

**Janssen Research & Development**

**Statistical Analysis Plan (Week 56 DBL)**

---

**A Multicenter, Randomized, Double-blind, Placebo-controlled, Proof-of-Concept Study of Ustekinumab in Subjects with Active Systemic Lupus Erythematosus**

---

**Protocol CNTO1275SLE2001; Phase 2a**

**STELARA® (ustekinumab)**

**Status:** Approved  
**Date:** 22 March 2018  
**Prepared by:** Janssen Research & Development, LLC  
**Document No.:** EDMS-ERI-159968603

**Compliance:** The study described in this report was performed according to the principles of Good Clinical Practice (GCP).

**Confidentiality Statement**

The information in this document contains trade secrets and commercial information that are privileged or confidential and may not be disclosed unless such disclosure is required by applicable law or regulations. In any event, persons to whom the information is disclosed must be informed that the information is privileged or confidential and may not be further disclosed by them. These restrictions on disclosure will apply equally to all future information supplied to you that is indicated as privileged or confidential.

## TABLE OF CONTENTS

<b>TABLE OF CONTENTS .....</b>	<b>2</b>
<b>AMENDMENT HISTORY .....</b>	<b>4</b>
<b>ABBREVIATIONS .....</b>	<b>4</b>
<b>1. INTRODUCTION.....</b>	<b>6</b>
1.1. Trial Objectives .....	6
1.2. Trial Design .....	6
1.3. Statistical Hypotheses for Trial Objectives.....	8
1.4. Sample Size Justification .....	8
1.5. Randomization and Blinding .....	8
<b>2. GENERAL ANALYSIS DEFINITIONS .....</b>	<b>10</b>
2.1. Visit Windows.....	10
2.2. Pooling Algorithm for Analysis Centers.....	10
2.3. Analysis Sets.....	11
2.3.1. Efficacy Analysis Set(s) .....	11
2.3.1.1. Primary Efficacy Analysis Set.....	11
2.3.2. Safety Analysis Set.....	11
2.3.3. Pharmacokinetics Analysis Set(s) .....	11
2.3.4. Immunogenicity Analysis Set.....	11
2.3.5. Pharmacodynamics Analysis Set .....	11
2.4. Definition of Subgroups.....	11
<b>3. INTERIM ANALYSIS AND DATA MONITORING COMMITTEE REVIEW.....</b>	<b>12</b>
<b>4. SUBJECT INFORMATION.....</b>	<b>12</b>
4.1. Demographics and Baseline Characteristics .....	12
4.2. Disposition Information.....	12
4.3. Treatment Compliance.....	12
4.4. Extent of Exposure .....	13
4.5. Protocol Deviations .....	13
4.6. Prior and Concomitant Medications .....	13
<b>5. EFFICACY .....</b>	<b>13</b>
5.1. Analysis Specifications.....	13
5.1.1. Level of Significance.....	13
5.1.2. Data Handling Rules.....	14
5.1.3. General Analysis Methods.....	14
5.1.4. Treatment Failure Criteria.....	14
5.1.5. Estimands .....	15
5.2. Efficacy Variable(s) .....	16
5.2.1. Definitions .....	17
5.2.2. Analysis Methods.....	21
<b>6. SAFETY .....</b>	<b>22</b>
6.1. Adverse Events .....	22
6.2. Clinical Laboratory Tests.....	24
6.3. Vital Signs and Physical Examination Findings .....	24
6.4. Electrocardiogram .....	25
<b>7. PHARMACOKINETICS/IMMUNOGENICITY/PHARMACODYNAMICS .....</b>	<b>25</b>
7.1. Pharmacokinetics.....	25
7.2. Immunogenicity (Antibodies to ustekinumab).....	26

---

7.3.	Pharmacodynamics.....	26
7.4.	Pharmacokinetic/Pharmacodynamic Relationships.....	27
<b>8.</b>	<b>HEALTH ECONOMICS .....</b>	<b>27</b>
	<b>REFERENCES.....</b>	<b>28</b>
	<b>APPROVALS.....</b>	<b>29</b>

## AMENDMENT HISTORY

### SLEDAI 2K SRI-4:

Note that the definition of the primary endpoint, SRI-4 response, has been updated from the Week 24 SAP. The following changes were adopted:

Week 24 SRI-4 Response: Defined as  $\geq$  4-point reduction from baseline in SLEDAI-2K score, **no new domain scores in either BILAG A or BILAG B** and no worsening from baseline in the physician's global assessment of disease activity.<sup>2</sup>

Week 56 SRI-4 Response: Defined as  $\geq$  4-point reduction from baseline in SLEDAI-2K score, **no new BILAG A and no more than 1 new BILAG B domain scores**, and no worsening from baseline in the physician's global assessment of disease activity.<sup>2</sup>

All analyses at the Week 56 DBL are based upon the Week 56 SRI-4 definition.

### Treatment Failure Rules:

Note that treatment failure rules were updated from the Week 24 SAP. The following changes were adopted:

The Week 24 SAP did not allow for replacement immunomodulator therapy, similar to the allowance for replace of immunomodulators (i.e. dose equivalent replacement). The Week 56 SAP has been updated to allow a replacement immunomodulator as long as the dose equivalent is the same or less than the current immunomodulator.

## ABBREVIATIONS

ACE	angiotensin-converting enzyme
AE	adverse event
ALT	alanine aminotransferase
ANCOVA	Analysis of Covariance
APAC	Asia Pacific
ARB	Angiotensin II receptor blocker
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
AUC	area under the curve
AZA/6 MP	Azathioprine / 6-mercaptopurine
NSAIDS	nonsteroidal anti-inflammatory drugs
BICLA	BILAG-based Combined Lupus Assessment
BILAG	The British Isles Lupus Assessment Group
BMI	body mass index
BP	blood pressure
CI	confidence interval
CLASI	Cutaneous Lupus Erythematosus Disease Area and Severity Index
CRF	case report form
DBL	database lock
DMC	Data Monitoring Committee
DPS	Data Presentation Specifications
eCRF	electronic case report form
EEU	Eastern European Union
FDA	Food and Drug Administration
FSS	Fatigue Severity Scale
HR	heart rate
IA	Interim Analysis

---

IAC	Interim Analysis Committee
ICF	informed consent form
IQ	interquartile
IV	intravenous
IWRS	interactive web response system
LA	Latin America
LN	Lupus Nephritis
LOCF	last observation carried forward
mITT	modified intent to treat
MMF	mycophenolate mofetil
MPA	mycophenolic acid
MTX	methotrexate
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
NSAID	non-steroidal anti-inflammatory drug
P/C	protein/creatinine ratio
PBO	placebo
PD	pharmacodynamics
PGA	Physicians Global Assessment
RBC	red blood cells
S2K RI-50	SLEDAI-2K Responder Index
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SC	subcutaneous
SD	standard deviation
SF-36	Short-form 36 questionnaire
SLE	Systemic lupus erythematosus
SLEDAI-2K	Systemic Lupus Erythematosus Disease Activity Index 2000
SLICC	Systemic Lupus International Collaborating Clinics
SRI	Systemic Lupus Erythematosus Responder Index
TEAE	treatment emergent adverse event
US	Unites States
VAS	Visual Analogue Scale
WBC	white blood cells
WEU	Western European Union
WHO	World Health Organization

## 1. INTRODUCTION

This Statistical Analysis Plan (SAP) contains definitions of analysis sets, derived variables, and statistical methods for the analysis of efficacy, safety, tolerability, pharmacokinetics (PK), pharmacodynamics (PD) and immunogenicity of ustekinumab. This Week 56 SAP is specific to the Week 56 database lock (DBL) for the CNTO1275SLE2001 study.

