success (defined as a weekly prorated rounded average PCS ≤ 1 [none - mild]) at week 24

The response status will be analyzed using a CMH test as described in Section 4.7.3.1 for ITT and PP populations.

4.7.3.8 Proportion of subjects with EASI-50, EASI-75 or EASI-90 (defined as achieving 50%, 75% or 90% reduction from baseline in EASI score) at each visit up to week 24

The following response status will be analyzed using a CMH test as described in Section 4.7.3.1 for ITT and PP populations.

- Proportion of subjects with EASI-50 (defined as achieving 50% reduction from baseline in EASI score) at each visit up to week 24

- Proportion of subjects with EASI-75 (defined as achieving 75% reduction from baseline in EASI score) at each visit up to week 24
- Proportion of subjects with EASI-90 (defined as achieving 90% reduction from baseline in EASI score) at each visit up to week 24

Sensitivity analysis using MI approach will be performed.

A histogram over time will be produced.

4.7.3.9 Absolute and percent change from baseline in SCORAD at week 24

The following endpoints will be analyzed as the primary efficacy endpoint using a MMRM approach for ITT and PP populations.

- Absolute change from baseline in SCORAD at week 24
- Percent change from baseline in SCORAD at week 24

4.7.4 Other Efficacy Endpoints

4.7.4.1 Disease severity scores (EASI, IGA, SCORAD, BSA, pruritus NRS, sleep disturbance NRS and PCS) at each visit including at follow up visit (week 28 and 32)

Data will be summarized descriptively by visit (Analysis population: ITT).

4.7.4.2 Change and percentage change from baseline in disease severity scores (EASI, SCORAD, pruritus NRS, and sleep disturbance NRS) at each visit

Data will be summarized descriptively by visit (Analysis population: ITT).

4.7.4.3 Change and percentage change in disease severity scores (EASI, SCORAD, pruritus NRS, and sleep disturbance NRS) from week 24/ET to follow up visits (week 28 and 32)

Data will be summarized descriptively by visit (Analysis population: ITT).

4.7.4.4 Proportion of subjects with an improvement of weekly average pruritus NRS ≥ 4 from baseline to each visit

The number and percentage of “responders” will be summarized by visit (Analysis population: ITT).

4.7.4.5 Proportion of subjects achieving PCS success (defined as a weekly prorated rounded average PCS ≤ 1 [none - mild]) at each visit including follow-up visits

The number and percentage of “responders” will be summarized by visit (Analysis population: ITT).

4.7.4.6 DPS at hour 2, 4, 8, 24, 48, and 72

Data will be summarized descriptively at each timepoint (Analysis population: ITT).

4.7.4.7 Monthly and total amount of TCS used during the treatment period (from baseline to week 24)

Data will be summarized descriptively (Analysis population: ITT).

Monthly and total amount of TCS used during the follow-up will be summarized separately.

Subjects who received rescue medication will be excluded from all TCS analyses.

4.7.4.8 Time to rescue therapy

Time to rescue therapy (days) from date of randomization will be analyzed using KM and Cox proportional hazard model and displayed graphically when appropriate (Analysis population: ITT). Median event times and two-sided 95% CI for each median will be provided when appropriate. The KM survival curves will be displayed for each treatment group along with p-value for differences (each nemolizumab regimen over placebo) based on the Log-rank test stratified by baseline IGA severity (3, 4).

The estimates of the hazard ratio (HR) comparing each nemolizumab regimen with placebo will be derived using the Cox proportional hazard model for the time to event served as a response variable and treatment group, baseline IGA severity as factors. The HRs (each nemolizumab regimen over placebo), corresponding 95% CIs, and two sided p-values for the HRs will be presented.

Time to rescue therapy will be calculated for treatment period only. Subjects who do not experience the event at any time during the treatment period will be censored at the last evaluation time (up to the end of treatment period).

4.7.4.9 Change from baseline in 5-D itch scale total score at week 2

The endpoint will be analyzed using an analysis of covariance (ANCOVA), with treatment group, baseline IGA severity (3, 4) and baseline 5-D itch scale total score as covariates in the model for ITT population. A 95% CI for the difference between each nemolizumab dose regimen and placebo in LSM means and p-value will be calculated based on contrast test statistics.

4.7.4.10 Evaluation of the prurigo nodularis lesions over time

In subjects with lesions of prurigo nodularis at baseline, the lesions are evaluated at each subsequent visit (weeks 1, 2, 4, 8, 12, 16, 20, 24, 28, 32 or ET when applicable) using a 5-point scale (2: very much improved; 1: improved; 0: no change; -1: worsened; -2: very much worsened).

Number and percentage of subjects in each of five categories will be summarized by visit for ITT population.

4.8 Safety Evaluation

All safety summaries and analyses will be based upon the SAF population as defined in Section 4.4.

