Official Title of Study:

A PHASE 2, MULTICENTER, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED STUDY TO EVALUATE THE EFFICACY AND SAFETY OF CC-220 IN SUBJECTS WITH ACTIVE SYSTEMIC LUPUS ERYTHEMATOSUS

NCT Number: NCT03161483

Document Date (Date in which document was last revised): June 20, 2019

STATISTICAL ANALYSIS PLAN

A PHASE 2, MULTICENTER, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED STUDY TO EVALUATE THE EFFICACY AND SAFETY OF CC-220 IN SUBJECTS WITH ACTIVE SYSTEMIC LUPUS ERYTHEMATOSUS

STUDY DRUG: **CC-220**

PROTOCOL NUMBER: **CC-220-SLE-002**

DATE FINAL: **20 Jun 2019**

TABLE OF CONTENTS

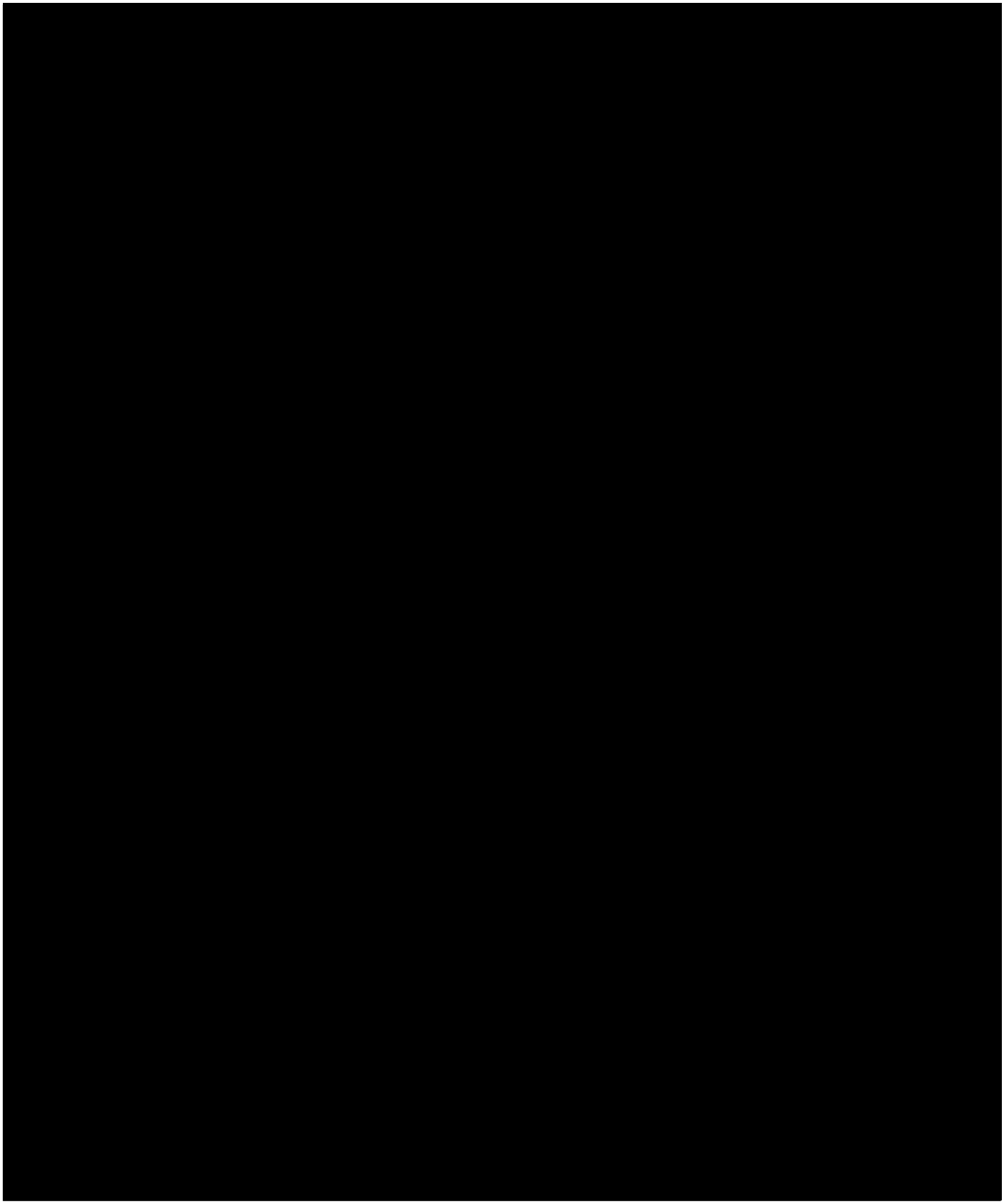
[redacted]

[redacted]

[redacted]

LIST OF TABLES

LIST OF FIGURES



1. LIST OF ABBREVIATIONS

Table 1: Abbreviations and Specialist Terms

Abbreviation	Description
ACR	American College of Rheumatology
AE	Adverse event
ALT	Alanine aminotransferase/ serum glutamic-pyruvic transaminase (SGPT)
AST	Aspartate aminotransferase/ serum glutamic-oxaloacetic transaminase (SGOT)
ATC	Anatomical Therapeutic Chemical
AUC	Area Under Curve
BILAG	British Isles Lupus Assessment Group 2004
BMI	Body mass index
CI	Confidence interval
CLASI	Cutaneous Lupus Area and Severity Index
CMH	Cochran-Mantel-Haenszel
CP	Conditional power
CTCAE	Common Terminology Criteria for Adverse Events
CV	Coefficient of variation
DMC	Data Monitoring Committee
ds-DNA	Double-stranded deoxyribonucleic acid
ECG	Electrocardiogram
eCRF	Electronic case report form
eGFR	Estimated glomerular filtration rate
ESR	Erythrocyte sedimentation rate
Hgb	Hemoglobin
hs-CRP	High sensitivity C-reactive protein
IFN	Interferon
IgA	Immunoglobulin A
IgG	Immunoglobulin G
IgM	Immunoglobulin M

Abbreviation	Description
IKZF	Ikaros
IL	Interleukin
IP	Investigational product
ITT	Intent-to-treat
LupusPRO	Lupus patient reported outcome
MCMC	Markov Chain Monte Carlo
MCS	Mental Component Summary
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple Imputation
mRNA	Messenger ribonucleic acid
PD	Pharmacodynamic
pDC	Plasmacytoid dendritic cell
PG	Pharmacogenetic
PGA	Physician's Global Assessment
PK	Pharmacokinetic
PP	Per protocol
PCS	Physical Component Summary
PT	Preferred term
QD	Once daily
SAP	Statistical analysis plan
SD	Standard deviation
SE	Standard error
SF-36	Short Form 36
SGOT	Serum glutamic oxaloacetic transaminase
SGPT	Serum glutamic pyruvic transaminase
SLE	Systemic lupus erythematosus
SLEDAI	Systemic lupus erythematosus Disease Activity Index
SLEDAI 2K	Systemic Lupus Erythematosus Disease Activity Index 2K

Abbreviation	Description
SLICC	Systemic Lupus International Collaborating Clinics
SOC	System Organ Class
SRI	SLE Responder Index
SRI(4)	SLE Responder Index response with 4 point reduction in the SLEDAI 2K score
SNP	Single Nucleotide Polymorphism
TEAE	Treatment-emergent adverse event
ULN	Upper limit of normal
WBC	White blood cells
WHODD	World Health Organization Drug Dictionary

[redacted]

3. STUDY OBJECTIVES

Table 2: Study Objectives

Primary Objective
– To evaluate the clinical efficacy of three doses of CC-220 (0.45 mg once per day [QD], 0.3 mg QD or 0.15 mg QD) compared to placebo, for the treatment of active systemic lupus erythematosus (SLE) using the SLE Responder Index at Week 24
Secondary Objectives
– To evaluate additional measures of clinical disease activity of CC-220 (0.45 mg once per day [QD], 0.3 mg QD or 0.15 mg QD) compared to placebo for the treatment of subjects with active SLE at Week 24 – To assess the reduction in steroid use – To assess the reduction in fatigue – To evaluate the safety and tolerability of three doses of CC-220 (0.45 mg QD, 0.3 mg QD, and 0.15 mg QD) compared to placebo in subjects with active SLE
[redacted]
[redacted]
[redacted]
[redacted]

4. INVESTIGATIONAL PLAN

4.1. Overall Study Design and Plan

CC-220-SLE-002 is a phase 2 randomized, placebo-controlled, double-blind, parallel group study with three active treatment groups.

The study consists of 5 phases:

- Screening Phase up to 5 weeks
- Randomized, Double-Blind, Placebo-Controlled Phase of up to 24 weeks
- Randomized, Double-Blind, Active Treatment Phase of up to 28 weeks
- Long-term Extension Phase of up to 52 weeks
- Observational Follow-up Phase of 4 weeks for females and 12 weeks for males

The total duration of the study is 113 weeks for females and 121 weeks for males. The total duration of the treatment period is 104 weeks.

Approximately 280 subjects will be randomized 2:2:1:2 to receive CC-220 (0.45 mg QD, 0.3 mg QD or 0.15 mg QD) or identically appearing placebo. There will be approximately 80 subjects randomized into the CC-220 0.45 mg QD and 0.3 mg QD dosing arms; 40 subjects randomized into the CC-220 0.15 mg QD dosing arm; and approximately 80 subjects in the placebo arm using an Interactive Response Technology (IRT). The treatment assignment will be stratified by baseline corticosteroid dose (≥ 10 mg/d and < 10 mg/d) and screening SLEDAI 2K score (≥ 10 points and < 10 points).

At Week 24, all remaining placebo subjects will be re-randomized 1:1 in a blinded fashion by IRT to either CC-220 0.45 mg QD or 0.3 mg QD. All subjects who were initially randomized to CC-220 (0.45 mg QD, 0.3 mg QD, or 0.15 mg QD) will be blindly re-randomized by IRT to the same dose group to which they were originally assigned.

Subjects who complete the Treatment Phase may be eligible to roll over into a Long-term Extension of up to one-year in duration. Subjects who enter this phase will maintain the CC-220 dosage they were re-randomized to at Week 24.

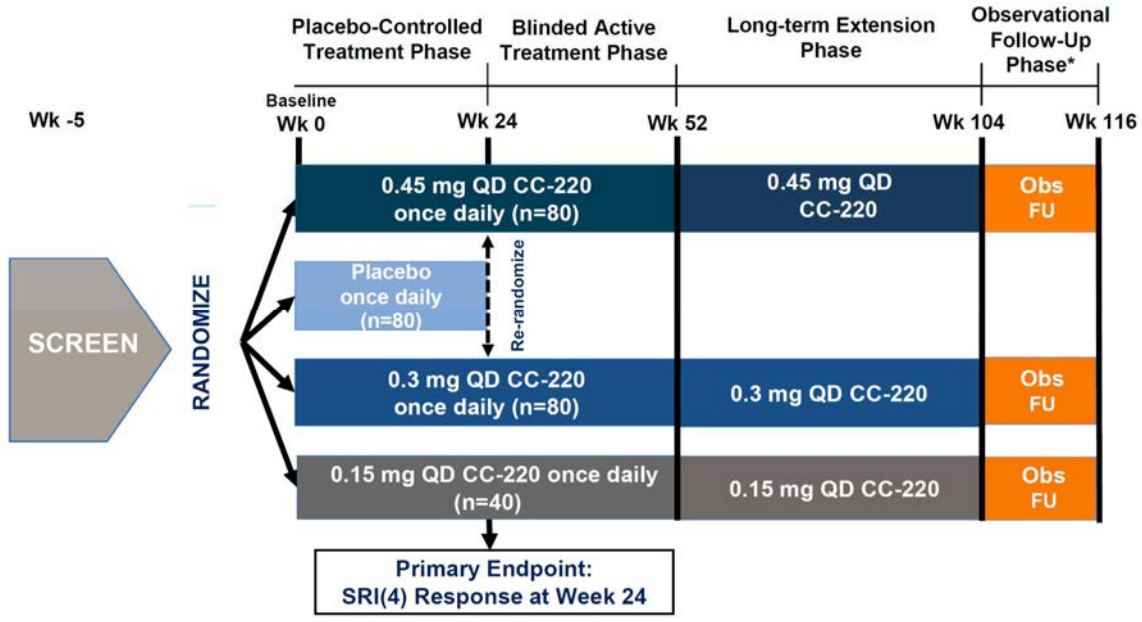
To maintain the blind at the site and subject level, the individual subject treatment assignments will not be revealed to the Investigators until after the 52-week database lock and after all final analyses are completed and the final results have been released.

Once an optimal dose is identified, subjects will be transitioned to that dose. If more than one dose is identified for further study, subject will be re-randomized accordingly as per the IRT.

If a subject discontinues investigational product (IP) early (ie, prior to completing Visit 15 during the Active Treatment Phase or prior to completing Visit 28 during the Long-term Extension), they will be required to complete an Early Termination Visit as soon as possible and enter into a 4-week Observational Follow-up Phase with an additional 12-week Observational Follow-up visit for males. In addition, if a subject completes the Randomized, Double-Blind, Active Treatment Phase, but opts not to enter the Long-term Extension Phase or completes the Long-term Extension Phase, the subject will enter into a 4-week Observational Follow-up Phase with an additional 12-week Observational Follow-up visit for males.

A 4-week Observational Follow-up Phase is included for all subjects to monitor subject safety after cessation of treatment. An additional 12-week Observational Follow-up visit is included for males to monitor hormone levels (testosterone, LH, FSH).

Figure 1: Overall Study Design



*Female subjects will have a 4-week Observational Follow-Up Phase. Males will have a 12-week Observational Follow-Up Phase.
Obs FU= Observational Follow Up Phase; QD=once per day; SRI(4)=SLE Responder Index(4); Wk=week.

4.2. Study Endpoints

4.2.1. Primary Endpoint

The primary endpoint is the percentage of subjects who achieve SRI(4) at Week 24. A subject achieves SRI(4) if the following criteria are all met:

- Change of ≤ -4 from baseline in the SLEDAI 2K score at Week 24

  The SLEDAI 2K score will be calculated as the sum of the scores of individual items.

- No new BILAG 2004 A score or more than one new BILAG 2004 B score compared to baseline at Week 24

  New A or B scores are determined for each organ system, rather than based on the total number of A or B scores across all organ systems. An A score is new if the baseline score for the same organ system is B, C, D or E. A B score is new if the baseline score for the same organ system is C, D, or E.

- Change of < 0.30 from baseline in the PGA score at Week 24
- No event of treatment failure (Section 5.6) before the date of the last SRI assessment (ie, SLEDAI 2K, BILAG 2004, or PGA, if they are assessed on different dates) at Week 24

Missing data handling is described in Section 5.5.

4.2.2. Secondary Endpoints

4.2.2.1. SLEDAI 2K

The secondary endpoint evaluating SLEDAI 2K is the percentage of subjects with an improvement of ≥ 4 from baseline in the SLEDAI 2K score at Week 24, which is a component of the primary endpoint. A subject achieves this endpoint if the following criteria are all met:

- Change of ≤ -4 from baseline in the SLEDAI 2K score at Week 24

  The SLEDAI 2K score will be calculated as the sum of the scores of individual items.

- No event of treatment failure (Section 5.6) before the date of the SLEDAI 2K assessment at Week 24

Missing data handling is described in Section 5.5.

4.2.2.2. CLASI

The secondary endpoint evaluating CLASI is the percentage of subjects with an improvement of ≥ 50% from baseline in the CLASI activity score at Week 24, in subjects with baseline CLASI activity score ≥ 10. A subject achieves this endpoint if the following criteria are all met:

- Percent change of ≤ -50% from baseline in the CLASI activity score at Week 24

  The CLASI activity score will be calculated as the sum of the individual activity scores (erythema in 13 locations, scale/hypertrophy in 13 locations, mucous membrane lesions, recent hair loss, and non-scarring alopecia).