The primary DBL for the CNTO1275SLE2001 study occurred at Week 24 and the details of the Week 24 analyses were included in a separate SAP. The analyses for the Week 56 DBL are to evaluate safety and durability of effect through Week 56 (i.e. the end of the study).

### 1.1. Trial Objectives

#### Primary Objectives

The primary objectives are:

- To evaluate the efficacy of ustekinumab in the reduction of signs and symptoms of systemic lupus erythematosus (SLE) as measured by the SLE Responder Index (SRI-4) response rate at Week 24.

#### Secondary Objectives

The secondary objectives are to evaluate:

- The safety and tolerability of ustekinumab in SLE.
- The effect of ustekinumab administration on health-related quality of life in subjects with SLE.
- Pharmacokinetics and immunogenicity of ustekinumab in subjects with active SLE.
- The effects of ustekinumab on cutaneous manifestations of SLE.

#### Exploratory Objectives

The exploratory objectives are to evaluate:

- Biomarkers related to lupus disease (genetic, systemic, and skin-related).
- Additional composite clinical endpoints or methods of calculation of clinical response with potential for greater sensitivity to improvement and/or worsening of SLE.

### 1.2. Trial Design

CNTO1275SLE2001 is a Phase 2a, proof-of-concept, multicenter, randomized, double-blind, placebo-controlled study of the efficacy and safety of ustekinumab compared with placebo in subjects with active SLE. Enrolled subjects are expected to have active disease (SLICC classification criteria, SLEDAI score  $\geq 6$ , ANA and/or anti-dsDNA positivity as well as 1 BILAG A and/or 2 BILAG B scores) despite conventional treatment with standard lupus therapy (corticosteroids, anti-malarial drugs, non-steroidal anti-inflammatory drugs (NSAIDs), and/or immunosuppressive drugs).

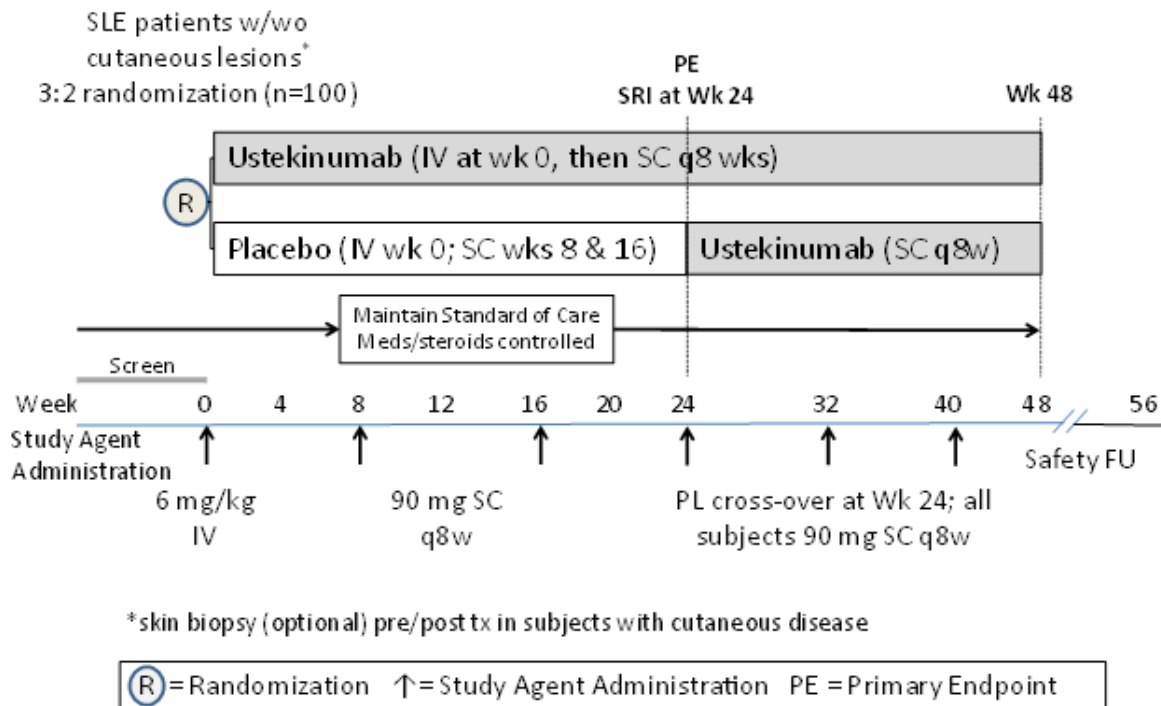
Approximately 100 subjects will be randomly assigned in a 3:2 ratio to receive either intravenous (IV) ustekinumab 6 mg/kg or IV placebo initially at Week 0 followed by continued treatment with subcutaneous (SC) ustekinumab 90 mg or placebo at Weeks 8 and 16. Subjects with cutaneous disease who sign a separate informed consent form (ICF) will be enrolled in the cutaneous lupus sub-study for additional assessments including photographs of skin lesions (optional consent) and/or collection of skin biopsies (optional consent) to be performed prior to study agent administration at Week 0 and at Week 24. There will not be any restrictions on the number of subjects with cutaneous disease who can enroll into either the main study or the cutaneous lupus sub study. Dynamic randomization will be utilized in this study to balance as best as possible the distribution of subjects to achieve the randomization ratio at both the study level and within each stratification factor: skin biopsy, presence of lupus nephritis, baseline SLE use and SLEDAI score (combined factor), site/region and race.

At Week 24, subjects receiving placebo will crossover and all subjects will then receive SC ustekinumab 90 mg at Weeks 24, 32, and 40 in a blinded fashion followed by 16 weeks of safety follow-up through Week 56 (i.e. more than 5 half-lives) after last SC administration.

Interim efficacy and safety analyses (IA) were conducted when approximately 1/3 and 2/3 of subjects reached Week 24. In the first IA, safety and an interim assessment of feasibility were assessed for the purpose of advising the study sponsor whether marked efficacy was observed, thus providing an early indicator to advance planning for a Phase 3 clinical trial in this indication. In the second IA, safety and evidence of marked efficacy as well as futility were analyzed. Database locks (DBLs) were to occur at Weeks 24 and 56.

A diagram of the study design is provided in Figure 1.

Figure 1: Study Design Schema



### 1.3. Statistical Hypotheses for Trial Objectives

The hypothesis is that ustekinumab is significantly superior to placebo as measured by the Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K) Responder Index (SRI-4) composite measure at Week 24.

