NOVARTIS

Clinical Development

Secukinumab (AIN457)

CAIN457Q12301 / NCT04181762

A two-year, phase III randomized, double-blind, parallel-group, placebo-controlled trial to evaluate the safety, efficacy, and tolerability of 300 mg s.c. secukinumab versus placebo, in combination with SoC therapy, in patients with active lupus nephritis

Statistical Analysis Plan (SAP)

Amendment 3

Author: [redacted], [redacted], Trial Statisticians

Document type: SAP Documentation

Document status: Final

Release date: 09-Oct-2023

Number of pages: 46

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
20-Oct-2020		FDA feedback on protocol	Added Kaplan-Meier analysis for time to study treatment discontinuation	2.4.3 Time to study treatment discontinuation
			Added treatment policy estimand for primary objective, added analysis of components of the estimands for primary objective.	2.3.4.2 Supportive analyses
			Added treatment policy estimand as supplementary estimand for secondary objectives with binary variables	2.6.2 Statistical hypothesis, model and method of analysis
			Added treatment policy and composite estimands as supplementary estimands for secondary objectives with continuous variables.	
			Added methodology for trimmed means	5.4.1.4 Trimmed means
15-Mar-2023	Prior to DBL of interim analysis 1	To align with protocol amendment 1	Stated SAP amendment 2 is to align with protocol amendment 1.	1
			Mentioned the total sample size is changed to 400.	1.1
			Explained the study now has two interim analyses and one final primary endpoint analysis.	
			Added the table of objectives and endpoints.	1.2
			Added the estimand framework.	
			Clarified the normal range of estimated glomerular filtration rate (eGFR).	
			Clarified the components of the primary endpoint.	2.5

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
			Amended the intercurrent events for the primary estimand and the strategies to handle them.	
			Updated the strategies to handle missing data not related to intercurrent events.	
			Revised the sensitivity analyses and the supportive analyses for the primary objective.	
			Clarified the two Week 104 analyses and mentioned they are not in the testing hierarchy.	2.6
			Updated the supportive analyses for the secondary objectives.	
			Categorized the interim analyses into the first (with approximately 138 participants) and second (with approximately 308 participants) interim analysis.	2.14
			Changed the alpha spending function from O'Brien-Fleming to Pocock for the superiority analysis in the second interim analysis.	
			Added a futility analysis in the second interim analysis.	
			Changed the total sample size from 460 to 400 based on an updated response rate of placebo.	3
			Re-calculated the powers for the primary and secondary efficacy analyses.	
			Incorporated the intercurrent events and the strategies to handle them in the definition of efficacy variables.	5.1
			Updated the definition of active urinary sediment.	

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
			Created two visit window groups specifically for eGFR and 24h UPCR.	5.3
			Amended the details of logistic regression.	5.4
			Updated reference.	6
21-Aug-2023	Prior to the final DBL of the study	The purpose of this amendment is to update the analysis plan due to the early termination of the study	Various changes to reflect the updated analysis due to early termination.	Entire Document

Table of contents

List of tables

List of figures

List of abbreviations

AE	adverse event
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
BMI	Body Mass Index
BSL	Baseline
CKD	chronic kidney disease
[redacted]	[redacted]
Cmax	Maximum concentration
Cmin	Minimum concentration
CO	Country Organization
CRF	Case Report/Record Form (paper or electronic)
CRP	C-Reactive Protein
CRR	Complete Renal Response
cs-DMARD	conventional synthetic DMARD
CSR	Clinical Study Report
CTCAE	Common Criteria for Adverse Events
Ctrough	trough concentration
CYC	Cyclophosphamide
[redacted]	[redacted]
DMARD	Disease Modifying Anti-rheumatic Drug
DMC	Data monitoring committee
DNA	DeoxyriboNucleic Acid
EC	Ethics committee
ECG	Electrocardiogram
eGFR	estimated glomerular filtration rate
ESR	Erythrocyte sedimentation rate
ESRD	end-stage renal disease
FACIT-Fatigue	Functional Assessment of Chronic Illness Therapy-Fatigue
FAS	Full Analysis Set
FDA	Food and Drug Administration
GGT	gamma glutamyl transferase
H	Hour
HDL	High Density Lipoprotein
HLA	Human Leukocyte Antigen
hsCRP	high sensitivity C-Reactive Protein
i.v.	Intravenous
[redacted]	[redacted]
IRT	Interactive Response Technology
ITT	Intent-to-Treat
LDL	Low Density Lipoprotein