- No event of treatment failure (Section 5.6) before the date of the CLASI assessment at Week 24

Missing data handling is described in Section 5.5.

4.2.2.3. BILAG 2004

The secondary endpoint evaluating BILAG 2004 is the percentage of subjects with no new BILAG 2004 A score or more than one new BILAG 2004 B score compared to baseline at Week 24, which is a component of the primary endpoint. A subject achieves this endpoint if the following criteria are all met:

- No new BILAG 2004 A score or more than one new BILAG 2004 B score compared to baseline at Week 24

  New A or B scores are determined for each organ system, rather than based on the total number of A or B scores across all organ systems. An A score is new if the baseline score for the same organ system is B, C, D or E. A B score is new if the baseline score for the same organ system is C, D, or E.

- No event of treatment failure (Section 5.6) before the date of the BILAG 2004 assessment at Week 24

Missing data handling is described in Section 5.5.

4.2.2.4. PGA

Percentage of subjects with no worsening from baseline (defined as change of < 0.30 from baseline) in the PGA score at Week 24

This endpoint is a component of the primary endpoint. A subject achieves this endpoint if the following criteria are all met:

- Change of < 0.30 from baseline in the PGA score at Week 24
- No event of treatment failure (Section 5.6) before the date of the PGA assessment at Week 24

Missing data handling is described in Section 5.5.

Change from baseline in the PGA score at Week 24

This endpoint will be evaluated by the mean change from baseline in the PGA score at Week 24. Data after the date of treatment failure will be considered missing for the efficacy analysis.

4.2.2.5. Joint Counts

The secondary endpoints evaluating joint counts are the mean change from baseline in swollen joint count at Week 24 in subjects with ≥ 2 swollen joints at baseline and the mean change from baseline in tender joint count at Week 24 in subjects with ≥ 2 tender joints at baseline. The handling of missing joint assessments is described in Section 5.5. Data after the date of treatment failure will be considered missing for the efficacy analysis.

4.2.2.6. Fatigue

The secondary endpoint evaluating fatigue is the mean change from baseline in the FACIT-Fatigue score at Week 24. Data after the date of treatment failure will be considered missing for the efficacy analysis. The FACIT-Fatigue score will be calculated as the sum of the 13 item scores ranging from 0 to 4.

4.2.2.7. Corticosteroid Reduction

Percentage of subjects with oral corticosteroid (OCS) dose reduced to ≤ 7.5 mg/day by Week 16 and maintained through Week 24 with no flares between Week 16 and Week 24, in subjects with baseline OCS dose ≥ 10 mg/day

A subject achieves this endpoint if the following criteria are all met:

- Daily OCS doses ≤ 7.5 mg/day from the date after the last SRI assessment (ie, SLEDAI 2K, BILAG 2004, or PGA, if they are assessed on different dates) at Week 16 to the date of the last SRI assessment at Week 24

  If no SRI assessment is found at Week 16 or Week 24 and thus the date of the last SRI assessment at Week 16 or Week 24 cannot be derived, the subject will be considered not achieving this endpoint.

- No flares as defined by the SLE Flare Index between Week 16 and Week 24, inclusive

  Missing flare assessments will not be imputed and will be considered as no flares.

- No event of treatment failure (Section 5.6) before the date of the last SRI assessment at Week 24

The baseline OCS dose is defined as the daily OCS dose on the date of the first dose of IP. The calculation of daily OCS doses is detailed in Appendix C.

Percentage of subjects with OCS dose reduced to < 10 mg/day by Week 16 and maintained through Week 24 with no flares between Week 16 and Week 24, in subjects with baseline OCS dose ≥ 10 mg/day

This endpoint will be derived in the same way as the one above, with < 10 mg replacing ≤ 7.5 mg.

Percent change from baseline in OCS dose at Week 24 in subjects with the baseline OCS dose ≥ 10 mg/day

This endpoint will be evaluated by the mean percent change and median percent change from baseline in OCS dose at Week 24 in subjects with the baseline OCS dose ≥ 10 mg/day. The baseline OCS dose is defined as the daily OCS dose on the date of the first dose of IP. The OCS dose at each post-baseline analysis visit is defined as the OCS dose on the date after the last SRI assessment (ie, SLEDAI 2K, BILAG 2004, or PGA, if they are assessed on different dates) at the analysis visit. Subjects without an OCS dose on the specified dates will have a dose of 0 at the corresponding analysis visit. The calculation of daily OCS doses is detailed in Appendix C. If no SRI assessment is found at an analysis visit and thus the date of the last SRI assessment at the analysis visit cannot be derived, the OCS dose at the analysis visit will be considered missing. OCS data after the date of treatment failure will still be used for this endpoint.

Standardized total OCS dose in Weeks 0-24

This endpoint will be evaluated by the mean standardized total OCS dose in Weeks 0-24.

The total OCS dose in Weeks 0-24 will be calculated using the records subjected to the calculation of daily OCS doses in Appendix C that have the calculated daily doses between the date of the first dose of IP and the date of the Week 24 visit, inclusive, ie, the start date is before the date of the first dose of IP and the end date (or the "effective" end date as described in Appendix C) is on or after the date of the first dose of IP or ongoing, or the start date is between the date of the first dose of IP and the date of the Week 24 visit, inclusive.

For each of these records, the duration (days) will be calculated as min(end date, date of the Week 24 visit) – max(start date, date of the first dose of IP) + 1, where the end date can be an "effective" end date as described in Appendix C, or date of the Week 24 visit – max(start date, date of the first dose of IP) + 1 if ongoing. The total dose for each record will be calculated as the calculated daily dose (mg/day) multiplied by the duration (days), and the total dose in Weeks 0-24 will be calculated as the sum of the total doses of these records. The standardized total dose in Weeks 0-24 will be calculated as the total dose in Weeks 0-24 divided by the actual number of days in Weeks 0-24 (date of the Week 24 visit – date of the first dose of IP + 1) multiplied by the normalizing duration of 168 (ie, 7 days/week × 24 weeks).

For subjects with no calculated daily doses between the date of the first dose of IP and the date of the Week 24 visit, inclusive, the standardized total dose in Weeks 0-24 will be 0.

For subjects who do not have the Week 24 visit, whether or not completing the placebo-controlled phase according to the placebo-controlled phase disposition eCRF, the date of completion or the date of discontinuation as collected in the placebo-controlled phase disposition eCRF will replace the date of the Week 24 visit in the above calculation.

4.2.4. Safety Endpoints

4.2.4.1. Adverse Events

Adverse events (AEs) will be coded according to the Medical Dictionary for Regulatory Activities (MedDRA). Treatment-emergent AEs (TEAEs) will be flagged for the safety analysis periods (Section 5.2) according to their start dates. Imputation rules of partial or completely missing start dates of AEs are given in Appendix B. AE intensity, relationship to IP, and seriousness will be ensured to be non-missing during medical monitoring.

Exposure-adjusted incidence rate (EAIR) per 100 person-years is defined as 100 times the number of subjects with the specific event divided by the total time (in years) at risk among subjects included in the analysis. Subjects with multiple occurrences of the specific event in the specific analysis period will be counted only once in the numerator. The time (in days) at risk for a subject without the specific event is the duration of the specific analysis period as described in Section 5.2, whereas the time (in days) at risk for a subject with the specific event is the duration of the specific analysis period up to the start date (inclusive) of the first occurrence of the specific event. The total time (in years) at risk will be calculated by dividing the sum of time (in days) at risk over all subjects included in the analysis by 365.25.

In the summary of TEAEs by time interval, the time intervals relative to the date of the first dose of IP (or the date of the first dose of CC-220) are the following:

- Placebo-controlled phase (Section 5.2): ≤ 4, > 4 to 8, > 8 to 12, > 12 weeks
- Weeks 0-52 CC-220-exposure period (Section 5.2): ≤ 4, > 4 to 8, > 8 to 12, > 12 to 24, > 24 to 36, > 36 weeks

For each time interval, a subject is counted once in the numerator of the percentage or EAIR for each applicable specific event if an occurrence starts in the time interval (eg, if the study day of

the start of an occurrence is 80, then the occurrence starts in the time interval of > 8 to 12 weeks). The denominator of the percentage is the number of subjects with the time at risk reaching the lower bound of the time interval, and the denominator of the EAIR is the sum of the time at risk during the time interval (up to the start date [inclusive] of the first occurrence of the specific event for each subject with an occurrence starting in the time interval) among the same number of subjects as in the denominator of the corresponding percentage.

In the summary of TEAEs by maximum severity, a subject with multiple occurrences of a specific event within a safety analysis period (Section 5.2) will be counted only once by the maximum severity in the safety analysis period.

4.2.4.2. Clinical Laboratory Tests

The endpoints of clinical laboratory tests include:

- The change from baseline by time point (including the follow-up visit) in the protocol-specified laboratory analytes of hematology, serum chemistry, immunology, and dipstick urinalysis (numeric parameters only) in subjects with a baseline value and a value at the respective post-baseline time point
- Shifts from baseline by time point (including each post-baseline time point, the maximum value, the minimum value, and the last value in a safety analysis period [Section 5.2]) in abnormality category (low/normal/high, or normal/abnormal where applicable, eg, for some urinalysis parameters) according to the reference range for the protocol-specified laboratory analytes of hematology, serum chemistry, immunology, and dipstick urinalysis

  For the summary of the maximum and minimum values, all data in the period, including any multiple values within the same analysis visit window, will be used to derive this endpoint. Percentages will be based on subjects with a baseline value and at least one post-baseline value in the period (at the respective time point for the summary by time point).
- The number and percentage of subjects with laboratory marked abnormalities (Appendix D) by time point and at least once in a safety analysis period (Section 5.2)

  For the summary of at least once in a given safety analysis period, all data in the period, including any multiple values within the same analysis visit window, will be used to derive this endpoint. Percentages will be based on subjects with a baseline value and at least one post-baseline value in the period (at the respective time point for the summary by time point) for the marked abnormality criteria referencing the baseline, or subjects with at least one post-baseline value in the period (at the respective time point for the summary by time point) for the marked abnormality criteria not referencing the baseline.
- Neutrophil ("neutrophils, segmented and band form"), lymphocyte, and hemoglobin shifts from baseline by time point (including each post-baseline time point, the minimum value, and the last value in a safety analysis period [Section 5.2]) in CTCAE grade for overall subjects and by baseline use of methotrexate (yes, no),

azathioprine (yes, no), mycophenolic compounds (yes, no), oral tacrolimus or oral cyclosporine (yes, no)

For the summary of the minimum value, all data in the period, including any multiple values within the same analysis visit window, will be used to derive this endpoint. Percentages will be based on subjects with a baseline value and at least one post-baseline value in the period (at the respective time point for the summary by time point).

- Time (days) to the first neutropenia ("neutrophils, segmented and band form") of CTCAE grade 3 or above (ie, < 1 10^9/L)

  The time (days) to the first neutropenia of CTCAE grade 3 or above is defined as the study day as defined in Section 5.3.2 for the first neutropenia of CTCAE grade 3 or above. For subjects who do not have a neutropenia of CTCAE grade 3 or above in a given safety analysis period (Section 5.2), the time will be censored on the date of the last neutrophil collection in the period.

Only the data in a given safety analysis period (Section 5.2) will be included in the analysis of that period.

4.2.4.3. Vital Signs and Weight

The endpoints of vital signs and weight include:

- The change from baseline by time point (including the follow-up visit) in the vital signs (including weight) in subjects with a baseline value and a value at the respective post-baseline time point
- Shifts from baseline by time point (including each post-baseline time point, the maximum value, the minimum value, and the last value in a safety analysis period [Section 5.2]) in abnormality category (low/normal/high) according to the reference range for the applicable vital signs

  For the summary of the maximum and minimum values, all data in the period, including any multiple values within the same analysis visit window, will be used to derive this endpoint. Percentages will be based on subjects with a baseline value and at least one post-baseline value in the period (at the respective time point for the summary by time point). The reference ranges are defined as: 60 to 100 beats/min for pulse, 90 to 140 mmHg for systolic blood pressure, and 60 to 90 mmHg for diastolic blood pressure.
- The number and percentage of subjects with vital signs marked abnormalities (Appendix E) by time point and at least once in a safety analysis period (Section 5.2)

  For the summary of at least once in a given safety analysis period, all data in the period, including any multiple values within the same analysis visit window, will be used to derive this endpoint. Percentages will be based on subjects with a baseline value and at least one post-baseline value in the period (at the respective time point for the summary by time point) for the marked abnormality criteria referencing the baseline, or subjects with at least one post-baseline value in the period (at the

respective time point for the summary by time point) for the marked abnormality criteria not referencing the baseline.

Only the data in a given safety analysis period (Section 5.2) will be included in the analysis of that period.

4.2.4.4. ECG

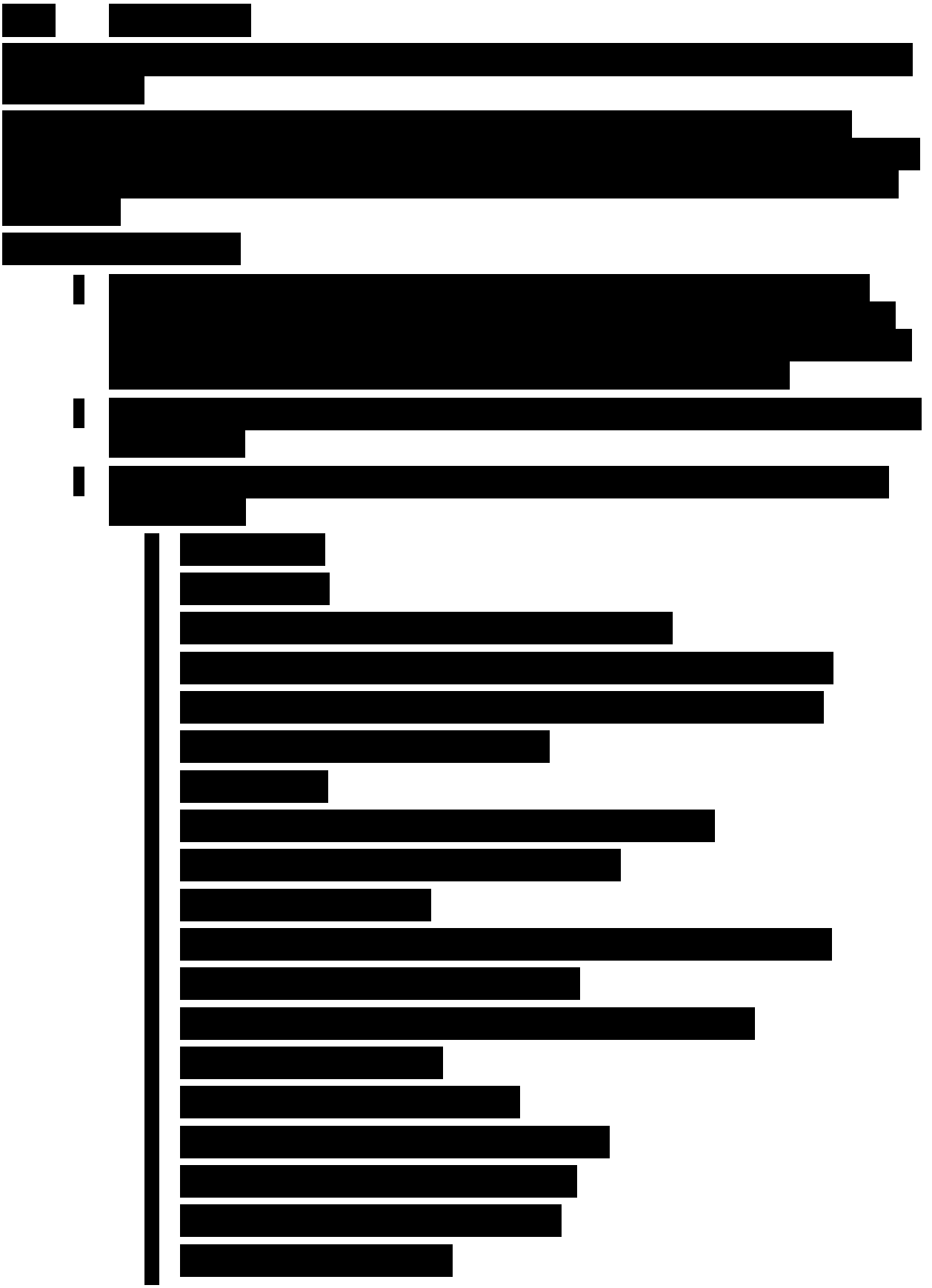
The ECG endpoints include:

- The change from baseline by time point in the following ECG variables in subjects with a baseline value and a value at the respective post-baseline time point:
  - Summary (mean) heart rate (beats/min)
  - Summary (mean) PR duration (msec)
  - Summary (mean) QRS duration (msec)
  - Summary (mean) QT duration (msec)
  - Summary (mean) RR duration (sec)
  - QTcB - Bazett's correction formula (msec)
  - QTcF - Fridericia's correction formula (msec)
- Shifts from baseline by time point (including each post-baseline time point, the worst result, and the last result in a safety analysis period [Section 5.2]) in abnormality category (normal/abnormal, not clinically significant/abnormal, clinically significant ECG interpretation results)

  For the summary of the worst result, all data in the period, including any multiple results within the same analysis visit window, will be used to derive this endpoint. Percentages will be based on subjects with a baseline result and at least one post-baseline result in the period (at the respective time point for the summary by time point). The ECG interpretation result will be based on the investigator's clinical interpretation.
- The number and percentage of subjects with ECG potentially clinically significant values/changes from baseline (Appendix F) by time point and at least once in a safety analysis periods (Section 5.2)

  For the summary of at least once in a given safety analysis period, all data in the period, including any multiple values within the same analysis visit window, will be used to derive this endpoint. Percentages will be based on subjects with a baseline value and at least one post-baseline value in the period (at the respective time point for the summary by time point) for the potentially clinically significant changes from baseline, or subjects with at least one post-baseline value in the period (at the respective time point for the summary by time point) for the potentially clinically significant values.

Only the data in a given safety analysis period (Section 5.2) will be included in the analysis of that period.



[redacted]

4.3. Stratification, Randomization, and Blinding

Approximately 280 subjects will be randomized 2:2:1:2 to receive CC-220 (0.45 mg QD, 0.3 mg QD or 0.15 mg QD) or identically appearing placebo. There will be approximately 80 subjects randomized into the CC-220 0.45 mg QD and 0.3 mg QD dosing arms; 40 subjects randomized into the CC-220 0.15 mg QD dosing arm; and approximately 80 subjects in the placebo arm using an Interactive Response Technology (IRT). The treatment assignment will be stratified by baseline corticosteroid dose (≥ 10 mg/d and < 10 mg/d) and screening SLEDAI 2K score (≥ 10 points and < 10 points).

At Week 24, all remaining placebo subjects will be re-randomized 1:1 in a blinded fashion by IRT to either CC-220 0.45 mg QD or 0.3 mg QD. All subjects who were initially randomized to CC-220 (0.45 mg QD, 0.3 mg QD, or 0.15 mg QD) will be blindly re-randomized by IRT to the same dose group to which they were originally assigned.

Subjects who complete the Treatment Phase may be eligible to roll over into a Long-term Extension of up to one-year in duration. Subjects who enter this phase will maintain the CC-220 dosage they were re-randomized to at Week 24.

To maintain the blind at the site and subject level, the individual subject treatment assignments will not be revealed to the Investigators until after the 52-week database lock and after all final analyses are completed and the final results have been released.

Once an optimal dose is identified, subjects will be transitioned to that dose. If more than one dose is identified for further study, subject will be re-randomized accordingly as per the IRT.

4.4. Sample Size Determination

Approximately 280 subjects will be randomized 2:2:1:2 to receive CC-220 (0.45 mg QD, 0.3 mg QD or 0.15 mg QD) or placebo. There will be approximately 80 subjects randomized into each CC-220 0.45 mg QD, CC-220 0.3 mg QD, and placebo arm and approximately 40 subjects into the CC-220 0.15 mg QD arm.

A sample size of 80 subjects in the CC-220 0.45 mg QD, CC-220 0.3 mg QD, and placebo dose groups provides approximately 80% power to detect a true 21% difference (55% versus 34%) between one of the CC-220 group and the placebo group, using a two-group chi-square test with a 0.1 two-sided significance level, for the percentage of subjects achieving an SRI(4) response at Week 24. Based upon clinical trials of belimumab and anifrolumab in a similar patient population with SLE, the observed placebo response rate ranged from 34% to 44%. A 21% treatment effect provides a clinically meaningful improvement compared to placebo and is approximately 50% higher than the treatment differences achieved in published trials in SLE (Human Genome Sciences, Inc., 2010; Furie, 2016).

The sample size of the 0.15 mg QD dose group is limited to 40 subjects to assess and determine the lowest effective dose.

5. GENERAL STATISTICAL CONSIDERATIONS

5.1. Reporting Conventions

All statistical analyses will be conducted using SAS® Version 9.2 or higher.

All analyses adjusting for the randomization stratification factors and the subgroup analyses with respect to the randomization stratification factors will use the actual stratifications derived from the clinical database, rather than the stratifications as randomized in the IRT.

P-values will be 2-sided and presented with 3 decimal places. P-values that are rounded to 0.000 will be presented as "< 0.001" and p-values that are rounded to 1 "> 0.999". Confidence intervals (CIs) will be presented as 2-sided 95% CIs and formatted to one more decimal place than the measured value.

Descriptive statistics for continuous variables include sample size (n), mean, standard deviation (SD), median, minimum, the 25th (Q1) and 75th (Q3) percentiles, and maximum. All mean, median, Q1, Q3, and SD values will be formatted to one more decimal place than the measured value. Minimum and maximum values will be presented to the same number of decimal places as the measured value. A frequency summary for categorical variables includes number and percentage. Number and percentage will be presented in the form XX (XX.X), where the percentage is in the parentheses. All percentages will be rounded to one decimal place, except 100% which will be displayed without any decimal places. Percentages will not be displayed for zero counts.

The change from baseline is calculated as the post-baseline value minus the baseline value. The percent change from baseline is calculated as the change from baseline divided by the baseline value multiplied by 100.

All laboratory data will be reported using standard international units.

Listings will be sorted by treatment group and subject number.

5.2. Safety Analysis Periods

For the safety analysis, the following analysis periods for the post-baseline data are defined. Safety data in these periods are considered treatment-emergent.

- Placebo-controlled phase: This period encompasses the post-baseline data through Week 24. It starts on the date of the first dose of IP, and ends on the date before the first Week 24 dose of CC-220 for subjects who receive at least one dose of CC-220 after Week 24, or 28 days after the last dose of IP or on the date of the last follow-up, whichever comes first, for subjects who do not receive at least one dose of CC-220 after Week 24. See Appendix B for the imputation of partial or completely missing dates of the last dose of IP.

  For the analysis of measured (ie, visit-based) safety data (ie, labs, vital signs, and ECG), this period is exclusive of the date of the first dose of IP and inclusive of the date of the first Week 24 dose of CC-220.

At the time of the interim analysis or the Week 24 analysis, if the date of the first Week 24 dose of CC-220 (derived from the Study Drug Exposure eCRF which may not be entered by sites until the next visit after Week 24) is not yet available for subjects with CC-220 dispensed at Week 24 and on or before the (possibly derived) date of the last dose of IP as of the time of analysis (ie, no evidence that the subject has discontinued without receiving at least one dose of CC-220 after Week 24), the date of CC-220 dispensing at Week 24 will replace the date of the first Week 24 dose of CC-220 in the above definition for these subjects.

- Weeks 0-52 CC-220-exposure period: This period encompasses all CC-220-exposure data through Week 52, irrespective of when the CC-220 exposure starts (at Week 0 or Week 24). It starts on the date of the first dose of CC-220, and ends on the date before the first Week 52 dose of CC-220 for subjects who receive at least one dose of CC-220 after Week 52, or 28 days after the last dose of IP or on the date of the last follow-up, whichever comes first, for subjects who do not receive at least one dose of CC-220 after Week 52.

  For the analysis of measured (ie, visit-based) safety data (ie, labs, vital signs, and ECG), this period is exclusive of the date of the first dose of CC-220 and inclusive of the date of the first Week 52 dose of CC-220.

  At the time of the Week 52 analysis, if the date of the first Week 52 dose of CC-220 (derived from the Study Drug Exposure eCRF which may not be entered by sites until the next visit after Week 52) is not yet available for subjects with CC-220 dispensed at Week 52 and on or before the (possibly derived) date of the last dose of IP as of the time of analysis (ie, no evidence that the subject has discontinued without receiving at least one dose of CC-220 after Week 52), the date of CC-220 dispensing at Week 52 will replace the date of the first Week 52 dose of CC-220 in the above definition for these subjects.

- Weeks 0-104 CC-220-exposure period: This period encompasses all CC-220-exposure data through Week 104, irrespective of when the CC-220 exposure starts (at Week 0 or Week 24). It starts on the date of the first dose of CC-220, and ends 28 days after the last dose of IP or on the date of the last follow-up, whichever comes first.

  For the analysis of measured (ie, visit-based) safety parameters (ie, labs, vital signs, and ECG), this period is exclusive of the date of the first dose of CC-220.

- Extension phase: This period encompasses the data in Weeks 52-104. It starts on the date of the first Week 52 dose of CC-220, and ends 28 days after the last dose of IP or on the date of the last follow-up, whichever comes first.

Testosterone, follicle stimulating hormone (FSH), and luteinizing hormone (LH) that are collected at the 12-week follow-up visit for males only will be included in the above safety analysis periods, irrespective of their timing relative to the end of the respective period.

5.3. Time Points

5.3.1. Screening Value and Baseline Value Definitions

Where applicable (eg, the screening SLEDAI 2K score), the screening value is defined as the last non-missing value with the recorded visit being the screening visit.

Unless otherwise specified, the baseline value is defined as the last non-missing value on or before the date of the first dose of IP. For the safety analysis, the baseline value is defined as the last non-missing value on or before the date of the first dose of IP for the analysis of the placebo-controlled phase, and the last non-missing value on or before the date of the first dose of CC-220 for the analysis of an CC-220-exposure period.

For the above definitions, in case of multiple non-missing values on the same date, the one with the earliest time will be used if the time of assessment/collection is available (eg, lab and ECG); otherwise the average of the multiple non-missing values will be used.

5.3.2. Analysis Visit Windows

All visit-based data, except for those with the recorded visit being a follow-up visit, will be assigned to analysis visits based on study day (the date of assessment/collection relative to the reference date) and the defined analysis visit windows (Appendix A), rather than the recorded visit. The only exception is that data with the recorded visit being a follow-up visit will be assigned to the analysis visit corresponding to the follow-up visit.

Unless otherwise specified, study day is defined as date of assessment/collection – date of the first dose of IP + 1 if the date of assessment/collection is on or after the date of the first dose of IP, or date of assessment/collection – date of the first dose of IP if the date of assessment/collection is before the date of the first dose of IP. For visit-based safety data (ie, labs, vital signs, and ECG) in the CC-220-exposure period for subjects who switch from placebo to CC-220 at Week 24, study day is defined as date of assessment/collection – date of the first dose of CC-220 + 1 so that the time points for the analyses of visit-based safety data in the CC-220-exposure period are relative to the same reference date, ie, the date of the first dose of CC-220.

Unless otherwise specified, the analysis visit windows are continuous, mutually exclusive, and stretch from the date after the reference date to the midpoint between the first two scheduled visits for the endpoint in question, from midpoint to midpoint between each successive pair of scheduled visits, and from the last midpoint onward for the last scheduled visit.

For visit-based analyses, if multiple non-missing values are within the same analysis visit window, then the value closest to the scheduled study day of the analysis visit will be used; if two values are equidistant from the scheduled study day, then the later value will be used. In case of multiple non-missing values on the same date, the one with the earliest time will be used if the time of assessment/collection is available (eg, lab and ECG); if the time is unavailable, the one with the earliest visit number will be used; if the visit numbers are also the same, the worst of the multiple non-missing values will be used.

For non-visit-based analyses (eg, lab marked abnormalities), all data, including any multiple values within the same analysis visit window, will be used.

5.4. Analysis Populations

5.4.1. Intent-to-treat Population

The intent-to-treat (ITT) population will include all subjects who are randomized and receive at least 1 dose of IP. The analysis of placebo-controlled (Weeks 0-24) efficacy and PD data in this study will be based on the ITT population unless otherwise specified and subjects will be included in the randomized treatment group.

5.4.2. Safety Population

The safety population will include all subjects who are randomized and receive at least 1 dose of IP. The analysis of placebo-controlled safety data in this study will be based on the safety population and subjects will be included in the treatment group corresponding to the IP they actually received.

5.4.3. Per-protocol Population

The per-protocol (PP) population will include all subjects in the ITT population who have no protocol deviations that may substantially affect the efficacy results. These protocol deviations could include, but are not limited to, not meeting key inclusion/exclusion criteria, gross IP non-compliance, receiving incorrect IP, treatment unblinding by the investigator, concomitant use of protocol-prohibited medications that are potentially effective for SLE for an unrelated comorbid condition, not having the required assessment at the designated time point without experiencing a treatment failure, or discontinuing due to lack of efficacy or a drug-related AE prior to the time point. The final determination of protocol deviations and thereby the composition of the PP population will be made prior to the DBL for the Week 24 analysis and will be separately documented. A supportive analysis of the primary endpoint will be based on the PP population and subjects will be included in the randomized treatment group.

[redacted]

5.4.5. CC-220 Subjects as Randomized Population

The efficacy analysis in Weeks 0-52 will be based on the CC-220 subjects as randomized population, which will include all subjects who are either initially randomized (at Week 0) or re-randomized (at Week 24) to a CC-220 treatment group, and receive at least one dose of CC-220 after (re-)randomization. Subjects will be included in the (re-)randomized CC-220 treatment group.

5.4.6. CC-220 Subjects as Treated Population

The safety analysis for the Weeks 0-52 and Weeks 0-104 CC-220-exposure periods will be based on the CC-220 subjects as treated population, which will include all subjects who are either initially randomized (at Week 0) or re-randomized (at Week 24) to a CC-220 treatment group,

and receive at least one dose of CC-220 after (re-)randomization. Subjects will be included in the CC-220 treatment group corresponding to the CC-220 dose they actually received.

5.4.7. Extension Phase Population

The efficacy and safety analysis in the extension phase (Weeks 52-104) will be based on the extension phase population, which will include subjects who receive at least one dose of CC-220 after Week 52. Subjects will be included in the (re-)randomized CC-220 treatment group.

5.5. Missing Data

Binary efficacy endpoints

Missing data for binary efficacy endpoints will be handled by nonresponder imputation (NRI), by which a subject will be considered a nonresponder at a given time point if the subject 1) does not have sufficient data (including the baseline data for the endpoints assessing the change from baseline) assessed within the analysis visit window for response determination, or 2) has had an event of treatment failure (Section 5.6) before the date of assessment (or the date of the last assessment in the case of a composite endpoint involving multiple criteria that may be assessed on different dates) for the time point.