There is no formal statistical hypothesis testing for the Week 56 DBL. The data from the Week 56 DBL are supportive to Week 24 and will evaluate longer term safety and durability of effect.

### 1.4. Sample Size Justification

The overall study sample size calculation was based upon the proportion of SRI responders at Week 24. To have approximately 80% power to detect a clinically significant odds ratio in the group proportions of 2.79, corresponding to a clinically significant increase in responders of 25% (35% placebo and 60% active) using a 2-sided Z-test at a 0.1 level of significance with pooled variance, 60 subjects with active treatment and 40 subjects with placebo are required. The assumption of a 35% responder rate for placebo is based upon a previous study in which a similar SLE population was treated.<sup>1</sup>

### 1.5. Randomization and Blinding

#### Randomization

The study employed an adaptive design using continuous enrollment. Dynamic randomization was used to randomize subjects to treatment group stratified by the following factors:



- collection of cutaneous skin biopsy (Y/N)
- presence of Lupus Nephritis (Y/N)
- composite of baseline SLE medications and SLEDAI score (high medications and SLEDAI  $\geq$  10, high medications and SLEDAI  $<$  10, medium medications and SLEDAI  $\geq$  10, medium medications and SLEDAI  $<$  10)
  - High medications defined as  $\geq$  15 mg/wk MTX, or  $\geq$  1.5 mg/kg/day AZA/6-MP, or  $\geq$  1.5 g/day MMF/MPA, and/or  $\geq$  15mg/day prednisolone or equivalent
  - Medium medications defined as  $<$  15 mg/wk MTX, or  $<$  1.5 mg/kg/day AZA/6-MP, or  $<$  1.5 g/day MMF/MPA, and/or  $<$  15 mg/day prednisolone or equivalent
- site/region
- race (white, black, not white or black)

Dynamic randomization was utilized to balance these features as best as possible and is described in detail in the Interactive Web Response System (IWRS) randomization specification documents.

Approximately 100 subjects were to be randomly assigned in a 3:2 ratio to receive either IV ustekinumab 6 mg/kg or IV placebo initially at Week 0 followed by continued treatment with SC ustekinumab 90 mg or placebo at Weeks 8 and 16. The interactive web response system (IWRS) assigned a unique treatment code based upon the best possible balance of these stratification factors, which dictated the treatment assignment and matching study drug kit for the subject. The requestor used his or her own user identification and personal identification number when contacting the IWRS, and then gave the relevant subject details to uniquely identify the subject.

### Blinding

Blinded treatment was used to reduce potential bias during data collection and evaluation of clinical endpoints as well as adverse events. Subjects received either ustekinumab (6 mg/kg IV followed by 90 mg SC) or its matching placebo through week 16 dosing, then all subjects received active treatment from week 24 through week 40.

The investigator was not provided with randomization codes. The codes were maintained within the IWRS, which has the functionality to allow the investigator to break the blind for an individual subject.

Data that may potentially unblind the treatment assignment (i.e., study agent serum concentrations, study agent preparation/accountability data, treatment allocation, biomarker or other specific laboratory data) was handled with special care to ensure that the integrity of the blind was maintained and the potential for bias is minimized. This can include making special provisions, such as segregating the data in question from view by the investigators, clinical team, or others as appropriate until the time of end of study database lock (DBL) and unblinding.

Under normal circumstances and unless otherwise stated, the original treatment blind should not be broken until all subjects have completed the study at Week 56 or terminated study participation and the database is finalized. Otherwise, the blind should be broken only if specific emergency treatment/course of action would be dictated by knowing the treatment status of the

subject. In such cases, the investigator may in an emergency determine the identity of the treatment by contacting the IWRS. It is recommended that the investigator contact the sponsor or its designee if possible to discuss the particular situation, before breaking the blind. Telephone contact with the sponsor or its designee will be available 24 hours per day, 7 days per week. In the event the blind is broken, the sponsor must be informed as soon as possible. The date and reason for the unblinding must be documented by the IWRS, and/or in the appropriate section of the electronic case report form (eCRF). The documentation received from the IWRS indicating the code break must be retained with the subject's source documents in a secure manner.

The Data Monitoring Committee (DMC) and Statistical Support Group (SSG) may be unblinded to subject treatment groups. The contents of the unblinded data to which the DMC and SSG have access should not be divulged, in any way, to members of the study team or to any members of the Sponsor Committee unless specifically requested by the Sponsor Committee Chairperson until the study has completed.

Subjects who have had their treatment assignment unblinded should continue to return for scheduled evaluations.

In general, randomization codes will be disclosed fully only if the study is completed and the clinical database is closed. For the specified interim analyses, the randomization codes and the translation of randomization codes into treatment and control groups will be disclosed to those authorized as detailed in the DMC Charter. The Sponsor will be blinded through the Week 24 evaluation and until the database is cleaned and finalized for Week 24 planned analyses. The clinical site, subjects, investigators, and site personnel will remain blinded through the end of the study until Week 56 data are finalized.

## **2. GENERAL ANALYSIS DEFINITIONS**

### **2.1. Visit Windows**

Nominal visits will be used for all by visit analyses. The study visits scheduled after randomization should occur at the times delineated in the Time and Events Schedule.

### **2.2. Pooling Algorithm for Analysis Centers**

Data from all investigational centers/sites will be pooled for analyses since analyses will not be adjusted for investigator site.

## 2.3. Analysis Sets

### 2.3.1. Efficacy Analysis Set(s)

#### 2.3.1.1. Primary Efficacy Analysis Set

The primary efficacy analysis set will be based upon the Full Analysis Set (FAS) defined as randomized subjects who receive at least 1 dose (partial or complete, IV or SC) of ustekinumab or placebo. Subjects will be analyzed according to their randomized treatment group.

#### 2.3.2. Safety Analysis Set

The safety analysis set is defined as the set of all randomized subjects who have received at least 1 dose (partial or complete, IV or SC) of ustekinumab or placebo. Subjects will be analyzed according to the actual treatment received. Randomized subjects confirmed not to have taken any study medication will not be included in the safety analysis set.

### 2.3.3. Pharmacokinetics Analysis Set(s)

The PK analysis set is defined as all subjects who received at least 1 complete dose of ustekinumab and have at least one post initial dose sample collection with measurable concentrations. Subjects will be analyzed according to the actual treatment received.

### 2.3.4. Immunogenicity Analysis Set

The immunogenicity analysis set is defined as all subjects who received at least 1 dose (partial or complete) of ustekinumab and have at least one post initial dose sample collection. Subjects will be analyzed according to the actual treatment received.

### 2.3.5. Pharmacodynamics Analysis Set

The PD analysis set is defined as all subjects who received at least one dose (complete or partial, IV or SC) of study agent. Subjects will be analyzed according to the actual treatment received.