4.8.1 Extent of Exposure

The extent of treatment will be summarized by treatment group as follows:

- Duration of study drug (days)
- Cumulative dose of nemolizumab (mg) received
- Relative dose (%) of nemolizumab
- Number of nemolizumab administrations
- Proportion of subjects receiving scheduled study drug administrations at each study drug administration visit by treatment group

4.8.2 Adverse Events

Analysis of severity: The severity of each AE will be summarized as assessed by the investigator (Mild, Moderate, Severe). Within the same MedDRA PT, only the most severe AE for each subject will be counted in tabulations by severity. Within a MedDRA SOC, subjects with more than one MedDRA PT will be counted only once for the most severe AE reported. AEs for which the severity is missing will be imputed to be Severe; this imputation will take place prior to determining the most severe AE within a SOC or PT for a given subject.

Analysis of causality: The relationship of each AE to the study drug and/or study procedure will be summarized as assessed by the investigator (Reasonable Possibility, No Reasonable Possibility). Within the same MedDRA PT, only the AE with the highest ranked relationship to study drug and/or study procedure for each subject will be counted in tabulations by causality. Within a MedDRA SOC, subjects with more than one MedDRA PT will be counted only once for the AE that is most related to study drug and/or study procedure. AEs for which the relationship to study drug is missing will be considered as Reasonable Possibility related. The imputation for a missing relationship will take place prior to determining the most related AE within a SOC or PT for a given subject.

High-level summary of TEAEs will be presented based on 1) all causalities, 2) study drug related, and 3) study procedure (including TCS) related. The table will include the number and percentage of subjects for the following categories:

- Number of TEAEs
- Subjects with TEAEs
- Subjects with treatment-emergent SAEs (TESAEs)
- Subjects with severe TEAEs
- TEAE with fatal outcome
- Subjects temporarily discontinued from study treatment due to TEAEs
- Subjects permanently discontinued from study treatment due to TEAEs
- Subjects discontinued from study due to TEAEs

The number and percentage of subjects reporting any TEAE will be summarized by SOC and PT.

- Summary of TEAEs by MedDRA SOC and PT based on 1) all causalities; 2) study drug related, and 3) study procedure (including TCS) related
- Summary of Severe TEAEs by MedDRA SOC and PT based on 1) all causalities; 2) study drug related, and 3) study procedure (including TCS) related
- Summary of the most common TEAEs ($\geq x\%$) by MedDRA SOC and PT based on all causalities
- Summary of TEAEs by MedDRA SOC and PT based on all causalities for subgroups as described in Section 4.7.1.6

AE summaries will be ordered in terms of decreasing frequency for SOC, and PT within SOC, in the treatment group (combined, nemolizumab 90 mg, 30 mg, 10 mg, then placebo), and then alphabetically for SOC, and PT within SOC. If deemed as appropriate, a different order may be applied.

All AEs will be provided in a by-subject listing which will include both the term reported on the eCRF (verbatim term) and the PT and SOC to which it is coded. Relative start and stop days will be included along with the actual onset and resolution dates. Pre-treatment AEs will be listed separately.

4.8.3 Deaths, Other Serious Adverse Events, and Other Significant Adverse Events

Other significant TEAEs are those TEAEs reported as leading to permanent or temporary discontinuation of study treatment, leading to discontinuation of study, AESIs, and selected AEs.

The following summaries will be produced:

- Summary of TESAEs by MedDRA SOC and PT based on 1) all causalities; 2) study drug related, and 3) study procedure (including TCS) related

- Summary of TEAEs associated with permanent discontinuation of study treatment by MedDRA SOC and PT based on 1) all causalities; 2) study drug related, and 3) study procedure (including TCS) related
- Summary of AEs associated with discontinuation of study participation by MedDRA SOC and PT based on 1) all causalities; 2) study drug related, and 3) study procedure (including TCS) related
- Summary of AESIs based on 1) all causalities; 2) study drug related
- Summary of selected AEs based on 1) all causalities; 2) study drug related.

In addition, separate listings will be provided for the following:

- Deaths and other SAEs
- Permanent discontinuations of study treatment due to AEs
- Discontinuations of study participation due to AEs
- Temporary discontinuations of study treatment due to AEs
- AESIs
- Selected AEs

4.8.4 Clinical Laboratory Evaluation

All laboratory values will be reported in SI units.

Laboratory data (absolute values and change from baseline) will be summarized descriptively by visit and treatment group.

In addition, the number and percent of subjects below, within, and above the laboratory reference ranges and the number and percentage of subjects who met criteria of PCSV (see Appendix 8) will be summarized by treatment group. Shift tables will be generated using the reference ranges.

A by-subject listing of all laboratory data will be provided by treatment group, with abnormal values flagged. Laboratory reference ranges should also be listed.

4.8.5 Vital Signs, Weight, Physical Findings and Other Observations Related to Safety

4.8.5.1 Vital Signs and Weight

All vital signs and weight data (absolute values and change from baseline) will be summarized by visit and treatment group.

In addition, the number and percent of subjects with blood pressure (systolic and diastolic), pulse rate, and weight values meeting criteria for PCSV will be summarized by treatment group.

A by-subject listing of vital signs data will be provided, with potentially clinically significant values flagged.

4.8.5.2 Peak Expiratory Flow (PEF)

For subject with medical history of asthma, PEF measurements (absolute values and change from baseline for actual PEF and %predicted PEF) will be summarized by visit and treatment group.

Number and percentage of subjects who met the following criteria will be summarized.