Continuous efficacy and PD endpoints through Week 52 when assessed by the mean change or percent change from baseline

Missing data for continuous efficacy and PD endpoints through Week 52 when they are assessed by the mean change or percent change from baseline will be handled by an adaptive approach (Mehrotra, 2012) that uses either a longitudinal data analysis (LDA) model (Liu, 2009) (in the absence of severe departures from normality) or multiple imputation in conjunction with a robust regression model that uses M-estimation (in the presence of severe departures from normality). Unless otherwise specified, data after the date of treatment failure will be considered missing for the efficacy analysis.

The LDA model assumes a common mean across treatment groups at baseline and a different mean for each treatment group at each of the post-baseline time points. In this model, the response vector consists of the baseline and post-baseline values for the analysis of a change from baseline, or the baseline value and the post-baseline percent changes from baseline for the analysis of a percent change from baseline. Time is treated as a categorical variable so that no restriction is imposed on the trajectory of the means over time. Where specified, the model will also adjust for the randomization stratification factors (baseline corticosteroid dose [$\geq$ 10 mg/d and < 10 mg/d], screening SLEDAI 2K score [$\geq$ 10 points and < 10 points]) and their interactions with time. An unstructured covariance matrix will be used to model the correlation among repeated measurements. The Kenward-Roger adjustment will be used with restricted maximum likelihood to make proper statistical inference. If the model with an unstructured covariance fails to converge with the default Newton-Raphson algorithm, the Fisher scoring algorithm or other appropriate methods can be used to provide initial values of the covariance parameters. In the rare event that none of the above methods yields convergence, a structured covariance such as the heterogeneous Toeplitz or Toeplitz structures will be used to model the correlation among repeated measurements. In this case, the empirical option will be used because

the sandwich variance estimator is asymptotically unbiased while the model-based variance estimator can grossly overestimate or underestimate the true variance.

The LDA model assumes the vector of the model-based residuals follows a multivariate normal distribution. Under severe departures from normality, the LDA model can be inefficient or potentially misleading. Accordingly, the vector of the residuals from the LDA model, scaled by the inverse Cholesky root of the estimated variance-covariance matrix, will be subjected to a test for normality. If normality is not rejected at the significance level of 0.001, then the LDA model will be used for the analysis. However, if normality is rejected, then the analysis will be conducted using multiple imputation in conjunction with a robust regression model that uses M-estimation. Of note, the significance level of 0.001 for the normality test is chosen so that the LDA is abandoned only under a clear departure from normality.

Continuous efficacy [redacted] through Week 52 when assessed by the median change or percent change from baseline

Missing data for continuous efficacy and PD endpoints through Week 52 when they are assessed by the median change or percent change from baseline will be handled by multiple imputation in conjunction with a quantile regression model (using PROC QUANTREG). Unless otherwise specified, data after the date of treatment failure will be considered missing for the efficacy analysis.

Continuous efficacy data after Week 52

Continuous efficacy data after Week 52 will be summarized based on observed cases and missing data will not be imputed. Efficacy data after the date of treatment failure will be considered missing for the efficacy analysis.

Safety data

Safety data will be summarized based on observed cases and missing data will generally not be imputed. However, lab values of the form of "< x" or "> x" will be imputed by "x" in the calculation of descriptive statistics but displayed as "< x" or "> x" in the listings.

Joint counts

Joints not assessed (eg, recorded as "NA") at baseline will be imputed by the corresponding joint assessments at screening, if assessed; otherwise, these joints will be excluded from joint count for any analysis and reporting. Joints not assessed (eg, recorded as "NA") at post-baseline visits will be imputed using the last observation (baseline or post-baseline) carried forward (LOCF) method. If no assessment is performed for any joint at a given visit, LOCF will not apply and joint count at that visit will be missing.

BILAG 2004

Letter scoring (A, B, C, D, or E) of the 9 BILAG 2004 organ systems will be performed by the vendors in the raw data with the following rules:

- An item indicated as not due to SLE activity will be considered normal (ie, ignored in the scoring), regardless of the value recorded.
- For a given organ system,

- If any (but not all) of the items within the organ system has a missing value (eg, an item that should be evaluated but is not evaluated [ie, marked as NOT DONE], or a missing lab test value due to any reasons), it will be imputed by the last non-missing value within the last 70 days of the date of the current BILAG 2004 assessment (ie, up to the last 2 visits or 56 days from the current assessment plus additional 14 days to allow for variations of visit scheduling, and to be consistent with the defined analysis visit windows [Appendix A]). For the items in the renal organ system that involve a comparison with the previous visit, a missing value at the previous visit will be imputed by the last non-missing value within the last 98 days of the date of the ***current*** BILAG 2004 assessment (ie, the same 70-day window described above applies to the previous visit, which itself is 28 days from the current assessment). In the event that any item still has a missing value after imputation, the organ system will be scorable only if a score can be determined based on the available items ***and*** all items involved in the evaluation of the higher score(s) are present such that the higher score(s) can be precluded definitively; otherwise, the organ system will not be scorable and will have a missing letter score. See the additional exceptions below for the renal and hematological organ systems.
- If all items within the organ system have missing values, then no imputation will be performed, and the organ system will not be scorable and will have a missing letter score.

- Scores D and E will be differentiated. If an organ system is scorable and has no A, B, or C at the current visit, and has no A, B, C, or D at all previous visits including baseline, the score at the current visit will be E. If an organ system is scorable and has no A, B, or C at the current visit, but has A, B, C, or D at any of the previous visits including baseline, the score at the current visit will be D.
- For the renal organ system,
  - Renal biopsy is not mandatory in this study. Accordingly, if the item "active nephritis" (item 89 in Appendix C of the protocol; the corresponding scoring rule is "histological evidence of active nephritis within last 3 months") has a missing value (eg, marked as NOT DONE), it will be ignored in the scoring. However, a non-missing value (eg, a value of Yes) of this item will still be included in the scoring.
  - 24-hour urine protein and urine albumin-creatinine ratio are not done in this study and therefore ignored in the scoring.
  - The reference visit for assessing the change at the current visit is the previous visit (or an unscheduled visit, if applicable, between the current visit and the previous visit).
  - Urine dipstick protein will be ignored in the scoring as long as urine protein-creatinine ratio is available (subject to the aforementioned imputation) at both the current visit and the reference visit. Otherwise, urine dipstick protein, if available (subject to the aforementioned imputation) at both the current visit and the reference visit, will be used as the substitute.

- For the hematological organ system,
    - The Coombs' test is not required in this study. Accordingly, if the item "Coombs' test positive (isolated)" (item 97 in Appendix C of the protocol; the corresponding scoring rule is "isolated Coombs' test positive") has a missing value (eg, marked as NOT DONE), it will be ignored in the scoring. However, a non-missing value (eg, a value of Yes) of this item will still be included in the scoring.

Partially missing items within other efficacy measures

For SLEDAI 2K (total score, clinical score, and organ system scores), CLASI activity and damage scores, and SLICC/ACR SLE damage index, if any (but not all) of the items of the score has a missing value, it will be imputed by the last non-missing value within the last 70 days of the date of the current assessment (ie, up to the last 2 visits or 56 days from the current assessment plus additional 14 days to allow for variations of visit scheduling, and to be consistent with the defined analysis visit windows [Appendix A]). If a non-missing value cannot be found within the aforementioned time limit, then the score will be missing. If all items of the score have missing values, then no imputation will be performed and the score will be missing.

For the derivation of SF-36, FACIT-Fatigue, LupusPRO, and HAQ-DI, partially missing items will be handled according to the respective official scoring algorithm.

Missing dates

The handling of partial or completely missing dates for certain data is given in Appendix B.

5.6. Treatment Failures

Treatment failure is defined as protocol-prohibited initiation or dose increase of concomitant medications indicated for SLE. Treatment failures will be identified, separately for Weeks 0-52 and Weeks 52-104, by medical review of medications with the medical history of SLE as the primary reason or an additional indication (according to the prior and concomitant medications eCRF and the SLE medical history eCRF). Medications that can be used for SLE but are indicated for non-SLE uses, systemic or topical, will not be included in this medical review to identify treatment failures; as such, uses of such concomitant medications will not lead to treatment failures. The treatment failure rules in the efficacy and PD analyses, ie, data after the date of treatment failure will be considered missing, and a subject will be considered a nonresponder after the date of treatment failure, apply only to the treatment failures within the corresponding period (Weeks 0-52 or Weeks 52-104). In other words, treatment failures in Weeks 0-52 will have no effect in the analysis of Weeks 52-104.

6. SUBJECT DISPOSITION

The number of subjects screened, the number and percentage of subjects discontinuing the screening phase and the reasons for discontinuation (as collected in the screening phase disposition eCRF), and the number and percentage of the inclusion/exclusion criteria failed will be summarized for all screened subjects.

The number and percentage of subjects included in each analysis population will be summarized by treatment group and overall for all randomized subjects. A listing of subjects excluded from the PP population and the reason for exclusion will be provided.

The number and percentage of subjects who enter, complete, and discontinue the placebo-controlled phase, active-treatment phase, extension phase, and follow-up phase, respectively, and the reasons for discontinuation (as collected in the respective phase disposition eCRF) will be summarized by treatment group and overall for the ITT population. Overall summaries of completion and discontinuation (including the reasons) through the active-treatment phase and through the extension phase, respectively, will also be provided by randomized treatment group and overall for the ITT population. A listing of discontinued subjects including the phase discontinued and the reason for discontinuation will be provided.

The number and percentage of subjects in each region, each country, and each site will be summarized by treatment group and overall for the ITT population.

7. PROTOCOL DEVIATIONS

Protocol deviations will be classified as protocol deviations (PD) or important protocol deviations (IPD) according to Celgene or Contract Research Organization (CRO)'s standard procedure. The identification and classification of PD/IPD will be finalized and documented prior to the DBL for each applicable planned analysis (Section 2).

A summary of IPD by treatment group and overall will be provided for the screening phase and placebo-controlled phase, respectively, for the ITT population, and for the active-treatment phase and extension phase for subjects who receive at least one dose of CC-220 after Week 24 and Week 52, respectively.

A listing of IPD will be provided for the ITT population.

8. DEMOGRAPHIC AND BASELINE CHARACTERISTICS

8.1. Demographic and Baseline Characteristics

Demographic and baseline characteristics will be summarized descriptively by treatment group and overall for the ITT population. The comparability of the treatment groups for each relevant characteristic will be assessed descriptively; no statistical hypothesis tests will be performed on these characteristics.

Listings of demographic and baseline characteristics will be provided for the ITT population.

Demographic and baseline characteristic variables include:

- Age (years)
- Age (< 40, 40 to 50, > 50 to < 65, ≥ 65 years)
- Sex (male, female)
- Race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Not Collected or Reported, Other)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino, Not Reported, Unknown)
- Region (North America, Europe, South America [including Mexico], Russia)
- Weight (kg)
- Weight (< 55, 55 to < 70, 70 to < 85, 85 to < 100, ≥ 100 kg)
- Body mass index (BMI; kg/m^2)
- BMI (< 18.5, 18.5 to < 25, 25 to < 30, 30 to < 35, 35 to < 40, ≥ 40 kg/m^2)

Baseline disease characteristics include:

- Duration of SLE (years)
- Duration of SLE (≤ 1, > 1 to 2, > 2 to 5, > 5 to 10, > 10 years)
- Joint involvement (yes, no, unknown)
- Skin involvement (yes, no, unknown)
- Baseline cutaneous lupus subtype (acute, subacute, chronic)
- Screening SLEDAI 2K score per clinical database (≥ 10, < 10)
- Screening SLEDAI 2K score as randomized (≥ 10, < 10)
- Screening SLEDAI 2K score
- Baseline SLEDAI 2K score
- Screening clinical SLEDAI 2K score
- Baseline clinical SLEDAI 2K score
- BILAG 2004 global score

- BILAG 2004 1 A or 2 B scores (yes, no)
- BILAG 2004 mucocutaneous organ system score of A, B, or C (yes, no)
- BILAG 2004 renal organ system score of A, B, or C (yes, no)
- BILAG 2004 hematological organ system score of A, B, or C (yes, no)
- PGA score
- CLASI activity score
- CLASI activity score (≥ 10, < 10; ≥ 8, < 8)
- CLASI damage score
- Swollen joint count ≥ 6 and tender joint count ≥ 6 (yes, no)
- Swollen joint count ≥ 2 and tender joint count ≥ 2 (yes, no)
- Swollen joint count (≥ 2, ≤ 1)
- Tender joint count (≥ 2, ≤ 1)
- Swollen joint count
- Tender joint count
- Non-zero swollen joint count
- Non-zero tender joint count
- SLICC/ACR SLE damage index score
- Antinuclear antibodies $\geq$ 1:40 (yes, no)
- Antinuclear antibodies $\geq$ 1:80 (yes, no)
- anti-DNA antibodies (normal, high)
- Smith antibody (normal, abnormal)
- Sjogrens SS-A antibody (normal, abnormal)
- Sjogrens SS-B antibody (normal, abnormal)
- Sjogrens SS-A antibody abnormal and/or Sjogrens SS-B antibody abnormal (yes, no)
- Lupus anticoagulant (normal, high as determined by high in dilute Russell's viper venom time ratio)
- Cardiolipin antibody (any of IgG, IgM) (normal, high/abnormal)
- Beta-2 glycoprotein 1 antibody (any of IgG, IgM, IgA) (normal, high)
- High/abnormal in all of lupus anticoagulant, cardiolipin antibody (any of IgG, IgM), and Beta-2 glycoprotein 1 antibody (any of IgG, IgM, IgA) (yes, no)
- Ribonucleoprotein antibody (normal, abnormal)

- Baseline oral corticosteroids (OCS) (yes [alone, with antimalarials, with immuosuppressants], no)
- Baseline OCS dose (mg/d)
- Baseline OCS dose per clinical database (≥ 10 mg/d, < 10 mg/d)
- Baseline OCS dose as randomized (≥ 10 mg/d, < 10 mg/d)
- Baseline antimalarials (yes [hydroxychloroquine, quinacrine, chloroquine, alone, with immuosuppressants], no)
- Baseline immunosuppressants (yes [methotrexate, azathioprine, mycophenolic compounds, 6-mercaptopurine, leflunomide, sulfasalazine, oral tacrolimus, oral cyclosporine], no)

For categorical variables, a "missing" category will be added, where applicable, to account for subjects in the ITT population who have a missing value for the variable.

Unless otherwise indicated (eg, screening SLEDAI 2K score per clinical database), the above demographic and characteristic variables refer to the baseline values; see Section 5.3.1 for the screen value and baseline value definitions. In particular, baseline cutaneous lupus subtype (collected at the screening and baseline visits) is defined as the last value on or before the date of the first dose of IP (ie, the value at the screening visit will be used as the baseline value if the value at the baseline visit is not available). The derivations of age and disease duration are given in Appendix B. Joint involvement and skin involvement will be determined according to the SLE medical history eCRF.