## 2.4. Definition of Subgroups

To evaluate the consistency of the proportion of subjects with SRI-4 composite response at Week 48, subgroup analyses may be performed on the following subgroups of baseline characteristics:

1. Region (APAC, LA, EEU, WEU, US)
2. Race
3. Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
4. SLEDAI-2K Score (< Median, ≥ Median)
5. PGA Score (< Median, ≥ Median)
6. Renal disease presence (Y, N)
7. P/C Ratio (< Median, ≥ Median)
8. SLE Medications (High dose, Medium dose)
9. Age (18-25, 26-55, Over 55)
10. Body Weight (≥ 35 and ≤ 55, > 55 and ≤ 85, > 85)
11. Baseline OCS Use

### **3. INTERIM ANALYSIS AND DATA MONITORING COMMITTEE REVIEW**

Details of the Data Monitoring Committee safety reviews and interim analysis assessments are included in a separate DMC SAP. The DMC SAP contains definitions of analysis sets, derived variables, scope and objects of the DMC, data-cutoff points, statistical methods and output to be produced for the safety reviews and interim analyses of efficacy as required by the DMC Charter. Additionally, a dedicated DMC Data Presentation Specifications (DPS) document details output and programming specifications.

### **4. SUBJECT INFORMATION**

#### **4.1. Demographics and Baseline Characteristics**

Demographics and baseline characteristics will be summarized based upon the randomized analysis set. Descriptive statistics (mean, standard deviation, median, and range) will be provided for the subjects' age and baseline factors including SLEDAI-2K score, PGA score, weight and height, Body Mass Index (BMI), Fatigue Severity Score (FSS), SF-36 physical and mental component summary, number of joints with pain and inflammation, number of joints with pain, number of joints with inflammation, CLASI damage and activity score, disease duration in months, BILAG index total, antinuclear antibodies, anti-double stranded DNA and SLE medications. Subjects' sex, race, ethnicity, region, baseline BILAG categories, baseline P/C ratio, number of subjects with serum creatinine above normal, number of subjects with Lupus Nephritis at baseline, Smith antibody presence and number of subjects with low complement will be summarized as frequency distributions by treatment group. Prior and concomitant medications used at baseline will be listed.

Age will be calculated as  $(\text{date of informed consent} - \text{date of birth} + 1) / 365.25$  in units of years.

BMI will be calculated as  $\text{weight}/\text{height}^2$  with units of  $\text{kg}/\text{m}^2$ .

#### **4.2. Disposition Information**

Disposition information will be summarized by treatment group in terms of numbers of subjects randomized, treated, and completing scheduled visits. Disposition information will be based upon the randomized analysis set. The number of subjects who enrolled in the study extension will be summarized. Additionally, the number of subjects who discontinued study agent and the reason for discontinuation as well as those who terminated the study and reason for termination will be summarized.

#### **4.3. Treatment Compliance**

Study agent will be administered as an IV (initial dose) and SC injection by an authorized and qualified staff and the details of each administration will be recorded in the electronic case report form (eCRF) including date, start and stop times for infusion and start time for injection. Compliance with the treatment assignments will be controlled by the study site personnel.

#### **4.4. Extent of Exposure**

The number of subjects exposed to study agent, the number of administrations of study agent, and cumulative dose of study agent will be summarized by actual treatment group.

#### **4.5. Protocol Deviations**

Major protocol deviations will be tabulated by treatment group as defined in the following categories:

- Subjects who entered but did not satisfy entry criteria
- Subjects who received the wrong treatment of incorrect dose
- Subjects who developed withdrawal criteria but were not withdrawn
- Subjects who received a disallowed concomitant medication
- Other deviations (e.g. study procedures out of window visit, missed visits for reasons other than AEs, unblinding error, actions deemed as major deviations by the study team)

#### **4.6. Prior and Concomitant Medications**

Background medication use for active SLE will be summarized by treatment group. Subjects initiating new or altering dose (increasing or decreasing) of medication for active SLE during the study will be summarized over time. Subjects with increases in medication will be listed.

Prior and concomitant medications will be coded according to the World Health Organization (WHO) encoding dictionary, and the number and percent of subjects receiving each medication in each treatment group will be summarized by Level 3 Anatomical Therapeutic Chemical Classification (ATC) code (first four numbers in the code).

### **5. EFFICACY**

#### **5.1. Analysis Specifications**

##### **5.1.1. Level of Significance**

Unless otherwise stated, statistical significance is set to 2 sided  $\alpha$ -level of 0.10 as per protocol. Where appropriate, 90% confidence intervals will be provided.

Two interim analyses on the primary endpoint were performed when 1/3 (approximately 33 subjects) and 2/3 (approximately 66 subjects) of the subjects reached their Week 24 visit. Per protocol, the study was not to be stopped for efficacy at interims.

At the Week 24 primary DBL, the primary and major secondary endpoints were tested sequentially according to the order specified in the protocol. If previous endpoint proved not statistically significant, subsequent endpoint(s) were not to be tested and nominal p-values were provided.

### 5.1.2. Data Handling Rules

**Missing Data Rules:** Unless otherwise specified, in situations with the presence of missing data (e.g. lost to follow-up, missed study visit) the following rules will apply:

Missing baseline value: No imputation will be performed for missing baseline values.

Subjects missing values for any subgroup parameters (defined in Section 2.4) at baseline will be excluded from the subgroup analysis.

Continuous endpoints:

- No imputation will be performed for missing post-baseline continuous values. Missing continuous endpoints tested at Week 24 will be adjusted for in the statistical model (i.e. Mixed effect Model Repeated Measures)

Binary endpoints:

- For primary and major secondary composite binary endpoints (SRI, BICLA), non-responder imputation will be performed for missing post-baseline response values.
- For any other binary endpoints tested, missing data will be imputed using multiple imputation methods (MI). If the missing data is not monotone, a Markov chain Monte Carlo (MCMC) method will be used first to effectively make the missing pattern monotone. For monotone missing data, a serial logistic regression method will be used. The imputation models will include treatment group and baseline SLEDAI as a covariate.

### 5.1.3. General Analysis Methods

Descriptive statistics (e.g. mean, median, standard deviation, interquartile range, minimum, maximum) will be used to summarize continuous variables. Counts and percentages will be used to summarize categorical variables. Recurrence data will be summarized by mean as well as counts and percentages. Median will be reported for time to event variables. Graphical displays will also be used to present the data.

### 5.1.4. Treatment Failure Criteria

**Treatment Failure Criteria:** Subjects who have 1 or more of the following events will be defined as treatment failures:

Non-biologic Immunomodulators (Mycophenolate mofetil [MMF], Mycophenolic acid [MPA], Azathioprine [AZA], 6-mercaptopurine [6-MP], Methotrexate [MTX])

- Between the Week 12 and the Week 24 visit either; (a) the dose of an immunomodulator is higher than at baseline, or (b) a new immunomodulator (not present at randomization) has been added to the existing treatment regimen and this immunomodulator does NOT replace a current immunomodulator (i.e. dose equivalent is equal or less than current

dose) of similar mechanism of action (e.g. MMF substitution for MPA or vice versa, AZA substitution for 6-MP or vice versa)

- There was addition of a new immunomodulator to the existing treatment regimen before Week 12 and the subject still was receiving that immunomodulator after Week 12 and this immunomodulator does NOT replace a current immunomodulator (i.e. dose equivalent is equal or less than current dose) of similar mechanism of action (e.g. MMF substitution for MPA or vice versa, AZA substitution for 6-MP or vice versa)

#### Antimalarials

- Between the Week 12 and the Week 24 visit either; (a) the dose of an antimalarial is higher than at baseline, or (b) a new antimalarial (not present at randomization) has been added to the existing treatment regimen and this antimalarial does NOT replace a current antimalarial (i.e. dose equivalent is equal or less than current dose)
- There was addition of a new antimalarial to the existing treatment regimen before Week 12 and the subject still was receiving that antimalarial after Week 12 and this antimalarial does NOT replace a current antimalarial (i.e. dose equivalent is equal or less than current dose)

#### Glucocorticoids

- Between the Week 12 and Week 24 visit, initiation of treatment with (a) disallowed dose or (b) disallowed use of oral, IV or IM or other type of corticosteroid, for SLE or (c) increased the dose of oral corticosteroids for SLE above baseline.