- With %predicted PEF less than 80%
- With $\geq 20\%$ change (decrease) from baseline in PEF

A by-subject listing of respiratory assessments/PEF will be provided.

4.8.5.3 Electrocardiogram

A by-subject listing of ECG overall interpretation will be provided.

4.8.6 Safety Monitoring (IDMC)

An IDMC has been set-up to monitor safety data generated in this clinical trial on an ongoing basis and will make appropriate recommendations to the sponsor. The members of IDMC do not include any sponsor representative or any investigator of the study. During the clinical trial, the IDMC will review the Suspected Unexpected Adverse drug Reactions (SUSARs) and unblinded SAEs on an ongoing basis, other SAEs and AESI on a monthly basis, and accumulating safety data approximately every 4 months. The cut-off date for the first IDMC review meeting will occur once approximately 50 subjects have completed 3 months of treatment, or approximately 4 months after the first subject has been enrolled in the study (whichever is sooner). Unscheduled meetings may occur as needed at the discretion of Sponsor or the IDMC members. The IDMC will review data blinded by treatment group, but can unblind if judged to be necessary. Details on the IDMC, including the plan of analysis for IDMC outputs, the composition of the IDMC, the procedures, roles, responsibilities and their communications are provided in the IDMC charter.

4.9 Other Analyses

4.9.1 Pharmacokinetic parameters and anti-drug antibody analyses

The PK parameters derived using non-compartmental techniques will be regarded as primary endpoints for the PK analyses (see Section 3.2.4). Primary inference for all the PK parameters will be based on the PKAP population.

Descriptive statistics (n, arithmetic mean, SD, minimum, median, maximum) by treatment group will be calculated for all derived PK endpoints. When calculating the statistics on PK parameters, if the values of more than half of the measured subjects are BLQ, the statistics will not be summarized for that measurement time point, and NC (Not Calculated) will be displayed. Geometric means and between-subject coefficients of variation (CVb) will be calculated for log_e-transformed AUC_{0-28d}, AUC_{0-12w}, AUC_{0-24w}, AUC_{0-t}, AUC_{0-inf}, C_{trough} and C_{max} where:





where: SD is the standard deviation of the \log_e -transformed data.

Following \log_e -transformation, AUC_{0-28d} , AUC_{0-12w} , AUC_{0-24w} , AUC_{0-t} , C_{trough} and C_{max} will be analyzed separately using analysis of variance (ANOVA), fitting a model with treatment group as a fixed effect. The residual variance from the model will be used to calculate point estimates and 90% CIs for the least squares means for each treatment formulation on the \log_e scale. These estimates will be back transformed to give point estimates and 90% CIs on the original scale.

The concentration at each time point will be summarized as arithmetic mean, SD, median, minimum, and maximum, number of BLQs (Below the Limit of Quantification). PK parameters using geometric means will be compared to determine when steady state conditions are achieved during the treatment period.

The potential relationship between serum concentrations of nemolizumab and change in EASI, pruritus NRS, or other indicators of disease activity will be explored using PK/Pharmacodynamics (PD) modeling, as appropriate.

The following Exposure-Response analysis will be performed:

- Percent change in EASI vs individual serum concentrations (scatter plots):
 - at Week 4
 - AUC_{0-28d}
 - C_{trough} week 4
 - at Week 12
 - AUC_{0-12w}
 - C_{trough} week 12
 - At week 24
 - AUC_{0-24w}
 - C_{trough} week 24

- Percent change in weekly average of the peak pruritus NRS vs individual serum concentrations (scatter plots):
 - at Week 4
 - AUC_{0-28d}
 - C_{trough} week 4
 - at Week 12
 - AUC_{0-12w}
 - C_{trough} week 12
 - At week 24
 - AUC_{0-24w}
 - C_{trough} week 24



- Percent change in sleep disturbance NRS vs individual serum concentrations (scatter plots):
 - at Week 4
 - AUC_{0-28d}
 - C_{trough} week 4
 - at Week 12
 - AUC_{0-12w}
 - C_{trough} week 12
 - At week 24
 - AUC_{0-24w}
 - C_{trough} week 24
- Proportion of subjects achieving IGA success (defined as IGA 0 [Clear] or 1 [Almost Clear]) vs quartile of concentrations (bar plots):
 - at Week 12
 - AUC_{0-12w}
 - C_{trough} week 12
 - At week 24
 - AUC_{0-24w}
 - C_{trough} week 24
- EASI-75 response rate vs quartile of concentrations (bar plots):
 - at Week 12
 - AUC_{0-12w}
 - C_{trough} week 12
 - At week 24
 - AUC_{0-24w}
 - C_{trough} week 24
- Incidence of TEAEs in subjects treated with nemolizumab categorized by steady state trough serum concentration in 4 quartiles.

Incidence of positive ADA results will be summarized and plotted by treatment group (absolute occurrence and percent of subjects) at each visit. Incidence of treatment-related positive ADA (subjects who will become ADA positive only after Nemolizumab administration) will also be summarized and plotted by treatment group (absolute occurrence and percent of subjects) at each visit.