Baseline SLE medications are defined as all medications with the medical history of "SYSTEMIC LUPUS ERYTHEMATOSUS (SLE)", "JOINT INVOLVEMENT ARTHRITIS", or "SKIN INVOLVEMENT" as the primary reason or an additional indication (according to the prior and concomitant medications eCRF and the SLE medical history eCRF) with an intake on the date of the first dose of IP (ie, start date on or before the date of the first dose of IP, and end date on or after the date of the first dose of IP or ongoing). The handling of partial or completely missing dates of medications is given in Appendix B. Specific SLE medications are identified as follows, together with the medical history of "SYSTEMIC LUPUS ERYTHEMATOSUS (SLE)", "JOINT INVOLVEMENT ARTHRITIS", or "SKIN INVOLVEMENT" being the primary reason or an additional indication:

- OCS: ATC2 level = H02 according to the Anatomical Therapeutic Chemical (ATC) coding scheme of the World Health Organization Drug Dictionary (WHODD) and oral route of administration
- Hydroxychloroquine: Preferred names containing HYDROXYCHLOROQUINE
- Quinacrine: Preferred names containing MEPACRINE
- Chloroquine: Preferred names containing CHLOROQUINE
- Methotrexate: Preferred names containing METHOTREXATE
- Azathioprine: Preferred names containing AZATHIOPRINE

- Mycophenolic compounds: Preferred names containing MYCOPHENOLIC ACID or MYCOPHENOLATE MOFETIL
- 6-mercaptopurine: Preferred names containing MERCAPTOPURINE
- Leflunomide: Preferred names containing LEFLUNOMIDE
- Sulfasalazine: Preferred names containing SULFASALAZINE
- Oral tacrolimus: Preferred names containing TACROLIMUS and oral route of administration
- Oral cyclosporine: Preferred names containing CICLOSPORIN and oral route of administration

The baseline OCS (prednisone-equivalent) dose is defined as the daily OCS dose on the date of the first dose of IP. The calculation of daily OCS doses is detailed in Appendix C.

8.2. Medical History

Medical history will be coded according to the MedDRA, and summarized by treatment group (and overall), system organ class (SOC), and preferred term (PT) for the ITT population. A separate summary of ongoing medical history will be provided similarly. A listing of medical history will be provided for the ITT population.

8.3. Prior and Concomitant Medications

Prior medications are defined as medications with a start date before the date of the first dose of IP (whether or not the end date is before the date of the first dose of IP). Concomitant medications during a safety analysis period defined in Section 5.2 are defined as medications taken at any time during the safety analysis period (ie, the start date is before the start date of the analysis period and the end date is on or after the start date of the analysis period or ongoing, or the start date is in the analysis period). The handling of partial or completely missing dates of medications is given in Appendix B.

Prior and concomitant medications will be coded using the ATC coding scheme of the WHODD. Prior medications will be summarized by treatment group (and overall), ATC2 level, and preferred name for the ITT population. Concomitant medications will be summarized by treatment group, ATC2 level, and preferred name for the placebo-controlled phase for the safety population, and for the CC-220-exposure periods for the CC-220 subjects as treated population.

A listing of prior and concomitant medications will be provided for the ITT population.

8.4. Prior and concomitant Procedures

Prior procedures are defined as procedures with a start date before the date of the first dose of IP. Concomitant procedures during a safety analysis period defined in Section 5.2 are defined as procedures performed at any time during the safety analysis period (ie, the start date is in the

analysis period). The handling of partial or completely missing dates of medications is given in Appendix B.

Prior and concomitant procedures will be coded according to the MedDRA. Prior medications will be summarized by treatment group (and overall), SOC, and PT for the ITT population. Concomitant procedures will be summarized by treatment group, SOC, and PT for the placebo-controlled phase for the safety population, and for the CC-220-exposure periods for the CC-220 subjects as treated population.

A listing of prior and concomitant procedures will be provided for the ITT population.

9. STUDY TREATMENTS AND EXTENT OF EXPOSURE

9.1. Treatment Duration

Treatment duration in days in a safety analysis period defined in Section 5.2 will be calculated as date of the last dose of IP in the analysis period – date of the first dose of IP in the analysis period + 1. Treatment duration in weeks is treatment duration in days divided by 7. The total person-years of treatment duration is the sum of treatment duration in days among all subjects divided by 365.25. For subjects who receive at least one dose of CC-220 after Week 24, the date of the last dose of IP in the placebo-controlled phase is defined as the date before the first Week 24 dose of CC-220. Similarly, for subjects who receive at least one dose of CC-220 after Week 52, the date of the last dose of IP in the Weeks 0-52 CC-220-exposure period is defined as the date before the first Week 52 dose of CC-220.

Descriptive statistics of treatment duration in weeks, a frequency summary of treatment duration categories, and the total person-years of treatment duration will be provided by treatment group for the placebo-controlled phase for the safety population, and for the CC-220-exposure periods for the CC-220 subjects as treated population. The treatment duration categories are the following:

- Placebo-controlled phase: ≤ 8, > 8 to 16, > 16 to 24, > 24 weeks
- Weeks 0-52 CC-220-exposure period: ≤ 8, > 8 to 16, > 16 to 24, > 24 to 38, > 38 to 52, > 52 weeks
- Weeks 0-104 CC-220-exposure period: ≤ 8, > 8 to 16, > 16 to 24, > 24 to 38, > 38 to 52, > 52 to 65, > 65 to 78, > 78 to 91, > 91 to 104, > 104 weeks

A listing of study drug exposure records will be provided.

9.2. Overdose

A listing of overdose records will be provided.

9.3. Treatment Compliance

Treatment compliance will be assessed using the drug accountability records. The compliance rate (%) in a period will be calculated as 100 times the total number of capsules taken over the period (the total number of dispensed capsules in the period minus the total number of returned capsules with the IP kit numbers matching those dispensed) divided by the intended total number of capsules that should have been taken over the same period. The total number of dispensed capsules in a period will be calculated from the following dispensing records:

- Placebo-controlled phase: The dispensing records at the scheduled visits before Week 24 and any unscheduled visits with a dispensing date before the dispensing date at Week 24 for subjects who receive at least one dose of CC-220 after Week 24, or all dispensing records for subjects who do not receive at least one dose of CC-220 after Week 24
- Active-treatment phase: The dispensing records at the scheduled visits between Week 24 and Week 52 (inclusive of Week 24 and exclusive of Week 52) and any unscheduled

visits with a dispensing date on or after the dispensing date at Week 24 and before the dispensing date at Week 52 for subjects who receive at least one dose of CC-220 after Week 52, or the dispensing records at the scheduled visits at or after Week 24 and any unscheduled visits with a dispensing date on or after the dispensing date at Week 24 for subjects who do not receive at least one dose of CC-220 after Week 52

- Extension phase: The dispensing records at the scheduled visits at or after Week 52 and any unscheduled visits with a dispensing date on or after the dispensing date at Week 52

The intended total number of capsules that should have been taken over a period will be calculated as 2 × (date of the last dose of IP in the period – date of the first dose of IP in the period + 1). For subjects who receive at least one dose of CC-220 after Week 24, the date of the last dose of IP in the placebo-controlled phase is defined as the date before the first Week 24 dose of CC-220. Similarly, for subjects who receive at least one dose of CC-220 after Week 52, the date of the last dose of IP in the active-treatment phase is defined as the date before the first Week 52 dose of CC-220.

Descriptive statistics of treatment compliance rate and a frequency summary of treatment compliance rate categories (< 75%, 75% – 120%, and > 120%) will be provided by treatment group for the placebo-controlled phase for the ITT population, and for the active-treatment phase and extension phase for subjects who receive at least one dose of CC-220 after Week 24 and Week 52, respectively.

A listing of drug accountability records will be provided.

10. EFFICACY ANALYSIS

Unless otherwise specified, the efficacy analysis in Weeks 0-24 (placebo-controlled phase), Weeks 0-52 (placebo-controlled phase through active-treatment phase), and Weeks 52-104 (extension phase) will be based on the ITT population (Section 5.4.1), CC-220 subjects as randomized population (Section 5.4.5), and extension phase population (Section 5.4.7), respectively. Unless otherwise specified, the efficacy analysis tables will be presented by the following treatment groups:

- Weeks 0-24 (placebo-controlled phase): placebo, 0.15 mg QD, 0.3 mg QD, and 0.45 mg QD
- Weeks 0-52 (placebo-controlled phase through active-treatment phase): placebo/0.3 mg QD, placebo/0.45 mg QD, 0.15 mg QD, 0.3 mg QD, and 0.45 mg QD
- Weeks 52-104 (extension phase): placebo/0.3 mg QD, placebo/0.45 mg QD, 0.15 mg QD, 0.3 mg QD, 0.45 mg QD, 0.3 mg QD combined, 0.45 mg QD combined

For the efficacy endpoints defined over Weeks 0-104 (as opposed to by time point in Weeks 0-104), ie, standardized total OCS dose in weeks 0-104, annualized (severe) flare rate in Weeks 0-104, time to the first (severe) flare in Weeks 0-104, and percentage of subjects who achieve SRI(4) at all of Week 24, Week 52, and Week 104, and at ≥ 70% of the visits between Week 24 and Week 104 in subjects who achieve SRI(4) at Week 24, the analysis will be based on the CC-220 subjects as randomized population and presented by placebo/0.3 mg QD, placebo/0.45 mg QD, 0.15 mg QD, 0.3 mg QD, and 0.45 mg QD.

For all endpoints, data based on observed cases will be summarized and presented in the same tables of the corresponding analysis results (where applicable). For continuous endpoints, descriptive statistics of the absolute value for subjects with a value at the post-baseline time point, the corresponding baseline value and change from baseline (also percent change from baseline only if it is analyzed as an endpoint) for subjects with a baseline value and a value at the post-baseline time point will be provided by time point, and data after the date of treatment failure will be considered missing for the efficacy analysis (ie, excluded from observed cases). For binary endpoints, a frequency summary will be provided by time point, and subjects who have sufficient data (ie, observed cases) but have had an event of treatment failure before the date of assessment for the time point, while included in the analysis, will be nonresponders.

Planned figures will be provided only where explicitly specified.

10.1. Multiplicity

The statistical tests comparing 0.3 mg QD with placebo and 0.45 mg QD with placebo with respect to the primary endpoint will be conducted at a 2-sided significance level of 0.1, with the multiplicity adjusted by the Hochberg procedure as follows. If the larger of the 2 p-values associated with the 2 tests is ≤ 0.1, then statistical significance will be declared for both tests; otherwise, statistical significance will not be declared for the test associated with the larger p-value and the smaller p-value will be compared with 0.05. If the smaller p-value is ≤ 0.05, then statistical significance will be declared for the test associated with the smaller p-value; otherwise, no statistical significance will be declared for both tests.

No multiplicity adjustments will be made for other comparisons. P-values from the tests not considered to be statistically significant due to the aforementioned multiplicity adjustment and from the tests not subject to multiplicity adjustment will be considered nominal.

10.2. Primary Endpoint

The primary endpoint, the percentage of subjects who achieve SRI(4) at Week 24, will be analyzed by the Cochran-Mantel-Haenszel (CMH) test stratified by the following randomization stratification factors:

- Baseline OCS dose (≥ 10 mg/d and < 10 mg/d)
- Screening SLEDAI 2K score (≥ 10 points and < 10 points)

For each treatment comparison between a CC-220 dose and placebo, the unstratified difference in percentages, the stratified difference in percentages and the 2-sided 95% stratified Newcombe CI (Yan, 2010) for the difference using the CMH weights, and the 2-sided p-value from the CMH test will be provided. The validity of the chi-square approximation for the distribution of the CMH statistic will be assessed by the Mantel-Fleiss criterion, using the CMH (MANTELFLEISS) option in PROC FREQ. If the Mantel-Fleiss criterion is >= 5, the CMH test based on the chi-square distribution will be used; otherwise, an exact test that the common odds ratio across strata equals one, using the COMOR option in the EXACT statement and the "Sum <= Point" definition of the 2-sided p-value, will be used.

If one and only one of the 2 treatment groups being compared has no subject in a stratum, the stratified treatment comparison will not be performed, and the 2-sided 95% unstratified Newcombe CI for the difference and the p-value from the chi-square test will be provided; if both treatment groups being compared have no subject in a stratum, the stratified treatment comparison will stratify by the remaining strata. When used, the validity of the chi-square test will be assessed by the rule implemented in PROC FREQ, using the CHISQ (WARN=OUTPUT) option that adds a warning indicator variable to the output dataset when more than 20% of the table cells have expected frequencies < 5, in which case the 2-sided Barnard's unconditional exact test, using the BARNARD option in the EXACT statement, will be used.

The following sensitivity analyses will be performed for the primary endpoint. All sensitivity analyses, except the tipping point analysis, will be presented in the same table as the primary analysis.

Sensitivity analysis with a tipping point analysis

To assess the impact of nonresponder imputation due to insufficient data for the Week 24 SRI(4) response determination (mainly due to early discontinuation before Week 24) (Section 5.5), a tipping point analysis, which varies the assumptions on the responses among subjects with insufficient data, will be performed as follows:

- The percentage of subjects who achieve SRI(4) at Week 24 will be analyzed by the chi-square test, instead of the CMH test, since each step of the tipping point analysis involves changing a non-specific subject with insufficient data from not achieving SRI(4) to achieving SRI(4) and thus the randomization strata cannot be determined and used for the CMH test.

- For the subset of subjects who do not have sufficient data (including the necessary baseline data) assessed within the analysis visit window for the Week 24 SRI(4) response determination and have ***not*** had an event of treatment failure (Section 5.6) before the date of the last SRI assessment (ie, SLEDAI 2K, BILAG 2004, or PGA, if they are assessed on different dates) at Week 24 if such a date exists, or before the date of completion or discontinuation in the placebo-controlled phase (as collected in the placebo-controlled phase disposition eCRF), the Week 24 SRI(4) response status will change from not achieving SRI(4) (as in the primary analysis) to achieving SRI(4) in an iterative manner. The SRI(4) response at Week 24 for subjects that are not in the aforementioned subset will be the same as that in the primary analysis.
- At each step of the tipping point analysis, one of the subjects from the aforementioned subset is changed from not achieving SRI(4) to achieving SRI(4), and the chi-square test is re-run. The test results in terms of statistical significance at the 2-sided significance level of 0.1 will be presented in a grid where the x-axis and y-axis represent the number of subjects assumed to achieve SRI(4) in the placebo group and the given CC-220 treatment group, respectively. The grid is limited from the top by the expected number of subjects from the aforementioned subset in the CC-220 treatment group that could have achieved SRI(4) had they had sufficient data (calculated as the number of subjects from the aforementioned subset in the CC-220 treatment group multiplied by the percentage of subjects achieving SRI(4) at Week 24 among subjects in the CC-220 treatment group who have sufficient data for the Week 24 SRI(4) response determination), and limited from the right by the number of subjects from the aforementioned subset in the placebo group.
- The grid will be divided into 3 regions:
    - Likely: This region's right limit is the expected number of subjects from the aforementioned subset in the placebo group that could have achieved SRI(4) had they had sufficient data (calculated as the number of subjects from the aforementioned subset in the placebo group multiplied by the percentage of subjects achieving SRI(4) at Week 24 among subjects in the placebo group who have sufficient data for the Week 24 SRI(4) response determination)
    - Uncertain: This region's right limit is the expected number of subjects from the aforementioned subset in the placebo group that could have achieved SRI(4) had they had sufficient data (calculated as the number of subjects from the aforementioned subset in the placebo group multiplied by the percentage of subjects achieving SRI(4) at Week 24 among subjects in the CC-220 treatment group who have sufficient data for the Week 24 SRI(4) response determination)
    - Unlikely: This is the region to the right of the uncertain region

Sensitivity analysis with multiple imputation

To assess the impact of nonresponder imputation due to insufficient data for the Week 24 SRI(4) response determination (mainly due to early discontinuation before Week 24) (Section 5.5), a sensitivity analysis using multiple imputation will be performed as follows:

- For the subset of subjects who do not have sufficient data (including the necessary baseline data) assessed within the analysis visit window for the Week 24 SRI(4) response determination and have ***not*** had an event of treatment failure (Section 5.6) before the date of the last SRI assessment (ie, SLEDAI 2K, BILAG 2004, or PGA, if they are assessed on different dates) at Week 24 if such a date exists, or before the date of completion or discontinuation in the placebo-controlled phase (as collected in the placebo-controlled phase disposition eCRF), missing values of the SRI assessments (SLEDAI-2k, BILAG 2004, or PGA) will be imputed separately, and the SRI(4) response at Week 24 will then be determined based on the imputed values of the SRI assessments. The SRI(4) response at Week 24 for subjects that are not in the aforementioned subset will be the same as that in the primary analysis. Missing values of the SRI assessments will be imputed as follows:
    - SLEDAI 2K: Missing SLEDAI 2K total scores at baseline, Weeks 4, 8, 12, 16, 20, and 24 will be imputed as a continuous variable. Enough missing values will first be imputed, by treatment group and randomization stratification factors, with PROC MI using the MCMC IMPUTE=MONOTONE specification and a VAR statement specifying the variables of SLEDAI 2K scores at baseline, Weeks 4, 8, 12, 16, 20, and 24 in order of analysis visit to result in monotone missing data patterns. Then monotone missing values will be imputed, by treatment group, with PROC MI using the regression method and a VAR statement specifying first the randomization stratification factors and then the variables of SLEDAI 2K scores at baseline, Weeks 4, 8, 12, 16, 20, and 24 in order of analysis visit. MINIMUM=0 and MAXIMUM=105 will be specified to ensure imputed values are within the range of SLEDAI 2K scores.
    - BILAG 2004: Missing BILAG 2004 binary responses, ie, whether or not the criterion of "no new BILAG 2004 A score or more than one new BILAG 2004 B score compared to baseline" is met, at Weeks 4, 8, 12, 16, 20, and 24 will be imputed as a binary variable, by treatment group, with PROC MI using the FCS LOGISTIC specification, a VAR statement specifying first the randomization stratification factors, baseline BILAG 2004 global score, and then the variables of BILAG 2004 binary response at Weeks 4, 8, 12, 16, 20, and 24 in order of analysis visit (all VAR variables except for the baseline BILAG 2004 global score will also be specified in the CLASS statement).
    - PGA: Missing PGA scores will be imputed in the same way as the SLEDAI 2K total scores. MINIMUM=0 and MAXIMUM=3 will be specified to ensure imputed values are within the range of PGA scores.
- 100 imputed datasets will be generated, and the results will be combined as follows:
    - The combined estimates of the percentage within each treatment group, and the unstratified and stratified differences in percentages will be calculated as the average of the corresponding estimates over all imputed datasets.
    - To calculate the combined 95% CIs for the stratified difference in percentages, the point estimate (calculated only for the purpose of calculating

the combined 95% CI) and associated standard error will be obtained from an adjusted Wald 95% CI (following the idea in Agresti, 1998) as follows for each imputed dataset and then combined using PROC MIANALYZE. The 95% CI produced by PROC MIANALYZE will be reported as the combined 95% CI. Let $\widetilde{p_{ij}} = x_{ij}/n_{ij}$ denote the estimated proportion in treatment group i and stratum j, w_j denote the CMH weight for stratum j, and $z_{0.975}$ denote the 97.5th percentile of the standard normal distribution. The point estimate is $\sum_j w_j (\widetilde{p_{ij}} - \widetilde{p_{i'j}})$ and the standard error is $\sqrt{\sum_j w_j^2 (\frac{\widetilde{p_{ij}}(1-\widetilde{p_{ij}})}{\widetilde{n_{ij}}} + \frac{\widetilde{p_{i'j}}(1-\widetilde{p_{i'j}})}{\widetilde{n_{i'j}}})}$, where $\widetilde{n_{ij}} = n_{ij} + z_{0.975}^2$ and $\widetilde{p_{ij}} = (x_{ij} + z_{0.975}^2/2)/\widetilde{n_{ij}}$.

- The CMH test will be conducted for each imputed dataset. The value of the CMH general association test statistic will be transformed using the Wilson-Hilferty transformation to create a more normally distributed statistic: $z = \frac{CMH^{1/3} - 7/9}{(2/9)^{1/2}}$, which is approximately normally distributed with a mean of 0 and a standard deviation of 1. The resulting transformed values from the 100 imputed datasets will be combined using PROC MIANALYZE and the combined p-value for the CMH test will be obtained as the upper-tailed p-value from the test produced by PROC MIANALYZE.

Sensitivity analysis with an observed case analysis

A sensitivity analysis with an observed case analysis will be performed based on the subset of subjects who have sufficient data (including necessary the baseline data) assessed within the analysis visit window for the Week 24 SRI(4) response determination, whether or not having had an event of treatment failure (Section 5.6) before the date of the last SRI assessment (ie, SLEDAI 2K, BILAG 2004, or PGA, if they are assessed on different dates) at Week 24 (subjects with an event of treatment failure before the date of the last SRI assessment at Week 24, while included in this analysis, will be nonresponders). The same analysis methods described for the primary analysis will then be applied to this subset.

Sensitivity analysis with a per-protocol analysis

A sensitivity analysis using the per-protocol population will be performed, with the same analysis methods described for the primary analysis applied to the per-protocol population.

Dose-response modeling

Dose-response modeling with different assumed dose-response shapes may be explored.

10.3. Secondary Endpoints

10.3.1. SLEDAI 2K

The secondary endpoint evaluating SLEDAI 2K is the percentage of subjects with an improvement of ≥ 4 from baseline in the SLEDAI 2K score at Week 24. The analysis methods described for the primary analysis of the primary endpoint (simply referred to as the analysis of the primary endpoint hereafter) in Section 10.2 will be used. The results will be presented in the

table of the longitudinal presentation of the results by time point in Weeks 0-24 rather than in a standalone table.

10.3.2. CLASI

The secondary endpoint evaluating CLASI is the percentage of subjects with an improvement of ≥ 50% from baseline in the CLASI activity score at Week 24, in subjects with baseline CLASI activity score ≥ 10. The analysis methods described for the analysis of the primary endpoint in Section 10.2 will be used. The results will be presented in the table of the longitudinal presentation of the results by time point in Weeks 0-24 rather than in a standalone table.

10.3.3. BILAG 2004

The secondary endpoint evaluating BILAG 2004 is the percentage of subjects with no new BILAG 2004 A score or more than one new BILAG 2004 B score compared to baseline at Week 24. The analysis methods described for the analysis of the primary endpoint in Section 10.2 will be used. The results will be presented in the table of the longitudinal presentation of the results by time point in Weeks 0-24 rather than in a standalone table.

10.3.4. PGA

Percentage of subjects with no worsening from baseline (defined as change of < 0.30 from baseline) in the PGA score at Week 24

The analysis methods described for the analysis of the primary endpoint in Section 10.2 will be used. The results will be presented in the table of the longitudinal presentation of the results by time point in Weeks 0-24 rather than in a standalone table.

Change from baseline in the PGA score at Week 24

The change from baseline in the PGA score at Week 24 will be analyzed by the adaptive approach that uses either a LDA model (in the absence of severe departures from normality) or multiple imputation in conjunction with a robust regression model that uses M-estimation (in the presence of severe departures from normality).

The LDA model assumes a common mean across treatment groups at baseline and a different mean for each treatment group at each of the post-baseline time points. The response vector consists of the baseline value and the post-baseline values at Weeks 4, 8, 12, 16, 20, and 24. Time is treated as a categorical variable so that no restriction is imposed on the trajectory of the means over time. The model will also adjust for the randomization stratification factors (baseline OCS dose [≥ 10 mg/d and < 10 mg/d], screening SLEDAI 2K score [≥ 10 points and < 10 points]), and their interactions with time. An unstructured covariance matrix will be used to model the correlation among repeated measurements. The Kenward-Roger adjustment will be used with restricted maximum likelihood to make proper statistical inference. If the model with an unstructured covariance fails to converge with the default Newton-Raphson algorithm, the Fisher scoring algorithm or other appropriate methods can be used to provide initial values of the covariance parameters. In the rare event that none of the above methods yields convergence, a structured covariance such as the heterogeneous Toeplitz or Toeplitz structures will be used to model the correlation among repeated measurements. In this case, the empirical option will be used because the sandwich variance estimator is asymptotically unbiased while the model-based

variance estimator can grossly overestimate or underestimate the true variance. Adjusted within-group means and the associated SEs, adjusted treatment differences in means between CC-220 doses and placebo and the associated SEs and 2-sided 95% CIs and 2-sided p-values will be derived from the LDA model.

The LDA model assumes the vector of the model-based residuals follows a multivariate normal distribution. Under severe departures from normality, the LDA model can be inefficient or potentially misleading. Accordingly, the vector of the residuals from the LDA model, scaled by the inverse Cholesky root of the estimated variance-covariance matrix, will be subjected to a test for normality. If normality is not rejected at the significance level of 0.001, then the LDA model will be used for the analysis. However, if normality is rejected, then the analysis will be conducted using multiple imputation in conjunction with a robust regression model that uses M-estimation. Of note, the significance level of 0.001 for the normality test is chosen so that the LDA is abandoned only under a clear departure from normality. Multiple imputation will be performed in the same way as described for the sensitivity analysis of the primary endpoint with multiple imputation (Section 10.2). A robust regression model that uses M-estimation will be run for each imputed dataset by using PROC ROBUSTREG, with the change from baseline at a given post-baseline time point (Weeks 4, 8, 12, 16, 20, or 24) as the response variable and adjusting for treatment group, the baseline value, and the randomization stratification factors (baseline corticosteroid dose [≥ 10 mg/d and < 10 mg/d] and screening SLEDAI 2K score [≥ 10 points and < 10 points]). The COVOUT option will be specified in the PROC ROBUSTREG statement to output the estimated covariance matrix of the parameter estimates to the OUTEST= dataset. Adjusted within-group means are not provided directly by PROC ROBUSTREG and will be calculated as (assuming the placebo group is coded as the last level of the treatment term and therefore the reference treatment group)

[redacted]

for the placebo group, where $\widehat{\beta_0}$, $\widehat{\beta_{\text{dose}\,i}}$, $\widehat{\beta_c}$, $\widehat{\beta_s}$, $\widehat{\beta_b}$ are the parameter estimates for the intercept, CC-220 dose i ($i = 0.15$ mg QD, 0.3 mg QD, 0.45 mg QD), baseline corticosteroid dose, screening SLEDAI 2K score, and the baseline value, respectively, from the OUTEST= dataset, and baseline mean is the overall mean of the baseline values across the treatment groups in the intended analysis population (ITT population in this case). The SEs for the adjusted within-group means will be calculated as $\sqrt{L'\Sigma L}$, where Σ is the estimated 7-dimentional covariance matrix of the parameter estimates from the OUTEST= dataset and L is the 7-dementional vector of coefficients in the aforementioned calculation of adjusted within-group means, ie, $L' = (1, 1, 0, 0, 0.5, 0.5, baseline\ mean)$, $L' = (1, 0, 1, 0, 0.5, 0.5, baseline\ mean)$, $L' = (1, 0, 0, 1, 0.5, 0.5, baseline\ mean)$, and $L' = (1, 0, 0, 0, 0.5, 0.5, baseline\ mean)$ for the CC-220 treatment and placebo groups, respectively. Adjusted treatment differences in means between CC-220 doses and placebo and the associated SEs are simply the parameter estimates $\widehat{\beta_{\text{dose}\,i}}$ and the associated SEs provided in the ODS table ParameterEstimates. The above results from each imputed dataset will be combined using PROC MIANALYZE, from which adjusted within-group means and the associated SEs, adjusted treatment differences in means between

CC-220 doses and placebo and the associated SEs and 2-sided 95% CIs and 2-sided p-values will be provided.

The results at each post-baseline time point (Weeks 4, 8, 12, 16, 20, and 24) will be presented. Line plots of adjusted means with SEs (also showing the number of subjects and the adjusted mean) by time point will be provided.

10.3.5. Joint Counts

The secondary endpoints evaluating joint counts are the mean change from baseline in swollen joint count at Week 24 in subjects with ≥ 2 swollen joints at baseline and the mean change from baseline in tender joint count at Week 24 in subjects with ≥ 2 tender joints at baseline. The analysis methods described for the mean change from baseline in the PGA score in Section 10.3.4 will be used. Line plots of adjusted means with SEs (also showing the number of subjects and the adjusted mean) by time point will be provided.

10.3.6. Fatigue

The secondary endpoint evaluating fatigue is the mean change from baseline in the FACIT-Fatigue score at Week 24. The analysis methods described for the mean change from baseline in the PGA score in Section 10.3.4 will be used, with the appropriate post-baseline time points. Line plots of adjusted means with SEs (also showing the number of subjects and the adjusted mean) by time point will be provided.

10.3.7. Corticosteroid Reduction

Percentage of subjects with oral corticosteroid (OCS) dose reduced to ≤ 7.5 mg/day by Week 16 and maintained through Week 24 with no flares between Week 16 and Week 24, in subjects with baseline OCS dose ≥ 10 mg/day

The analysis methods described for the analysis of the primary endpoint in Section 10.2 will be used, except that the randomization stratification factors in the stratified analysis will be reduced to only screening SLEDAI 2K score (≥ 10 points and < 10 points).

Percentage of subjects with OCS dose reduced to < 10 mg/day by Week 16 and maintained through Week 24 with no flares between Week 16 and Week 24, in subjects with baseline OCS dose ≥ 10 mg/day

The analysis methods described for the above endpoint will be used. The analysis of this endpoint will be presented in the same table as the one above.

Percent change from baseline in OCS dose at Week 24 in subjects with the baseline OCS dose ≥ 10 mg/day

The results of the mean percent change and median percent change from baseline will be presented in one table.

To assess the mean percent change, the analysis methods described for the mean change from baseline in the PGA score in Section 10.3.4 will be used. The randomization stratification factors in the stratified analysis will be reduced to only screening SLEDAI 2K score (≥ 10 points and < 10 points). The response vector in the LDA model consists of the baseline value and the post-baseline percent changes from baseline, whereas the response variable in the robust regression

model is the percent change from baseline at a given post-baseline time point. As many zero changes from baseline in OCS dose are expected, model fitting issues may arise due to these excessive zeros, particularly with the robust regression model. In the event of unresolvable modeling fitting issues, the corresponding modeling results will not be provided at the time points impacted.

To assess the median percent change, multiple imputation in conjunction with a quantile regression model will be used. Multiple imputation will be performed in the same way as described for the sensitivity analysis of the primary endpoint with multiple imputation (Section 10.2). A quantile regression model will be run for each imputed dataset by using PROC QUANTREG, with the percent change from baseline at a given post-baseline time point (Weeks 4, 8, 12, 16, 20, or 24) as the response variable and adjusting for treatment group, the baseline value, and the randomization stratification factor (screening SLEDAI 2K score [≥ 10 points and < 10 points]). The CI=RESAMPLING option will be specified in the PROC QUANTREG statement and the QUANTILE=0.5 option will be specified in the MODEL statement. Adjusted within-group medians and the associated SEs will be estimated by using the ESTIMATE statement (eg, ESTIMATE '0.15 mg QD median' INTERCEPT 1 TREATMENT 1 0 0 0 BLOCS *baseline mean*, where the treatment term is coded in the order of 0.15 mg QD, 0.3 mg QD, 0.45 mg QD, placebo, BLOCS is the variable for baseline OCS dose, and the overall baseline mean OCS dose across the treatment groups in the intended analysis population [ITT population with the baseline OCS dose ≥ 10 mg/day in this case] is used as the coefficient for BLOCS) and the results are provided in the ODS table Estimates. Adjusted treatment differences in medians between CC-220 doses and placebo and the associated SEs are simply the parameter estimates and the associated SEs for the treatment term (coded in the order of 0.15 mg QD, 0.3 mg QD, 0.45 mg QD, placebo) provided in the ODS table ParameterEstimates. The above results from each imputed dataset will be combined using PROC MIANALYZE, from which adjusted within-group medians and the associated SEs, adjusted treatment differences in medians between CC-220 doses and placebo and the associated SEs and 2-sided 95% CIs and 2-sided p-values will be provided. Other quantiles (other than 50% or median) of the percent change from baseline may be explored in the same manner if no treatment effects are demonstrated for the median percent change from baseline. The results at each post-baseline time point (Weeks 4, 8, 12, 16, 20, and 24) will be presented in one table.