#### Angiotensin II Receptor Blocker or Angiotensin-Converting Enzyme

- Initiation of new angiotensin II receptor blocker (ARB) or angiotensin-converting enzyme (ACE) inhibitor therapy after the first dose of study agent and before Week 24 and this medication does NOT replace a current ARB or ACE respectively.

#### Discontinuation

- Discontinuation of study agent due to lack of efficacy or for an AE of worsening of SLE prior to Week 24

Treatment failure rules will NOT be applied after a subject discontinues study agent (i.e. after the date of discontinuation of study agent.)

### **5.1.5. Estimands**

The **Composite Estimand** is defined by the following:

- Population: Subjects in the FAS with active SLE despite receiving one or more standard-of-care treatments.
- Variable: Proportion of SRI-4 composite response
- Intercurrent event: The following intercurrent events to be taken into account:



1. If a subject takes rescue medication or discontinues treatment (see treatment failure rules for details), assume non-response from the point of treatment failure onward
  2. If a subject is lost to follow-up, assume non-response from the point of loss to follow-up onward
- Population-level summary: Proportion of subjects achieving SRI-4 composite response

The **Hypothetical Estimand** is defined by the following:

- Population: Subjects in the FAS with active SLE despite receiving one or more standard-of-care treatments.
- Variable: Proportion of SRI-4 composite response
- Intercurrent event: The following intercurrent events to be taken into account:
  1. If a subject takes rescue medication or discontinues treatment (see treatment failure rules for details), do not assume non-response and instead assume missing values
  2. If a subject is lost to follow-up, do not assume non-response and instead assume missing
- Population-level summary: Proportion of subjects achieving SRI-4 composite response

The **Treatment Policy (deFacto) Estimand** is defined by the following:

- Population: Subjects in the FAS with active SLE despite receiving one or more standard-of-care treatments.
- Variable: Proportion of SRI-4 composite response
- Intercurrent event: The occurrence of an intercurrent event is irrelevant. All observed values will be used regardless of occurrence of an intercurrent event.
  1. All observed values will be used regardless of occurrence of an intercurrent event.
  2. If a subject is lost to follow-up, no imputation is performed
- Population-level summary: Proportion of subjects achieving SRI-4 composite response.

## 5.2. Efficacy Variable(s)

- [REDACTED]

### SLE Disease Activity:

- The proportion of subjects with responses in SRI-4, SRI-5, SRI-6, SRI-7, SRI-8, S2K RI-50 response and BICLA over time



- [REDACTED]
- [REDACTED]
- The absolute change from baseline in SLEDAI-2K, S2K RI-50, PGA over time
- The percent change from baseline in serological activity (eg, ANA, anti-dsDNA, other autoantibodies, C3, C4) or SLEDAI-2K feature measurements over time
- [REDACTED]
- [REDACTED]
- The change from baseline in active, tender, and swollen joint counts in subjects with joint disease over time
- [REDACTED]
- [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

#### Medications:

- The proportion of subjects with meaningful changes in selected SLE medications over time

#### **5.2.1. Definitions**

SLEDAI 2K SRI-4 (Modified): This version was used in the Week 24 DBL and will NOT be applied to Week 56 SRI analyses. Defined as  $\geq 4$ -point reduction from baseline in SLEDAI-2K score, **no new domain scores in either BILAG A or BILAG B** and no worsening from baseline in the physician's global assessment of disease activity. (see Week 24 SAP)

**SLEDAI 2K SRI-4:** This version will be used in all Week 56 SRI analyses. Defined as  $\geq 4$ -point reduction from baseline in SLEDAI-2K score, **no new BILAG A and no more than 1 new BILAG B domain scores**, and no worsening from baseline in the physician's global assessment of disease activity.<sup>2</sup>

Composite response is defined as subjects who achieve SRI-4 response and do not meet the treatment failure criteria prior to the analysis time point (e.g. prior to Week 24 for the primary endpoint)

SLEDAI-2K SRI-5, SRI-6, SRI-7 and SRI-8 are similarly defined as above with response requiring a  $\geq 5$ -point,  $\geq 6$ -point,  $\geq 7$ -point or  $\geq 8$ -point reduction from baseline in SLEDAI-2K score respectively. Subjects evaluable for SRI-5 through SRI-8 response will include subjects with a baseline SLEDAI score of at least 5 through 8 respectively.

**SLEDAI-2K:** The SLE disease activity index 2000 (SLEDAI-2K is an established, validated SLE activity index. It is based on the presence of 24 features in 9 organ systems and measures disease activity in SLE patients in the previous 30 days. It is weighted according to the feature. Features are scored by the assessing physician if present within the last 30 days with more severe features having higher scores, and then simply added to determine the total SLEDAI 2K score, which ranges from 0 to 105.<sup>3</sup> The baseline measurement for the SLEDAI is defined as the closest measurement taken prior to the initiation of the Week 0 administration, unless otherwise specified.

**BILAG:** The BILAG<sup>4,5</sup> index scores subjects based on the need for alterations or intensification of therapy. The assessing physician will evaluate 97 items divided into 9 organ system domains. The physician considers presence of each item in the past 4 weeks and answers 0=not present, 1=improving, 2=same, 3=worse, or 4=new. Each organ/system domain is classified as BILAG A, B, C, D, or E based upon criteria specific to the domain. The baseline measurement for the BILAG is defined as the worst (highest) measurement at any time prior to the initiation of the Week 0 administration, unless otherwise specified. The baseline measurement for the BILAG is defined as the closest measurement taken during the study prior to the initiation of the Week 0 administration, unless otherwise specified.

**PGA:** The Physician's Global Assessment of Disease Activity<sup>6</sup> independent of subjects' assessment is recorded on a visual analogue scale (VAS; 0 to 10 cm). The scale for the assessments range for 'no Lupus activity' (0) to 'extremely active Lupus' (10). No worsening in PGA is defined as less than a 10 mm increase from baseline on 100 mm VAS. The baseline measurement for the PGA is defined as the closest measurement taken prior to the initiation of the Week 0 administration, unless otherwise specified.