An interim PK and Exposure-Response analysis will be done with the bionalytical data available at Week 24 analysis (see Section 4.7.1.5). The interim analysis will include:

- Serum pharmacokinetic parameters will be listed by cohort, subject and day, and summarized by treatment.
- Serum Nemolizumab concentrations will be plotted by treatment group (individual PK profiles, mean PK profile)

- Descriptive statistics for calculated PK parameters will include: n, arithmetic mean, SD, median, minimum and maximum values.
- Geometric mean, geometric coefficient of variation (CVb), lower quartile (Q1), and upper quartile (Q3) will be presented for AUC_{0-28d} , AUC_{0-12w} , AUC_{0-24w} , AUC_{0-t} , C_{trough} and C_{max} . PK parameters will also be summarized by weight category. Analysis of PK parameters, AUC_{0-28d} , AUC_{0-12w} , AUC_{0-24w} , AUC_{0-t} , C_{trough} and C_{max} , using an analysis of variance (ANOVA) will be presented.
- The Exposure-Response analysis described above.
- Incidence of positive ADA results on individual PK profiles (plot with 3 different identifications for the negative ADA, for the treatment-related positive ADA and for the positive ADA).

4.9.2 Biomarker Analyses

Primary inference for all biomarker analyses (including, but not limited to, eosinophil TARC, IgE, IL6, IL8, and IL18) will be based on the observed cases. All biomarker variables (using logarithm transformation) will be summarized across treatments at each time point.

Any observations in trends within the Biomarker data may be explored with PK/PD correlations/models. All of this analysis will be considered exploratory. The outputs will be in a separate report.

The detailed analysis will be carried out by Biomarker group in GALDERMA, and will be reported separately.

4.9.3 Quality of Life and Productivity data Analyses

Primary study population for all QoL and productivity data analyses will be based on ITT population. The DLQI, HADS, EQ5D and productivity data will be summarized by treatment group and visit.

By-subject listings of these data should be provided.

4.9.3.1 Change from baseline in DLQI at each visit up to week 24

This endpoint will be analyzed as the primary efficacy endpoint using a MMRM approach.

In addition, the following analyses will be performed:

- A shift table (by visit) will be generated using the categories as described in Section 3.2.6.1.
- Change from baseline in each sub-scale (see Section 3.2.6.1) at each visit up to week 24 will be analyzed as the primary efficacy endpoint using a MMRM approach.

- Proportion of subjects achieving a change in DLQI score of at least 4 points at each visit up to week 24 will be analyzed using a CMH test as described in Section 4.7.3.1.

4.9.3.2 Change from baseline in HADS at each visit up to week 24

The following endpoints will be analyzed as the primary efficacy endpoint using a MMRM approach.

- Change from baseline in HADS anxiety sub-score at each visit up to week 24
- Change from baseline in HADS depression sub-score at each visit up to week 24

In addition, the following analyses will be performed using a CMH test as described in Section 4.7.3.1:

- Proportion of subjects with HADS anxiety sub-score ≥ 11 at each visit up to week 24
- Proportion of subjects with HADS depression sub-score ≥ 11 at each visit up to week 24

4.9.3.3 Missed work/school day in the past 4 weeks by visit

This endpoint will be analyzed using an analysis of variance (ANOVA), with treatment group and baseline IGA severity (3, 4) as covariates in the model. A 95% CI for the difference between each nemolizumab dose regimen and placebo in LSMeans and p-value will be calculated based on contrast test statistics.

Total number of missing days during the treatment period will be summarized by treatment group.

4.9.3.4 Change from baseline in EQ5D at week 24

The following endpoints will be analyzed using ANCOVA, with treatment group, baseline IGA severity (3, 4) and baseline values of the corresponding response variable as covariates in the model. A 95% CI for the difference between each nemolizumab dose regimen and placebo in LSMeans and p-value will be calculated based on contrast test statistics.

- Change from baseline in EQ5D index at week 24
- Change from baseline in EQ5D VAS at week 24
- Change from baseline in each dimension score of EQ5D at week 24

In addition, number and percentage of subjects in each level of each dimension of EQ5D will be summarized at week 24. The difference between each nemolizumab dose regimen and placebo will be tested using CMH tests based on ridit scores stratified by baseline IGA severity (3, 4).

4.10 Determination of Sample Size

Sample size is estimated using MCP-Mod approach by selecting a set of candidate dose response shapes based on EASI response in previous Phase 2a study (CIM003JG study report 2017). Refer to Section 4.7.2 on the models considered as candidates in the estimation of sample size.

Assuming a placebo response of 35%, standard deviation 45% and a maximum treatment difference of 30%, in terms of percent change from baseline in EASI, 50 subjects per arm will provide at least 90% power to detect a statistically significant dose-response for at least one model at 1-sided significance level of 0.025.

4.11 Changes in the Conduct of the Study or Planned Analysis

Changes in the conduct of the study may be instituted through a protocol amendment. Planned analyses will be revised as appropriate. Similarly, planned analyses may be changed as a result of planned blinded data reviews. Changes will be finalized prior to database lock.