Standardized total OCS dose in Weeks 0-24

The analysis methods described for the mean change from baseline in the PGA score in Section 10.3.4 will be used. The response vector in the LDA model consists of the baseline OCS dose and the calculated post-baseline standardized total OCS dose, whereas the response variable in the robust regression model is the calculated post-baseline standardized total OCS dose (the baseline OCS dose still as a covariate). By construction of this endpoint, missing data is not possible, so multiple imputation will not be applicable if the robust regression model is used under a clear departure from normality.

The adjusted average daily OCS dose will be calculated as the adjusted mean standardized total OCS dose divided by the normalizing duration of 168 days.

[redacted]

10.5. Subgroup Analysis

To assess whether the treatment effect is consistent across various subgroups, treatment differences in percentages and nominal 2-sided 95% CIs for the primary endpoint (percentage of subjects who achieve SRI(4) at Week 24) will be provided within each category of the following classification variables:

- Age (< 40, 40 to < 65, ≥ 65 years)
- Sex (male, female)
- Race (Black or African American, White, others)
- Ethnicity (Hispanic or Latino, others)
- Region (North America, Europe, South America [including Mexico], Russia)
- Baseline weight (< 70, ≥ 70 kg)
- Duration of SLE (≤ 2, > 2 to 5, > 5 years)

- Baseline cutaneous lupus subtype (acute, subacute, chronic, any of acute, subacute, or chronic, any of subacute or chronic)
- Screening SLEDAI 2K score per clinical database (≥ 10, < 10)
- Baseline BILAG 2004 1 A or 2 B scores (yes, no)
- Baseline CLASI activity score (≥ 8, < 8)
- Baseline CLASI activity score (≥ 10, < 10)
- Baseline swollen joint count ≥ 6 and tender joint count ≥ 6 (yes, no)
- Baseline antinuclear antibodies ≥ 1:40 (yes, no)
- Baseline OCS dose per clinical database (≥ 10 mg/d, < 10 mg/d)
- Baseline antimalarials (yes, no)
- Baseline immuosuppressants (yes, no)

If the value of a classification variable cannot be determined, the subject will be excluded from the corresponding subgroup analysis.

In addition to the individual treatment groups, the 0.3 mg QD and 0.45 mg QD treatment groups will be combined to form an additional CC-220 treatment group in the subgroup analysis. The unstratified percentage within each treatment group will be provided. For each treatment comparison between a CC-220 group and the placebo group, the unstratified difference in percentages, the stratified difference in percentages and the 2-sided 95% stratified Newcombe CI (Yan, 2010) for the difference using the CMH weights will be provided by category of the classification variables listed above.

For the subgroup analysis of each of the 2 randomization stratification factors (baseline OCS dose and screening SLEDAI 2K score), the stratified analysis will only stratify by the other factor. If one and only one of the 2 treatment groups being compared has no subject in a stratum, the stratified treatment comparison will not be performed (the stratified difference in percentages and the 2-sided 95% stratified Newcombe CI for the difference using the CMH weights will not be provided) and the 2-sided 95% unstratified Newcombe CI for the difference will be provided; if both treatment groups being compared have no subject in a stratum, the stratified treatment comparison will stratify by the remaining strata.

Forest plots for the treatment differences in percentages by subgroup will be provided.

11. SAFETY ANALYSIS

The safety analysis for the placebo-controlled phase, the Weeks 0-52 and Weeks 0-104 CC-220-exposure periods, and the extension phase (Section 5.2) will be based on the safety population (Section 5.4.2), the CC-220 subjects as treated population (Section 5.4.6), and the extension phase population (Section 5.4.7), respectively. The safety analysis tables will be presented by the following treatment groups:

- Placebo-controlled phase: placebo, 0.15 mg QD, 0.3 mg QD, 0.45 mg QD, and CC-220 total
- CC-220-exposure periods and extension phase: 0.15 mg QD, 0.3 mg QD, 0.45 mg QD, and CC-220 total

Unless otherwise specified, all safety summaries will be provided separately for the placebo-controlled phase, the Weeks 0-52 CC-220-exposure period, and the Weeks 0-104 CC-220-exposure period. All AE summaries will also be provided for the extension phase, unless otherwise specified.

11.1. Adverse Events

Unless otherwise specified, all frequency summaries of AEs will provide the EAIR and be presented by system organ class (SOC) and preferred term (PT), with SOCs sorted in the standard international order and PTs within each SOC in descending order of percentage (and then alphabetically if needed). A subject with multiple occurrences will be counted only once for each applicable category (eg, any TEAE, or a given SOC or PT).

Overall summaries of subjects with at least one TEAE in the following categories will be provided:

- Any TEAE
- Any drug-related TEAE
- Any severe TEAE
- Any serious TEAE
- Any TEAE leading to drug interruption
- Any TEAE leading to drug withdrawal
- Any TEAE leading to death

Subjects with at least one TEAE will be summarized for the following categories:

- Any TEAE
- Any TEAE $\geq$ 5% in any treatment group, including CC-220 total (not provided for the extension phase)
- Any TEAE by time interval (not provided for the Weeks 0-104 CC-220-exposure period and the extension phase)
- Any drug-related TEAE

- Any TEAE by maximum severity (EAIR will not be provided for this summary)
- Any serious TEAE
- Any TEAE leading to drug withdrawal
- Any TEAE leading to death

Listings of all AEs (TEAEs and non-TEAEs), serious AEs (TEAEs and non-TEAEs), TEAEs leading to drug withdrawal, and all deaths will be provided.

11.2. Clinical Laboratory Tests

For the protocol-specified laboratory analytes of hematology, serum chemistry, immunology, and urinalysis, descriptive statistics of the baseline value, the absolute value and the change from baseline by time point (including the follow-up visit), and tables of shifts from baseline by time point (including each post-baseline time point, the maximum value [or the worst value for non-numeric urinalysis analytes], the minimum value [not applicable for non-numeric urinalysis analytes], and the last value in the safety analysis period [Section 5.2]) in abnormality category (low/normal/high, or normal/abnormal where applicable, eg, for some urinalysis parameters) according to the reference range will be provided. As noted in Section 5.2, these summaries will include testosterone, FSH, and LH that are collected at the 12-week follow-up visit for males only, irrespective of their timing relative to the end of the respective period.

A frequency summary of subjects with laboratory marked abnormalities (Appendix D) by time point and at least once in the safety analysis period (Section 5.2) will be provided.

A shift table of neutrophil ("Neutrophils, segmented and band form"), lymphocyte, and hemoglobin shifts from baseline by time point (including each post-baseline time point, the minimum value, and the last value in the safety analysis period [Section 5.2]) in CTCAE grade for overall subjects and by baseline use of methotrexate (yes, no), azathioprine (yes, no), mycophenolic compounds (yes, no), oral tacrolimus or oral cyclosporine (yes, no) will be provided.

For the time to the first neutropenia ("Neutrophils, segmented and band form") of CTCAE grade 3 or above, the numbers of subjects who have an event (neutropenia of CTCAE grade 3 or above) and who are censored will be summarized. The KM estimates of the 25th, median, and 75th percentile of the time to the first event and the 2-sided 95% CIs will be provided. The cumulative number of events, the number of subjects at risk, and the KM estimate of the event rate (ie, failure probability instead of survival probability) and SE will be provided by time point (Weeks 4, 8, 12, etc.). The KM plot of failure curves (instead of survival curves) with the number of subjects at risk at each time point will be presented.

A listing of all values of the applicable laboratory analytes will be provided for subjects with any laboratory marked abnormalities. A listing of all laboratory data will also be provided.

11.3. Vital Signs and Weight

Descriptive statistics of the baseline value, the absolute value and the change from baseline in vital signs and weight by time point (including the follow-up visit), and tables of shifts from baseline by time point (including each post-baseline time point, the maximum value, the

minimum value, and the last value in the safety analysis period [Section 5.2]) in abnormality category (low/normal/high) according to the reference range for the applicable vital signs will be provided.

A frequency summary of subjects with vital signs marked abnormalities (Appendix E) by time point and at least once in the safety analysis period (Section 5.2) will be provided.

A listing of all values of the applicable vital signs will be provided for subjects with any marked abnormalities. A listing of all vital signs and weight will also be provided.

11.4. ECG

Descriptive statistics of the baseline value, the absolute value and the change from baseline in ECG variables by time point (including the follow-up visit), and tables of shifts from baseline by time point (including each post-baseline time point, the worst result, and the last result in the safety analysis period [Section 5.2]) in abnormality category (normal/abnormal, not clinically significant/abnormal, clinically significant ECG interpretation results) will be provided.

A frequency summary of subjects with ECG potentially clinically significant values/changes from baseline (Appendix F) by time point and at least once in the safety analysis period (Section 5.2) will be provided.

A listing of all values of the applicable ECG variables will be provided for subjects with any potentially clinically significant values/changes from baseline. A listing of all ECG interpretation results will be provided for subjects with any clinically significantly abnormal interpretation result. A listing of all ECG variables and interpretation results will also be provided.

13. INTERIM ANALYSIS

One unblinded interim analysis is planned. The purpose of the interim analysis is to ensure that the study is conducted with an acceptable benefit/risk ratio by reviewing the safety and efficacy data, including the efficacy data in subgroups defined by baseline mRNA and DNA. The interim analysis will be conducted after approximately 50% of the subjects (approximately 140 subjects) have completed or discontinued prior to Week 24. An independent external Data Monitoring Committee (DMC) will review the interim analysis results and make recommendations, including discontinuing any CC-220 treatment group or the study. The operational details are documented in the DMC charter.

14. CHANGES TO THE STATISTICAL SECTION OF THE PROTOCOL

The following changes have been made to the statistical section of the protocol:

- Changes have been made to the exploratory efficacy endpoints as a result of expert input received after the finalization of the protocol version on which this SAP is based.
- The protocol states that no multiplicity adjustment will be conducted for this Phase 2 study. In this SAP, multiplicity adjustment using the Hochberg procedure will be applied to the statistical tests comparing 0.3 mg QD with placebo and 0.45 mg QD with placebo with respect to the primary endpoint.

[redacted]

16. APPENDICES

APPENDIX A. ANALYSIS VISIT WINDOWS

Table 3: Analysis Visit Windows

Analysis Visit	Scheduled Study Day[a]	Analysis Visit Window[a]
SLEDAI 2K, SLEDAI Flare Index, BILAG 2004, PGA, Swollen and Tender Joint Counts, ds-DNA, Complement Panel, Urine Protein/Creatinine, Chemistry, Urinalysis, eGFR, Vital Signs		
Baseline	≤ 1	≤ 1
Week 4	29	2 – 42
Week 8	57	43 – 70
Week 12	85	71 – 98
Week 16	113	99 – 126
Week 20	141	127 – 154
Week 24	169	• Except for the safety analysis in the CC-220-exposure period for subjects who switch from placebo to CC-220 at Week 24: 155 – Date of the first Week 24 dose of CC-220 (or date of CC-220 dispensing at Week 24 as a substitute[b]) for subjects who receive at least one dose of CC-220 after Week 24, or ≥ 155 for subjects who do not receive at least one dose of CC-220 after Week 24 • For the safety analysis in the CC-220-exposure period for subjects who switch from placebo to CC-220 at Week 24: 155 – 182
Week 28	197	• Except for the safety analysis in the CC-220-exposure period for subjects who switch from placebo to CC-220 at Week 24: Date after the first Week 24 dose of CC-220 – 210 • For the safety analysis in the CC-220-exposure period for subjects who switch from placebo to CC-220 at Week 24: 183 – 210
Week 32	225	211 – 238
Week 36	253	239 – 266
Week 40	281	267 – 294
Week 44	309	295 – 322
Week 48	337	323 – 350
Week 52	365	351 – 406
Week 64	449	407 – 490
Week 76	533	491 – 574
Week 88	617	575 – 672
Week 104	729	≥ 673
CLASI Activity		
Baseline	≤ 1	≤ 1
Week 4	29	2 – 42
Week 8	57	43 – 70
Week 12	85	71 – 98

Analysis Visit	Scheduled Study Day[a]	Analysis Visit Window[a]
Week 16	113	99 – 126
Week 20	141	127 – 154
Week 24	169	155 – Date of the first Week 24 dose of CC-220 (or date of CC-220 dispensing at Week 24 as a substitute[b]) for subjects who receive at least one dose of CC-220 after Week 24, or ≥ 155 for subjects who do not receive at least one dose of CC-220 after Week 24
Week 40	281	Date after the first Week 24 dose of CC-220 – 322
Week 52	365	323 – 406
Week 64	449	407 – 490
Week 76	533	491 – 574
Week 88	617	575 – 672
Week 104	729	≥ 673
CLASI Damage, HAQ-DI, Lupus Autoantibody Panel		
Baseline	≤ 1	≤ 1
Week 24	169	2 – Date of the first Week 24 dose of CC-220 (or date of CC-220 dispensing at Week 24 as a substitute[b]) for subjects who receive at least one dose of CC-220 after Week 24, or ≥ 2 for subjects who do not receive at least one dose of CC-220 after Week 24
Week 52	365	Date after the first Week 24 dose of CC-220 – 546
Week 104	729	≥ 547
FACIT Fatigue		
Baseline	≤ 1	≤ 1
Week 12	85	2 – 126
Week 24	169	127 – Date of the first Week 24 dose of CC-220 (or date of CC-220 dispensing at Week 24 as a substitute[b]) for subjects who receive at least one dose of CC-220 after Week 24, or ≥ 127 for subjects who do not receive at least one dose of CC-220 after Week 24
Week 36	253	Date after the first Week 24 dose of CC-220 – 308
Week 52	365	309 – 448
Week 76	533	449 – 630
Week 104	729	≥ 631
SF-36		
Baseline	≤ 1	≤ 1
Week 12	85	2 – 126

Analysis Visit	Scheduled Study Day[a]	Analysis Visit Window[a]
Week 24	169	127 – Date of the first Week 24 dose of CC-220 (or date of CC-220 dispensing at Week 24 as a substitute[b]) for subjects who receive at least one dose of CC-220 after Week 24, or ≥ 127 for subjects who do not receive at least one dose of CC-220 after Week 24
Week 52	365	Date after the first Week 24 dose of CC-220 – 448
Week 76	533	449 – 630
Week 104	729	≥ 631
LupusPRO		
Baseline	≤ 1	≤ 1
Week 24	169	2 – Date of the first Week 24 dose of CC-220 (or date of CC-220 dispensing at Week 24 as a substitute[b]) for subjects who receive at least one dose of CC-220 after Week 24, or ≥ 2 for subjects who do not receive at least one dose of CC-220 after Week 24
Week 52	365	Date after the first Week 24 dose of CC-220 – 406
Week 64	449	407 – 490
Week 76	533	491 – 574
Week 88	617	575 – 672
Week 104	729	≥ 673
SLICC/ACR SLE Damage Index		
Baseline	≤ 1	≤ 1
Week 52	365	2 – 546
Week 104	729	≥ 547
Hematology		
Baseline	≤ 1	≤ 1
Week 4	29	2 – 42
Week 8	57	43 – 70
Week 12	85	71 – 98
Week 16	113	99 – 126
Week 20	141	127 – 154
Week 24	169	• Except for the data in the CC-220-exposure period for subjects who switch from placebo to CC-220 at Week 24: 155 – Date of the first Week 24 dose of CC-220 (or date of CC-220 dispensing at Week 24 as a substitute[b]) for subjects who receive at least one dose of CC-220 after Week 24, or ≥ 155 for subjects who do not receive at least one dose of CC-220 after Week 24 • For the data in the CC-220-exposure period for subjects who switch from placebo to CC-220 at Week 24: 155 – 182