**BICLA:** The BILAG-based Combined Lupus Assessment (BICLA) requires patients to meet response criteria across 3 assessment tools: (1) the BILAG-2004 index (2) the SLEDAI index and (3) a PGA. Patients are identified as responders or non-responders based upon the following requirements:<sup>7</sup>

Table 2: BICLA Response Criteria

Requirements for BICLA Response (BILAG + SLEDAI-2K + PGA + Treatment Failure criteria must be met)	
BILAG	BILAG improvement classified as: <ul style="list-style-type: none"> <li>• All BILAG A scores at baseline improved to either BILAG B, C or D</li> <li>• All BILAG B scores at baseline improved to either BILAG C or D</li> <li>• No worsening in disease activity defined as no new BILAG A scores and <math>\leq 1</math> new BILAG B score</li> </ul>
SLEDAI-2K	No worsening of total SLEDAI-2K from baseline (change $\leq 0$ )
PGA	No significant deterioration ( $< 1$ cm increase) in 10 cm visual analogue PGA
Treatment Failure	No treatment failure (see SAP for definition of treatment failure)

SLEDAI-2K RI-50 (S2K RI-50): The SLEDAI-2K has been adapted and developed into the SLEDAI-2K Responder Index (S2K RI-50)<sup>15</sup>, a measure that can document partial improvement in the 24 disease features between SLEDAI-2K assessments.<sup>16</sup> A threshold of 50% improvement was judged to reflect clinically significant improvement and is scored as half the weight for the feature. “When a descriptor is recorded as present at the initial visit, 1 of 3 situations can follow: (1) the descriptor achieves complete remission at follow-up, in which case the score would be “0”; (2) the descriptor does not achieve a minimum of 50% improvement at follow-up, in which case the score would be identical to its corresponding SLEDAI-2K value; or (3) the descriptor improves by  $\geq 50\%$  (according to the S2K RI-50 definition) but has not achieved complete remission, in which case the score is evaluated as one-half the score that would be assigned for SLEDAI-2K.”<sup>17</sup> The S2K RI-50 score is the sum of the 24 scored items, which ranges from 0 to 105. The S2K RI-50 post-baseline visits will be compared to baseline.

[REDACTED]

- [REDACTED]

[REDACTED]

[REDACTED]

SLE SRI-4, SRI-5, SRI-6, SRI-7, SRI-8, S2K RI-50, BICLA, SLEDAI-2K, BILAG and PGA: These efficacy endpoints are defined in sections 5.2.1 and 5.3.1 above

Patient's Global Assessment (PtGA) of Disease Activity and Pain: The PtGA of Disease Activity will be recorded by the patient on a visual analogue scale (VAS; 0 to 10 cm). The scale for the assessments range from 'very well' (0) to 'very poor' (10). The PtGA of Pain is used to assess the patient reported pain intensity. The patient's will be asked their average pain during the past week on a visual analogue scale (VAS; 0 to 10 cm). The anchors of the instrument include 0 to represent 'no pain' and 10 to represent 'the worst possible pain'.

CLASI: The Cutaneous Lupus Erythematosus Disease Area and Severity Index (CLASI) consists of 2 scores; the first summarizes the activity of the disease while the second is a measure of the damage done by the disease (CLASI Activity and CLASI Damage scores). Activity is scored on the basis of erythema, scale/hyperkeratosis, mucous membrane involvement, acute hair loss and non-scarring alopecia. Damage is scored in terms of dyspigmentation and scarring, including scarring alopecia. The scores are calculated by simple addition based on the extent of the symptoms.<sup>8</sup> Higher activity and damage scores indicate worse disease activity. CLASI scores range from 0-70 for activity and 0-56 for damage.

FSS: The Fatigue Severity Scale (FSS) is a 9-item questionnaire designed to assess the severity of fatigue and its impact on daily living using 7 response options (1=Completely Disagree, 7=Completely Agree) during a recall period of the past week. It can be completed within 5 minutes by the subject. Scores above 36 of the total possible score of 63 reflect increasing severity of fatigue. The scale was developed for use in SLE.<sup>9</sup> The scores on the scale correlate with patient reported pain, sleep, depression, and with each subscale of the SF-36. The FSS has shown a high internal consistency, and differentiates patients from controls in studies with SLE subjects. The instrument was translated from the original English version and is available in several languages.

Clinically meaningful change in FSS defined as a change (decrease) that is  $\geq \frac{1}{2}$  the baseline standard deviation of the pooled treatment groups.

SF-36: The RAND short form (SF)-36 questionnaire is a self-administered multi-domain scale with 36 items. Eight health domains cover a range of functioning:

- Limitations in physical function
- Limitations in usual role activities
- Bodily pain
- General mental health (psychological distress and well-being)
- Vitality (energy and fatigue)
- Limitations in social functioning due to physical or mental health problems
- Limitations in usual role activities due to personal or emotional problems
- General health perception

The subscales are scored from 0 to 100. The scoring yields a Physical Component Summary (PCS) score and a Mental Component Summary (MCS) score, a total score, and subscale scores. Higher scores represent better outcomes. It is appropriate for persons over the age of 14 and may

be completed in 5 to 10 minutes. Translations are available in most languages; the instrument has undergone extensive linguistic and cultural validation. Version 2 acute will be used in the study.

The concepts measured by the SF-36 are not specific to any age, disease, or treatment group, allowing comparison of relative burden of different diseases and the benefit of different treatments.<sup>10</sup> A change of 3 points in any of the subscales or 5 points for the component score is associated with clinically meaningful change.<sup>11,12,13</sup> The SF-36 has been used extensively in clinical trials providing evidence of psychometric properties. Reliability estimates for physical and mental component summary scores exceeded 0.90 in early studies<sup>14</sup> and have been further confirmed in later studies. Construct validation was established through comparison to several other generic health surveys.



Minimally clinical important difference in SF-36 defined as at least a 5-point improvement (increase) from baseline in SF-36 PCS or MCS score.

### 5.2.2. Analysis Methods

For endpoints after Week 24:

SRI-4 response at Week 24 and Week 48 (8 Week Follow Up) in subjects that were randomized to placebo and crossed-over to receive ustekinumab at Weeks 24, 32 and 40 will be compared with appropriate methods (e.g. Kappa statistic or McNemar's test of paired proportions).

In general, no statistical comparisons will be performed on endpoints occurring after the Week 24 DBL. Descriptive statistics (e.g. mean, median, standard deviation, interquartile range, minimum and maximum) will be used to summarize continuous variables. Counts and percentages will be used to summarize categorical variables. Recurrence data will be summarized by mean as well as counts and percentages. Median will be reported for time to event variables. Graphical displays will also be used to present the data.

In general, analysis on SLEDAI-2K SRI and BICLA endpoints will be based on the Composite Estimand: if subjects meet treatment failure criteria or are lost to follow-up, they are assumed to

be non-responders. Sensitivity analyses may be performed using the Treatment Efficacy and DeFacto estimands.

In general, all other endpoints will be based upon the Treatment Efficacy estimand: if subjects meet treatment failure rules, their observations after meeting treatment failure criteria will be set to missing. Sensitivity analyses may be performed using the DeFacto estimand.