The following endpoints are added:

- Secondary efficacy endpoint: Proportion of subjects achieving IGA success (defined as IGA 0 [clear] or 1 [almost clear]) and a reduction of ≥ 2 points at each visit up to week 24
- Other efficacy endpoint: Proportion of subjects achieving PCS success (defined as a weekly prorated rounded average PCS ≤ 1 [none - mild]) at each visit including follow-up visits

The following endpoint is revised:

- Secondary efficacy endpoint: Proportion of subjects with EASI-50, EASI-75 or EASI-90 (defined as achieving 50%, 75% or 90% reduction from baseline in EASI score) at each visit up to week 24 (at week 24 per the protocol)

5 REFERENCES

Bretz F et al. Combining multiple comparisons and modeling techniques in dose-response studies. *Biometrics*, 2005; 61: 738–748.

European task force on atopic dermatitis. Severity scoring of atopic dermatitis: the SCORAD index: consensus report of the European task force on atopic dermatitis. *Dermatology* 1993; 186: 23-31.

Appendix 1 Schedule of assessments

Study Period Visit (V)	Run-in V1 ^a	V2 ^b	V3	V4	V5 ^b	Treatment V6 ^b	V7 ^b	V8 ^b	V9 ^b	V10 ^b	Follow-up V11	Follow-up V12/FINA L ^b	W32 D225	Unscheduled visit ^{b,c} (if applicable)	Early Termination Visit ^{b,d} (if applicable)
Week (W)	Screening	Baseline	W1	W2	W4	W8	W12	W16	W20	W24	W28	W32			
Day (D)	g/D -28 to -15	Baseline	D8	D15	D29	D57	D85	D113	D141	D169	D197	D225			
Visit window (±d)			± 1d	± 1d	± 1d	± 3d	± 5d	± 5d	± 5d	± 5d	± 7d				
ICF and country specific consent form (if applicable)	X														
Demographics	X														
Medical history/ Prior therapies	X														
Concomitant therapies	X	X	X	X	X	X	X	X	X	X	X	X	X	(X)	X
Inclusion/Exclusion criteria	X	X													
EFFICACY/SUBJECT-REPORTED OUTCOMES ASSESSMENTS															
DLQI	X	X		X		X	X			X					X
HADS	X	X				X	X			X					X
EQ5D	X	X				X	X			X					X
5-D itch	X	X		X		X	X			X					X
Missed work/school day				X		X	X			X					X
Pruritus NRS (daily) ^e	X	X		X		X	X			X					X
PCS (daily) ^e	X	X		X		X	X			X					X
Sleep disturbance NRS (daily) ^e	X	X		X		X	X			X					X
EASI	X	X	X	X	X	X	X	X	X	X	X	X	X	(X)	X
IGA	X	X	X	X	X	X	X	X	X	X	X	X	X	(X)	X
BSA	X	X	X	X	X	X	X	X	X	X	X	X	X	(X)	X
SCORAD	X	X	X	X	X	X	X	X	X	X	X	X	X	(X)	X
DPS ^f	X	X		X		X	X			X					X



CONFIDENTIAL

Study Period Visit (V)	Run-in V1 ^a	Treatment V5 ^b - V6 ^b	V7 ^b	V8 ^b	V9 ^b	V10 ^b	Follow-up V11 ^b - V12 ^b /FINA ^L _b	Unscheduled visit ^{b,c} (if applicable)	Early Termination Visit ^{b,e} (if applicable)	
Week (W) Day (D)	Screening g/D -28 to -15	V4 W2 D15 ± 1d	V5 W4 D29 ± 1d	V6 W8 D57 ± 3d	V7 W12 D85 ± 5d	V8 W16 D113 ± 5d	V9 W20 D141 ± 5d	V10 W24 D169 ± 5d	W28 D197 ± 7d	W32 D225 ± 7d
Visit window (±td)										
LABORATORY/SAFETY ASSESSMENTS										
AE recording (including review of laboratory values if applicable)	X	X	X	X	X	X	X	(X)	X	
Physical examination (including the evaluation of prurigo nodularis lesions if applicable)	X	X	X	X	X	X	X	(X)	X	
Height	X			X	X	X		(X)		
Weight	X			X	X	X		(X)		
ECG	X			X	X	X		(X)	X	
Vital signs	X	X	X	X	X	X	X	(X)	X	
Respiratory assessment ^g	X	X	X	X	X	X	X	(X)	X	
Urinalysis	X	X	X	X	X	X	X	(X)	X	
Hematology	X	X	X	X	X	X	X	(X)	X	
Blood chemistry	X	X	X	X	X	X	X	(X)	X	
TB test	X			X	X	X		(X)		
Hepatitis B and C test	X			X	X	X		(X)		
HIV test	X			X	X	X		(X)		
Pregnancy test ^h	X			X	X	X		(X)		
TARC	Serum	Urine	Urine	Urine	Urine	Urine	Urine	(Urine)	Urine	
Blood sample for biomarkers (including IgE)	X	X	X	X	X	X	X	(X)		
Stratum corneum sample for biomarkers	X	X	X	X	X	X	X			