Analysis Visit	Scheduled Study Day[a]	Analysis Visit Window[a]
Week 28	197	• Except for the data in the CC-220-exposure period for subjects who switch from placebo to CC-220 at Week 24: Date after the first Week 24 dose of CC-220 – 210 • For the data in the CC-220-exposure period for subjects who switch from placebo to CC-220 at Week 24: 183 – 210
Week 32	225	211 – 238
Week 36	253	239 – 266
Week 40	281	267 – 294
Week 44	309	295 – 322
Week 48	337	323 – 350
Week 52	365	351 – 378
Week 56	393	379 – 406
Week 60	421	407 – 434
Week 64	449	435 – 462
Week 68	477	463 – 490
Week 72	505	491 – 518
Week 76	533	519 – 546
Week 80	561	547 – 574
Week 84	589	575 – 602
Week 88	617	603 – 630
Week 92	645	631 – 658
Week 96	673	659 – 686
Week 100	701	687 – 714
Week 104	729	≥ 715
ECG		
Baseline	≤ 1	≤ 1
Week 4	29	2 – 98
Week 24	169	• Except for the data in the CC-220-exposure period for subjects who switch from placebo to CC-220 at Week 24: 99 - Date of the first Week 24 dose of CC-220 (or date of CC-220 dispensing at Week 24 as a substitute[b]) for subjects who receive at least one dose of CC-220 after Week 24, or ≥ 99 for subjects who do not receive at least one dose of CC-220 after Week 24 • For the data in the CC-220-exposure period for subjects who switch from placebo to CC-220 at Week 24: 99 – 182
Week 28	197	• Except for the data in the CC-220-exposure period for subjects who switch from placebo to CC-220 at Week 24: Date after the first Week 24 dose of CC-220 – 280 • For the data in the CC-220-exposure period for subjects who switch from placebo to CC-220 at Week 24: 183 – 280

Analysis Visit	Scheduled Study Day[a]	Analysis Visit Window[a]
Week 52	365	281 – 546
Week 104	729	≥ 547
IgA, IgM, IgG, Testosterone, FSH, LH		
Baseline	≤ 1	≤ 1
Week 24	169	• Except for the data in the CC-220-exposure period for subjects who switch from placebo to CC-220 at Week 24: 2 – Date of the first Week 24 dose of CC-220 (or date of CC-220 dispensing at Week 24 as a substitute[b]) for subjects who receive at least one dose of CC-220 after Week 24, or ≥ 2 for subjects who do not receive at least one dose of CC-220 after Week 24 • For the data in the CC-220-exposure period for subjects who switch from placebo to CC-220 at Week 24: 2 – 266
Week 52	365	• Except for the data in the CC-220-exposure period for subjects who switch from placebo to CC-220 at Week 24: Date after the first Week 24 dose of CC-220 – 448 • For the data in the CC-220-exposure period for subjects who switch from placebo to CC-220 at Week 24: 267 – 448
Week 76	533	449 – 630
Week 104	729	≥ 631
ESR, hsCRP		
Baseline	≤ 1	≤ 1
Week 12	85	2 – 126
Week 24	169	• Except for the data in the CC-220-exposure period for subjects who switch from placebo to CC-220 at Week 24: 127 – Date of the first Week 24 dose of CC-220 (or date of CC-220 dispensing at Week 24 as a substitute[b]) for subjects who receive at least one dose of CC-220 after Week 24, or ≥ 127 for subjects who do not receive at least one dose of CC-220 after Week 24 • For the data in the CC-220-exposure period for subjects who switch from placebo to CC-220 at Week 24: 127 – 210
Week 36	253	• Except for the data in the CC-220-exposure period for subjects who switch from placebo to CC-220 at Week 24: Date after the first Week 24 dose of CC-220 – 308 • For the data in the CC-220-exposure period for subjects who switch from placebo to CC-220 at Week 24: 211 – 308
Week 52	365	≥ 309
Lupus Anti-Phospholipid Profile		
Baseline	≤ 1	≤ 1
Week 12	85	2 – 126

Analysis Visit	Scheduled Study Day[a]	Analysis Visit Window[a]
Week 24	169	127 – Date of the first Week 24 dose of CC-220 (or date of CC-220 dispensing at Week 24 as a substitute[b]) for subjects who receive at least one dose of CC-220 after Week 24, or ≥ 127 for subjects who do not receive at least one dose of CC-220 after Week 24
Week 36	253	Date after the first Week 24 dose of CC-220 – 308
Week 52	365	309 – 406
Week 64	449	407 – 490
Week 76	533	491 – 574
Week 88	617	575 – 672
Week 104	729	≥ 673
[redacted]		
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

[a] Unless otherwise specified, study day is defined as date of assessment/collection – date of the first dose of IP + 1 if the date of assessment/collection is on or after the date of the first dose of IP, or date of assessment/collection – date of the first dose of IP if the date of assessment/collection is before the date of the first dose of IP. For visit-based safety data (ie, labs, vital signs, and ECG) in the CC-220-exposure period for subjects who switch from placebo to CC-220 at Week 24, study day is defined as date of assessment/collection – date of the first dose of CC-220 + 1.

[b] At the time of the interim analysis or the Week 24 analysis, if the date of the first Week 24 dose of CC-220 (derived from the Study Drug Exposure eCRF which may not be entered by sites until the next visit after Week 24) is not yet available for subjects with CC-220 dispensed at Week 24 and on or before the (possibly derived) date of the last dose of IP as of the time of analysis (ie, no evidence that the subject has discontinued without receiving at least one dose of CD-220 after Week 24), the date of CC-220 dispensing at Week 24 will replace the date of the first Week 24 dose of CC-220 in the above definition for these subjects.

APPENDIX B. HANDLING OF PARTIAL OR COMPLETELY MISSING DATES

Age

Age (in years) will be calculated as the integer part of (date of informed consent – date of birth + 1) / 365.25 when the complete date of birth is collected; otherwise, the age recorded will be used.

Disease Duration

Disease duration (in years) will be calculated as (date of informed consent – date of diagnosis + 1) / 365.25. In case of a partial date of diagnosis, July 1 of the year will be used if only year is present, or the 16th of the month will be used if both year and month are present. Disease duration will be missing if the date of diagnosis is completely missing.

Date of the Last Dose of IP (as of the Time of Analysis)

The date of the last dose of IP as of the time of analysis will be determined from the last dose of study treatment eCRF or derived as follows:

- If a complete date exists in the last dose of study treatment eCRF, the date of the last dose of IP will be set to that date.
- If a partial date exists in the last dose of study treatment eCRF, the date of the last dose of IP will be set to the maximum of the following:
    - The earliest possible date given the non-missing field(s) of the partial date (ie, January 1 of the year will be used if only year is present, or the first date of the month will be used if both year and month are present), or
    - The minimum of the latest possible date given the non-missing field(s) of the partial date (ie, December 31 of the year will be used if only year is present, or the last date of the month will be used if both year and month are present), the date of the last follow-up, the date of the Week 104 visit, or the date of the early termination visit.
- If the date is completely missing in the last dose of study treatment eCRF, the date of the last dose of IP will be set to the minimum of the date of the last follow-up, the date of the Week 104 visit, or the date of the early termination visit.

Medical History

Partial or completely missing dates of medical history will not be imputed other than the handling described above in the derivation of disease duration.

Prior and Concomitant Medications/Procedures

Completely missing dates of medications/procedures will not be imputed. Medications with a completely missing start date will be considered prior and/or concomitant according to the response to the question of "did the medication start before the first dose of study treatment" in the prior and concomitant medications eCRF, and the (possibly imputed) end date or the ongoing indication. In case the response to the question is also missing, the medication will be considered prior (and concomitant if indicated by the [possibly imputed] end date or as ongoing). Procedures with a completely missing start date will be considered prior.

Partial start dates of medications/procedures and partial end dates of medications will be imputed. For medications, if the imputed start date is after the (possibly imputed) end date, then the start date will be imputed by the end date.

Partial start dates of medications/procedures will be imputed by the earliest possible date given the non-missing field(s) of the partial date, ie, January 1 of the year will be used if only year is present, or the first date of the month will be used if both year and month are present. For medications, however, if the imputed start date is before the date of the first dose of IP whereas the response to the question of "did the medication start before the first dose of study treatment" in the prior and concomitant medications eCRF is "no", then the start date will be imputed by the date of the first dose of IP.

Partial end dates of medications will be imputed by the latest possible date given the non-missing field(s) of the partial date, ie, December 31 of the year will be used if only year is present, or the last date of the month will be used if both year and month are present.

Adverse Events

Partial or completely missing start dates of AEs will be imputed. Partial end dates of AEs will also be imputed to ensure that the imputed start date is not after the imputed end date, whereas completely missing end dates of AEs will not be imputed. If the imputed start date is after the (possibly imputed) end date, then the start date will be imputed by the end date.

The imputation rules for partial or completely missing start dates of AEs are given in Table 4. The principle of the imputation rules is to treat the AE as treatment-emergent, ie, starting on or after the date of the first dose of IP, if possible.

Table 4: Imputation Rules for Partial or Completely Missing Start Dates of Adverse Events

Scenario	Imputation Rule
1. Partial start date with only year present	
a. $Y_{Event} < Y_{IP}$	$31/12/Y_{Event}$
b. Otherwise, ie, $Y_{Event} \geq Y_{IP}$	Max (date of the first dose of IP, $1/1/Y_{Event}$)
2. Partial start date with both year and month present	
a. $Y_{Event} < Y_{IP}$, or ($Y_{Event} = Y_{IP}$ and $M_{Event} < M_{IP}$)	Last date of M_{Event}/Y_{Event}
b. Otherwise, ie, $Y_{Event} > Y_{IP}$, or ($Y_{Event} = Y_{IP}$ and $M_{Event} \geq M_{IP}$)	Max (date of the first dose of IP, $1/M_{Event}/Y_{Event}$)
3. Completely missing start date	Date of the first dose of IP
$D_{Event}/M_{Event}/Y_{Event}$ represents the AE start date. $D_{IP}/M_{IP}/Y_{IP}$ represents the date of the first dose of IP.	

Partial end dates of AEs will be imputed by the latest possible date given the non-missing field(s) of the partial date, ie, December 31 of the year will be used if only year is present, or the last date of the month will be used if both year and month are present.

APPENDIX C. DERIVATIOIN OF DAILY ORAL CORTICOSTEROID DOSES

Daily oral corticosteroid (OCS) (prednisone-equivalent) doses will be calculated for the records in the prior and concomitant medications eCRF satisfying all the following:

- ATC2 level = H02
- The medical history of "SYSTEMIC LUPUS ERYTHEMATOSUS (SLE)", "JOINT INVOLVEMENT ARTHRITIS", or "SKIN INVOLVEMENT" as the primary reason or an additional indication (according to the prior and concomitant medications eCRF and the SLE medical history eCRF)
- Oral route of administration

For the records with a dosing frequency of less frequent than daily, the recorded dose will be divided equally according to the dosing frequency to calculate the daily dose (eg, for the dosing frequency of 1 time per week, the daily dose will be the recorded dose divided by 7). The "effective" end date corresponding to the daily dose for such records will be derived by start date + number of administrations × dosing frequency – 1, where the number of administrations will be determined as (end date – start date + 1) / dosing frequency, rounded up to the next integer. For example, if the start and end dates are July 1 and July 15, respectively, and the dosing frequency is 1 time per week, the number of administrations is (July 15 – July 1 + 1) / 7 = 3, and the "effective" end date corresponding to the daily dose is July 1 + 3 × 7 – 1 = July 21.

Corticosteroids administered PRN will not be considered in the calculation of the daily dose.

If more than one record contributes doses to a given date (due to, for example, doses from different records with overlapping start/end dates, or doses allocated from records with a dosing frequency of less frequent than daily), the daily dose on that date will be the sum of these doses.

Daily OCS doses will be expressed as prednisone-equivalent doses. Conversions to prednisone-equivalent doses are given in Table 5.

Table 5: Conversions to Prednisone-equivalent Doses

Corticosteroid (Preferred Term)	Equivalent Dose
Prednisone	10 mg
Betamethasone	1.2 mg
Cortisone	50 mg
Deflazacort	12 mg
Dexamethasone	1.5 mg
Hydrocortisone	40 mg
Methylprednisolone	8 mg
Prednisolone	10 mg
Triamcinolone	8 mg

APPENDIX D. LABORATORY MARKED ABNORMALITY CRITERIA

Table 6: Laboratory Marked Abnormality Criteria

Category/Analyte	SI Unit	Criteria
Chemistry		
Alanine Aminotransferase (SGPT)	U/L	> 3 x ULN
Albumin	Kg/m^3	< 25
Alkaline Phosphatase	U/L	> 2 x ULN
Aspartate Aminotransferase (SGOT)	U/L	> 3 x ULN
Total Bilirubin	µmol/L	> 2 x ULN
Blood Urea Nitrogen	mmol/L	> 15
Calcium	mmol/L	< 1.8 > 3.0
Creatinine	µmol/L	> 2 x ULN
Glucose	mmol/L	< 2.8 > 13.9
Lactate Dehydrogenase	U/L	> 3 x ULN
Magnesium	mmol/L	> 1.2
Phosphate	mmol/L	< 0.64 > 1.60
Potassium	mmol/L	< 3.0 > 5.5
Sodium	mmol/L	< 130 > 150
Urine Protein Creatinine Ratio	g/mol	> 226 Increase > 56.5 from baseline
[redacted]	[redacted]	[redacted]
Hematology		
Hemoglobin	g/L	< 80 Male: > 185, Female: > 170
Leukocytes	10^9/L	< 1.5
Lymphocytes	10^9/L	< 0.5
Neutrophils	10^9/L	< 1.0
Platelets	10^9/L	< 75 > 600
Eosinophils	10^9/L	> 1.0

APPENDIX E. VITAL SIGNS MARKED ABNORMALITY CRITERIA

Table 7: Vital Signs Marked Abnormality Criteria

Vital Sign	Unit	Criteria
Diastolic Blood Pressure	mmHg	Increase from baseline > 5 but ≤ 10 Increase from baseline > 10 but ≤ 15 Increase from baseline > 15
Systolic Blood Pressure	mmHg	Increase from baseline > 10 but ≤ 15 Increase from baseline > 15 but ≤ 20 Increase from baseline > 20

APPENDIX F. ECG POTENTIALLY CLINICALLY SIGNIFICANT VALUES/CHANGES FROM BASELINE

Table 8: ECG Potentially Clinically Significant Values/Changes from Baseline

ECG Variable	Unit	Criteria
QTcB - Bazett's correction formula	msec	Male: ≥ 450, Female: ≥ 470 Increase from baseline > 0 but ≤ 30 Increase from baseline > 30 but < 60 Increase from baseline ≥ 60
QTcF - Fridericia's correction formula	msec	Male: ≥ 450, Female: ≥ 470 Increase from baseline > 0 but ≤ 30 Increase from baseline > 30 but < 60 Increase from baseline ≥ 60
Summary (mean) QT duration	msec	> 500