#### Subgroup Analyses:

Subgroup analyses of the SRI-4 response may be performed on the following:

- Region (APAC, LA, EEU, WEU, US)
- Race
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- SLEDAI-2K Score (< Median, ≥ Median)
- PGA Score (< Median, ≥ Median)
- Renal disease presence (Y, N)
- P/C Ratio (< Median, ≥ Median)
- SLE Medications (High dose, Medium dose)
- Age (18-25, 26-55, Over 55)
- Body Weight (≥ 35 and ≤ 55, > 55 and ≤ 85, > 85)
- Baseline OCS Use

For subgroup analyses, the proportion of subjects who achieved composite SRI-4 response at Week 48 (8 Week Follow Up) will be presented by treatment sequence.

## **6. SAFETY**

Unless otherwise stated, all safety analyses will be based upon the safety analysis set as described in Section 2.3.3 above. No formal statistical comparison is planned.

### **6.1. Adverse Events**

Adverse Event (AE) data will be collected at every study visit from screening through the end of the study. Analyses of AEs will be performed on those events that are considered to be treatment emergent adverse events (TEAEs). A TEAE is defined as any event that occurs at or after the initial administration of study agent. AEs that occur prior to the administration of study agent but then increase in intensity after dosing are also considered to be treatment emergent.

Note that if the event occurs on the day of initial administration and either the event time or time of administration are missing, then the event will be assumed to be treatment emergent.

For the placebo group, AEs that occur before the first dose of ustekinumab (before crossing over) will be considered to have occurred while on placebo treatment. For AEs that occur in the placebo group during or after the first dose of ustekinumab (after crossing over) will be

considered to have occurred while on ustekinumab treatment. Summaries will clearly describe this separation (e.g. example heading below).

Placebo – Ustekinumab	
Placebo (Up to Week 24)	Ustekinumab (After Week 24 Crossover)

TEAEs will be tabulated by treatment group and Medical Dictionary for Regulatory Activities (MedDRA) System Organ Class (SOC) and be sorted in descending frequency of the Preferred Term (PT) within the SOC. The following TEAEs will be summarized by treatment group:

- All TEAEs
- Serious TEAEs
- TEAEs by severity (mild, moderate, severe)
- TEAEs reasonably related to study agent
- Treatment emergent infections
- Treatment emergent infusion/injection reactions

In addition, the following will be listed by subject:

- Serious TEAEs
- TEAEs leading to discontinuation of study agent
- Any deaths

A reasonably related AE is defined as any event with the relationship to study agent as ‘very likely’, ‘probable’, or ‘possible’ on the AE eCRF page or if the relationship to study agent is missing.

An infusion reaction is defined as an AE that occurs during or within 1 hour following the infusion of study agent AND the question on the AE eCRF page ‘Was this an infusion related reaction’ = ‘Yes’, with the exception of laboratory abnormalities. If the AE onset time is missing and the AE onset date is the same as the date of infusion, an AE will be considered as an infusion reaction only if the question on the AE eCRF page ‘Was this an infusion related reaction’ = ‘Yes’.

An injection site reaction is defined as an AE with the question on the AE eCRF page ‘Was this an injection site reaction’ = ‘Yes’.

An infection is defined as any AE that was characterized by the investigator as an infection on the eCRF.

Since safety should be assessed relative to exposure and follow-up, all AE summary tables will summarize the average weeks of follow-up and average exposure for each treatment group.

## 6.2. Clinical Laboratory Tests

Routine laboratory data from clinical chemistry, hematology, serology and urinalysis will be collected at screening and at Visits according to the time and events schedule in the study protocol. The laboratory data to be summarized are as follows:

**Clinical Blood Chemistry:** Chemistry panel including alanine aminotransferase (ALT), albumin, alkaline phosphatase, aspartate aminotransferase (AST), bicarbonate, calcium, chloride, creatinine, creatinine kinase, GFR, glucose, phosphate, potassium, protein/creatinine ration, sodium, total bilirubin, total protein and urea nitrogen. If total bilirubin is abnormally elevated, both direct and indirect bilirubin will be assessed.

**Hematology:** hemoglobin, hematocrit, platelet count, red blood cell (RBC) count, white blood cell (WBC) total and differential (basophils, eosinophils, lymphocytes, monocytes, neutrophils)

**Urinalysis:** Dipstick (bilirubin, blood, glucose, ketones, leukocytes, nitrite, pH, specific gravity, and urobilinogen). If dipstick is abnormal, flow cytometry will be used to measure sediment (bacteria, crystals, epithelial cells, granular casts, heme casts, RBC casts, WBC casts).

Additional lab data may be summarized as needed.

Descriptive statistics for each clinical laboratory analyte and for change from baseline at each scheduled post-baseline visit will be provided by treatment group. The proportion of subjects with post-baseline values by maximum toxicity grade for clinical laboratory tests will be summarized by treatment group. The National Cancer Institute's Common Terminology Criteria for Adverse Events (NCI-CTCAE; version 4.0) will be used. Subjects with toxicity grades  $\geq 2$  will be listed.

For those labs without NCI toxicity grades, the incidence of markedly abnormal laboratory values will be presented by treatment group. Additionally, the markedly abnormal laboratory values will be listed. Markedly abnormal criteria will be defined in the Data Presentation Specifications (DPS) document.

## 6.3. Vital Signs and Physical Examination Findings

Weight, height, blood pressure, temperature and heart rate will be assessed.

Descriptive statistics of changes from baseline will be summarized by visit and treatment group. Abnormal findings will be summarized by treatment group. In addition, a listing of subjects with abnormal vital signs will be presented. Vital sign parameters will be considered abnormal using the criteria below:



Vital Sign	Abnormally Low	Abnormally High
Systolic BP (mm Hg)	Absolute value < 90 mmHg and a decrease from baseline $\geq$ 20 mmHg	Absolute value > 180 mmHg and an increase from baseline $\geq$ 20 mmHg
Diastolic BP (mm Hg)	Absolute value < 50 mmHG and a decrease from baseline $\geq$ 15 mmHg	Absolute value > 105 and an increase from baseline $\geq$ 15 mmHg
Pulse Rate (beats/minute)	Absolute value < 50 bpm and a decrease from baseline $\geq$ 15 bpm	Absolute value > 120 bpm and an increase from baseline $\geq$ 15 bpm
Temperature (°C)	< 36°C	> 38°C

#### 6.4. Electrocardiogram

This section does not apply to this study.

### 7. PHARMACOKINETICS/IMMUNOGENICITY/PHARMACODYNAMICS

#### 7.1. Pharmacokinetics

PK evaluations will be based on PK analysis set (section 2.3.4). No imputation for missing concentration data will be performed.

PK concentrations will be summarized by treatment group among PK evaluable subjects. Descriptive statistics, including arithmetic mean, SD, coefficient of variation, median, minimum, and maximum will be calculated for ustekinumab concentrations at each sampling time.

Concentrations below the lowest quantifiable concentration in a sample will be treated as zero in the summary statistics. All subjects excluded from the analysis will be clearly documented in the study report. Once a subject meets one of the following dosing deviation criteria, the subject's data will be excluded from the by-visit data analyses from that point onwards:

- Discontinued ustekinumab administrations.
- Skipped a ustekinumab administration.
- Received an incomplete/ incorrect ustekinumab dose.
- Received an incorrect study agent.
- Received an additional ustekinumab dose.
- Received an administration outside of visit windows (For PK analyses, the visit and the study agent administration should occur within +/- 7 days of the scheduled visit day relative to Week 0).