CONFIDENTIAL

Study Period Visit (V)	Run-in V1 ^a	V2 ^b	V3	V4	V5 ^b	Treatment V6 ^b	V7 ^b	V8 ^b	V9 ^b	V10 ^b	Follow-up V11 ^b	Unscheduled visit ^{b,c} (if applicable)	Early Termination Visit ^{b,e} (if applicable)
Week (W)	Screening	Baseline	W1	W2	W4	W8	W12	W16	W20	W24	W28	W32	
Day (D)	g/D -28 to -15	e/D 1	D8	D15	D29	D57	D85	D113	D141	D169	D197	D225	
Visit window (±d)			± 1d	± 1d	± 1d	± 3d	± 5d	± 5d	± 5d	± 5d	± 7d	± 7d	

PK AND ADA ASSESSMENTS

PK samples
 ADA samples

STUDY DRUG AND BACKGROUND THERAPIES

Randomization
 Subcutaneous study drug injection
 TCS dispensing (D) / Return (R) and moisturizer dispensing if necessary

^a Screening visit must be performed at least 14 days prior to the day 1 visit to ensure that each subject undertakes at least 14 days of TCS treatment prior to randomization.
^b Subjects are required to fast for at least 8 hours before the visit.
^c Assessments to be conducted at the unscheduled visit depend on the reason for the visit. PK and ADA analyses are obligatory at unscheduled visits for safety reasons.
^d Subjects discontinued before the week 24 visit should attend an early termination visit and a final visit 12 weeks after the last study drug injection.
^e Pruritus NRS and PCS to be recorded by subjects once daily in the evening. Sleep disturbance NRS to be recorded by subjects once daily in the morning.
^f DPS to be recorded by subjects at 2, 4, 8, 24, 48, and 72 hours after the study drug administration on day 1.
^g PEF will be measured for subjects with a medical history of asthma only.
^h Only for females of childbearing potential. Serum pregnancy test to be performed at screening visit, and urine pregnancy test for all other visits.
ⁱ Stratum corneum sample from lesional and non-lesional skin on day 1 and from lesional skin only at week 16.



CONFIDENTIAL

Appendix 2 Criteria for potentially clinically significant vital signs and weight

Pulse rate

≤ 50 bpm and decrease from baseline ≥ 20 bpm

≥ 120 bpm and increase from baseline ≥ 20 bpm

To be applied for all positions (including missing).

SBP

≤ 95 mmHg and decrease from baseline ≥ 20 mmHg

≥ 160 mmHg and increase from baseline ≥ 20 mmHg

To be applied for all positions (including missing).

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).

Weight

$\geq 5\%$ increase from baseline

$\geq 5\%$ decrease from baseline

Appendix 3 Dermatology Life Quality Index (DLQI)

The DLQI questionnaire is designed for use in adults, i.e. subjects over the age of 16. It is self-explanatory and can be simply handed to the subject who is asked to fill it in without the need for detailed explanation. It is usually completed in one to two minutes.

Scoring

The scoring of each question is as follows:

Response	Score
Very much	scored 3
A lot	scored 2
A little	scored 1
Not at all	scored 0
Not relevant	scored 0
Question unanswered	scored 0
Question 7: "prevented work or studying"	scored 3

The DLQI is calculated by summing the score of each question resulting in a maximum of 30 and a minimum of 0. The higher the score, the more quality of life is impaired. The DLQI can also be expressed as a percentage of the maximum possible score of 30.

****Please Note:** That the scores associated with the different answers should not be printed on the DLQI itself, as this might cause bias**

Meaning of DLQI Scores

0-1 = no effect at all on subject's life

2-5 = small effect on subject's life

6-10 = moderate effect on subject's life

11-20 = very large effect on subject's life

21-30 = extremely large effect on subject's life

Detailed analysis of the DLQI

The DLQI can be analyzed under six headings as follows:



Section	Questions	Score
Symptoms and feelings	Questions 1 and 2	Score maximum 6
Daily activities	Questions 3 and 4	Score maximum 6
Leisure	Questions 5 and 6	Score maximum 6
Work and School	Question 7	Score maximum 3
Personal relationships	Questions 8 and 9	Score maximum 6
Treatment	Question 10	Score maximum 3

The scores for each of these sections can also be expressed as a percentage of either 6 or 3.

Interpretation of incorrectly completed questionnaires

There is a very high success rate of accurate completion of the DLQI. However, sometimes subjects do make mistakes.

- If 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 two or more questions are left unanswered the questionnaire is not scored.
- 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. If "Not relevant" is ticked, the score for Question 7 is 0. If it is answered 'no', but the second half is left incomplete, the score will remain 0.
- If two or more response options are ticked, the response option with the highest score should be recorded.
- If there is a response between two tick boxes, the lower of the two score options should be recorded.
- The DLQI can be analyzed by calculating the score for each of its six sub-scales (see above). When using sub-scales, if the answer to one question in a sub-scale is missing, that sub-scale should not be scored.

Minimal Clinically Important Difference of the DLQI

For **general inflammatory skin conditions** a change in DLQI score of at least 4 points is considered clinically important (Basra et al, 2015). This means that a subject's DLQI score has to either increase or decrease by at least 4 points in order to suggest that there has actually been a

meaningful change in that subject's quality of life since the previous measurement of his/her DLQI scores.

Reference

Basra MK et al. Determining the minimal clinically important difference and responsiveness of the Dermatology Life Quality Index (DLQI): further data. *Dermatology*, 2015; 230(1):27-33.

Appendix 4 EQ5D Index Score

The EQ5D consists of the EQ5D descriptive system and the EQ VAS. The descriptive system is made up of 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has 3 levels: no problems, some problems, extreme problems (see scoring algorithm below). A unique health state is defined by combining 1 level from each of the 5 dimensions. There are 243 possible health states defined on this way. Each state is referred to in terms of a 5 digit code. Any missing values which are coded as '9' will be set to missing (and hence no health state value will be calculated). Ambiguous values (e.g., 2 boxes ticked for a single dimension) should be treated as missing values. If 1 or more domains are missing the health state value (HSV) will not be calculated.

For example, if a subject provides the following responses: Mobility=No problems, Self-care=No problems, Usual Activities=Some problems, Pain/Discomfort=Moderate problems, Anxiety/Depression=Extreme problems, his response sequence is 11223. The 243 theoretical possible sequences can then be mapped to a HSV to provide a summary across all dimensions. The tariff to be used for the mapping is that given by Dolan (1997). The scoring algorithm is as follows (the value set for UK is used in calculating EQ5D on this study):

Full health (11111)	1.000
Constant term (for any dysfunctional state)	-0.081
Mobility level 1	0
Mobility level 2	-0.069
Mobility level 3	-0.314
Self-care level 1	0
Self-care level 2	-0.104
Self-care level 3	-0.214
Usual activities level 1	0
Usual activities level 2	-0.036
Usual activities level 3	-0.094
Pain/discomfort level 1	0
Pain/discomfort level 2	-0.123
Pain/discomfort level 3	-0.386
Anxiety/depression level 1	0
Anxiety/depression level 4	-0.071
Anxiety/depression level 3	-0.236
At least one domain at 3 (N3)	-0.269

The HSV for the subject with example sequence 11223 = 0.255, calculated as follows:

- 1) Full health =1
- 2) Minus constant: -0.081
- 3) Minus mobility level 1: 0
- 4) Minus self-care level 1: 0
- 5) Minus usual activities level 2: -0.036
- 6) Minus pain/discomfort level 2: -0.123
- 7) Minus anxiety/depression level 3: -0.236
- 8) Minus N3: -0.269

Where level 1 corresponds to no problems, level 2 to some problems, and level 3 to extreme problems.

Reference

Dolan P. Modeling valuations for EuroQol health states. Med Care, 1997; 35(11):1095-108.

Appendix 5 Protocol deviations

A list of major PDs that could potentially impact efficacy analysis and may exclude a subject from PP population is as follows:

- Study treatment administration errors
- Eligibility criteria not met
- Disallowed/prohibited concomitant therapies
Note: The programming produces a draft list of potential prohibited medications based on ATC codes for the medical/clinician review on an ongoing basis; a final list of protocol deviations (includes prohibited medications) is determined before the database lock after a joint team review and included into the relevant dataset.
- EASI assessment was not performed at week 24 visit as per the protocol
- Blinding has been compromised
- Did not sign ICF at all
- Other reasons, as outlined in the most recent version of the Protocol Deviation Specification document

Final determination of protocol deviations as major or minor will occur during blinded review prior to the unblinding of the study.

Appendix 6 Intrinsic or extrinsic AD

Classification of intrinsic or extrinsic AD will be based on IgE and/or eosinophil counts and will be defined after reviewing the overall data before DBL.

Appendix 7 Validation of the NRS responder analysis

To assess the appropriateness of responder thresholds in the NRS responder analysis (an improvement of weekly average pruritus peak NRS ≥ 4), an approach combines an anchor-based approach and ROC curves will be performed to identify the responder thresholds that best predict classification based on an external criterion. First, an external anchor or another external criterion that is appropriate to identify responders (e.g., Pruritus PCS) is used to classify all subjects into responder (at least 1 grade improvement from baseline in PCS) or non-responder groups. Then, to check for appropriateness, the relationship between the PCS criterion and the NRS measure (change from baseline in weekly average pruritus peak NRS) is examined through correlation analyses. Finally, the predictive accuracy of how well the specific NRS change scores relate to the classifications is evaluated using logistic regression analyses and graphically displayed as an ROC curve. Each point on the curve provides the sensitivity (true positives) versus one minus the specificity (false positives) trade-off for identifying responders for a specific unit of change on the NRS measure. The entire range of change scores and the likelihood classifications are provided within the one curve. A diagonal line is included in the figure as a reference. Evaluations that produce curves close to the diagonal line should be viewed with caution because the diagonal line represents chance classification.

The area under curve (AUC) and its 95% CI will be presented, along with percent concordant and discordant.

The analyses described above will be performed using all data across all visits and the data at week 24.

The following cumulative distribution function (CDF) plots will be produced to illustrate the between treatment group differences and within-subject change from baseline to Week 24.