The definition of treatment group at Week 24 analysis is as follows:

- **Ustekinumab:** Subjects who received ustekinumab from Week 0.

The definition of treatment group at Week 56 analysis is as follows:

- **Ustekinumab**: Subjects who received ustekinumab from Week 0.
- **Placebo → Ustekinumab**: Subjects who received placebo at Week 0 and then crossed-over at Week 24 to receive ustekinumab

PK data may be displayed graphically. Plots of mean (SD) or Median (IQ) serum concentrations of ustekinumab over time, using semi-logarithmic scales, may be provided.

The relationship between PK and antibody to ustekinumab status may be assessed.

If feasible, a population PK analysis using nonlinear mixed effects modeling may be used to characterize the disposition characteristics of ustekinumab in the current study. The influence of important variables such as body weight and antibodies to ustekinumab status on the population PK parameter estimates may be evaluated. Results of the population PK analysis will be presented in a separate technical report.

## 7.2. Immunogenicity (Antibodies to ustekinumab)

The incidence and titers of antibodies to ustekinumab will be summarized for all subjects who received at least one dose (complete or partial) of ustekinumab and had appropriate serum samples for antibody detection. The incidence of NABs to ustekinumab will be summarized for subjects who are positive for antibodies to ustekinumab and have samples evaluable for NABs. A listing of subjects positive for antibodies to ustekinumab at any time will be provided.

The definition of treatment group at Week 24 analysis is as follows:

- **Ustekinumab**: Subjects who received ustekinumab from Week 0.

The definition of treatment group at Week 56 analysis is as follows:

- **Ustekinumab**: Subjects who received ustekinumab from Week 0.
- **Placebo → Ustekinumab**: Subjects who received placebo at Week 0 and then crossed-over at Week 24 to receive ustekinumab

## 7.3. Pharmacodynamics

Anti-Smith, SS-A (Ro), SS-B (La) autoantibodies will be collected after Week 0 only in subjects who are positive for these autoantibodies at baseline. Autoantibody levels and proportion of subjects with positive autoantibody over time will be summarized for those subjects positive at baseline.

Changes in the concentration of individual pharmacodynamics markers from baseline to the selected post treatment time points will be summarized. These include but are not limited to ANA, C3 and C4 complement, IgA, IgG, IgM, and SAA.

**7.4. Pharmacokinetic/Pharmacodynamic Relationships**

If data permit, the relationships between serum ustekinumab concentration and efficacy or pharmacodynamic measures may be analyzed graphically.

**8. HEALTH ECONOMICS**

Not applicable

---

## REFERENCES

1. Van Vollenhoven RF, Petri MA, Cervera R, et al. Belimumab in the treatment of systemic lupus erythematosus: high disease activity predictors of response. *Ann Rheum Dis*. 2012;71(8):1343-1349.
2. Furie RA, Petri MA, Wallace DJ, et al. Novel evidence-based systemic lupus erythematosus responder index. *Arthritis & Rheumatism*. 2009;61(9):1143-1151.
3. Touma Z, Gladman DD, Urowitz MB. SLEDAI-2K for a 30 day window. *Lupus*. 2010a;19(1):49-51. Epub 2009 Nov 12.
4. Hay EM, Bacon PA, Gordon C, et al. The BILAG index: a reliable and valid instrument for measuring clinical disease activity in systemic lupus erythematosus. *Quart J Medicine*. 1993;86:447-458.
5. Isenberg DA, Rahman A, Allen E, et al. BILAG 2004. Development and initial validation of an updated version of the British Isles Lupus Assessment Group's disease activity index for patients with systemic lupus erythematosus. *Rheumatology*. 2005;44:902-906.
6. Felson DT, Anderson JJ, Boers M, et al. American College of Rheumatology preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum*. 1995;38(6):727-735.
7. Wallace DJ. *Lupus: The essential clinician's guide*. New York, NY: Oxford University Press, Inc; 2008.
8. Albrecht J, Taylor L, Berlin JA, et al. The CLASI (Cutaneous Lupus Erythematosus Disease Area and Severity Index): an outcome instrument for cutaneous lupus erythematosus. *J Invest Dermatol*. 2005;125(5):889-894.
9. Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD. The fatigue severity scale: application to patients with multiple sclerosis and systemic lupus erythematosus. *Arch Neurol*. 1989;46(10):1121-1123.
10. Ware JE Jr, Sherbourne CD. The MOS 36 item short form health survey (SF 36), I: conceptual framework and item selection. *Med Care*. 1992;30(6):473-483.
11. Samsa G, Edelman D, Rothman ML, et al. Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. *Pharmacoeconomics*. 1999;15(2):141-155.
12. Ware JE, Kosinski M, Keller SK. *SF-36 Physical and Mental Health Summary Scales: A User's Manual*. Boston MA. The Health Institute, 1994.
13. Ware JE. SF-36 Health Survey Update. *Spine*. 2000;25(24):3130-3139.
14. McHorney CA, Ware JE Jr, Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care*. 1994;32(1):40-66.
15. Touma Z; Urowitz M, Gladman D. SLEDAI-2K Responder Index-50 (SRI-50). *Lupus*. The 9th International Congress on SLE June 24-27 2010b, Vancouver, Canada. Abstract PO2.D.7.
16. Touma Z, Urowitz M, Ibañez D, Gladman D. SLEDAI-2K 10 days versus SLEDAI-2K 30 days in a cross-sectional and longitudinal evaluation. *Lupus*. The 9th International Congress on SLE June 24-27 2010c, Vancouver, Canada. Abstract PO2.D.6.
17. Touma Z, Gladman DD, Ibanez D, Urowitz MB. Development and initial validation of the systemic lupus erythematosus disease activity index 2000 responder index 50. *J Rheumatol*. 2011; 38:2; doi:10.3899/jrheum.100724.
18. Franklyn K, Lau CS, Navarra SV et al. Definition and initial validation of a Lupus Low Disease Activity State (LLDAS). *Annals of the Rheumatic Diseases* 2016;75(9):1615-1621.
19. Golder V, Kandane-Rathnayake, R, Hoi, AY et al. Frequency and predictors of the lupus low disease activity state in a multi-national and multi-ethnic cohort. *Arthritis Res Ther*. 2016 Nov 9;18(1):260.

**APPROVALS**

CNTO1275 SLE2001 Statistical Analysis Plan

*A Multicenter, Randomized, Double-blind, Placebo-controlled, Proof-of-Concept Study of Ustekinumab (Stelara) in Subjects with Active Systemic Lupus Erythematosus*

*Version 1.0: Dated 22Mar2018*

**Approved by:**

---

██████████, MD  
Director, Clinical Research, Immunology  
Study Responsible Physician

Date (Day Month Year)

---

██████████, PhD  
Senior Director, Clinical Biostatistics  
Immunology Therapeutic Area Statistical Head

Date (Day Month Year)