- CDF plot illustrating the treatment differences from baseline to Week 24
- CDF plot of peak pruritus change score from baseline to Week 24 for all subjects by the different PCS response options at Week 24
- CDF plot of peak pruritus change score from baseline to Week 24 for all subjects by the PCS change scores (differences) between baseline and Week 24 (e.g. ± 3 , ± 2 , ± 1 , and 0 point change)

In addition, the mean change in peak pruritus NRS from baseline to Week 24 will be calculated in the following subjects:

- An improvement of ≥ 1 in PCS from baseline to week 24
- 50-74%, 75-89%, or 90-100% improvement in EASI from baseline to week 24
- An improvement of ≥ 2 in IGA scale from baseline to Week 24
- An IGA of 0 or 1 at Week 24

Appendix 8 Potentially clinically significant values (PCSV) for clinical laboratory tests

Parameter	SI Units	Conventional Units	PCSV
			<3.0 Giga/L & ≥3.0G/L at Baseline (non-Black)
WBC	x10e9/L	x10e3/μL	<2.0 Giga/L & ≥2.0G/L at Baseline (Black) : ≥16.0 G/L & <16G/L at Baseline :
Neutrophils ABS	x10e9/L	x10e3/μL	<1.5 Giga/L & ≥1.5G/L at Baseline (non-Black) <1.0 Giga/L & ≥1.0G/L at Baseline (Black)
Eosinophils ABS	x10e9/L	x10e3/μL	(>0.5 Giga/L and >ULN) and (≤0.5 Giga/L or ≤ULN at Baseline)
Basophils ABS	??	??	>0.1 Giga/L and ≤0.1 Giga/L at Baseline
Monocytes ABS	??	??	>0.7 Giga/L and ≤0.7 Giga/L at Baseline
Lymphocytes ABS	x10e9/L	x10e3/μL	>4.0 Giga/L and ≤4.0 Giga/L at Baseline <0.5 Giga/L and ≥0.5 Giga/L at baseline ≤100 g/L and Baseline >100 g/L for Male ;
Hemoglobin *Female *Male	g/L	g/dL	≤ 90 g/L and Baseline >90 g/L for Female ≥200 g/L and baseline <200 g/L for Male; ≥180 g/L and Baseline <180 g/L for Female
Platelets	x10e9/L	x10e3/μL	<100 Giga/L and ≥100G/L at Baseline ≥700 G/L & <700G/L at Baseline ≤130mmol/L and Baseline >130
Sodium	mmol/L	mmol/L	: ≥155 mmol/L and Baseline <155 <3mmol/L and Baseline ≥3
Potassium	mmol/L	mmol/L	>6.0 mmol/L and Baseline ≤6.0



Parameter	SI Units	Conventional Units	PCSV
Calcium (Total)			<7 mg/dL and Baseline ≥7 mg/dL
Corrected serum calcium	mmol/L	mg/dL	>12.5 mg/dL and Baseline ≤12.5 mg/dL
Glucose	mmol/L	mg/dL	<2.2mmol/L and ≥2.2 mmol/L at Baseline >13.9mmol/L and ≤13.9mmol/L at Baseline
Creatinine	µmol/L	mg/dL	≥150µmol/L (Adults) and Baseline <150µmol/L
AST	U/L		>3 ULN and Baseline ≤3 ULN
ALT	U/L		>3 ULN and baseline ≤3 ULN
Alkaline phosphatase	U/L		>2.5 ULN and Baseline ≤2.5 ULN
Bilirubin total	µmol/L	mg/dL	>2.0 ULN and Baseline ≤2.0 ULN
CPK	U/L		: >3.0 ULN and Baseline ≤3.0 ULN
Fibrinogen	g/L	mg/dL	Not provided
GGT	U/L		>2.5 ULN and Baseline ≤2.5 ULN
Albumin	g/L	g/dL	≤30 g/L and Baseline >30g/L
Uric acid			>600 µmol/L and ≤600 µmol/L at Baseline for Male;
*Male16-59			>500 µmol/L and ≤500 µmol/L at Baseline for Female
*Female>16	mmol/L	mg/dL	
*Male >60			
Cholesterol total	mmol/L	mg/dL	>10.34 mmol/L and ≤10.34 mmol/L at Baseline
Triglycerides	mmol/L	mg/dL	>5.7 mmol/L and ≤5.7 mmol/L at Baseline



CCI

Electronic Signature Page

This page is the manifestation of the electronic signature(s) used in compliance with CCI CCI electronic signature policies and procedures and in compliance with applicable regulations.

UserName: PPD
Title: Senior Statistical Programmer, CCI
Date: Tuesday, 21 August 2018, 08:59 AM GMT Standard Time
Meaning: Document contents approved.

UserName: PPD
Title: Biostatistician II, CCI
Date: Tuesday, 21 August 2018, 01:34 PM GMT Standard Time
Meaning: Document contents approved.

UserName: PPD
Title: Medical Director, CCI
Date: Tuesday, 21 August 2018, 02:58 PM GMT Standard Time
Meaning: Document contents approved.

UserName: PPD
Title: CCI Senior Scientist, CCI
Date: Tuesday, 21 August 2018, 03:33 PM GMT Standard Time
Meaning: Document contents approved.

UserName: PPD
Title: Principal Biostatistician, CCI
Date: Tuesday, 21 August 2018, 03:39 PM GMT Standard Time
Meaning: Document contents approved.