LFT	Liver function test
LLN	lower limit of normal
[redacted]	[redacted]
LN	Lupus Nephritis
LOCF	Last Observation Carried Forward
LupusQoL	Lupus Nephritis Quality of Life questionnary
MAR	missing at random
MCS	Mental Component Summary
MedDRA	Medical dictionary for regulatory activities
Mg	milligram(s)
Ml	milliliter(s)
MMRM	Mixed effect Model for Repeated Measurements
MPA	mycophenolic acid
PRR	partial renal response
PCS	Physical Component Summary
PD	pharmacodynamic(s)
PK	pharmacokinetic(s)
PRN	as needed
PRO	Patient Reported Outcome
q4w	Every 4 weeks
QoL	Quality of Life
SAE	serious adverse event
SD	standard deviation
SF-36	Medical Outcome Short Form (36) Health Survey
SF-36 PCS	Short Form-36 Physical Component Summary
SLE	systemic lupus erythematosus
[redacted]	[redacted]
SoC	Standard of Care
SPP	Safety Profiling Plan
[redacted]	[redacted]
ULN	upper limit of normal
[redacted]	[redacted]
UPCR	Urine Protein-to-Creatinine Ratio
[redacted]	[redacted]
vs	Versus

1 Introduction

This Statistical Analysis Plan (SAP) is for study CAIN457Q12301, "A two-year, phase III randomized, double-blind, parallel-group, placebo-controlled trial to evaluate the safety, efficacy, and tolerability of 300 mg s.c. secukinumab versus placebo, in combination with SoC therapy, in patients with active lupus nephritis (SELUNE)".

Data was originally planned to be analyzed by Novartis according to the data analysis Section 12 of the clinical study protocol CAIN457Q12301-Protocol-V1.0 17 Feb 2023 (protocol amendment 1). However, the study has been terminated early by Novartis due to futile results from interim analysis 1. There have been no safety related reasons for early termination or concerns for the participants in this study. Due to the program termination, protocol version 1.0 dated 17-Feb-2023 was not approved globally yet resulting in a mix of approved protocol version 1.0 dated 17-Feb-2023 and protocol version 00 dated 11-Oct-2019 across countries. The purpose of this version of SAP (SAP amendment 3) is to provide details on the implementation of analyses to be reported in the abbreviated Clinical Study Report (abbreviated CSR).

Analyses based on this SAP will be executed after all patients complete their planned visits based on the early termination plan.

Data analysis will be performed by [redacted] according to this SAP.

Data analyses for Data Monitoring Committee (DMC) meetings are specified in a separate DMC SAP.

The statistical methodology is described below and any deviations from Protocol-V1.0 are documented. Additional detailed information regarding the analysis methodology is contained in the Appendix section.

1.1 Study design

This is a Phase III, randomized, double-blind, placebo-controlled trial evaluating at Week 52 the efficacy and safety of secukinumab versus placebo in subjects with active Lupus Nephritis (LN) also receiving background Standard of Care (SoC) regimen. Long-term efficacy, safety and tolerability will be collected up to 2 years.

The SoC regimen will consist of induction therapy with mycophenolic acid (MPA) (which refers to Mycophenolate mofetil (MMF) (Cellcept® or generic equivalent), or enteric-coated MPA sodium (Myfortic® or generic equivalent) at equivalent doses (oral), or Cyclophosphamide (CYC) (i.v.), followed by maintenance therapy with MPA. The choice of background SoC induction therapy will be at investigator's discretion. At Randomization, subjects will be stratified on the basis of the SoC induction therapy they will receive during the study, MPA or CYC-based, to ensure a balanced representation in each of the treatment arms (secukinumab or placebo). The target will be to have a maximum of 25% of randomized subjects receiving CYC-based induction therapy.

In addition, steroids will be administered through i.v. pulses followed by oral daily doses as described in study protocol.

The study previously included three major analyses: two interim analyses and a final primary endpoint analysis. The first interim analysis was conducted when approximately 138 participants (approximately 35% of the total 400 participants) are expected to complete 52 weeks of treatment (including those who discontinue the treatment early). This interim analysis contains a futility analysis and a PK analysis. The second interim analysis will be performed when approximately 308 participants (approximately 77% of all 400 participants) complete 52 weeks of treatment (including those who discontinue the treatment early). The results from this interim analysis will support the decision-making concerning the current clinical study, i.e., to continue or to stop the trial based on efficacy (superiority or futility) and/or safety findings. The final primary endpoint analysis will be conducted when all 400 participants complete 52 weeks of treatment (including those who discontinue the treatment early). Additional safety analyses may be performed between the second interim analysis and the Week 52 final analysis if needed to support health authority interactions following a statistically significant result at the second interim analysis.

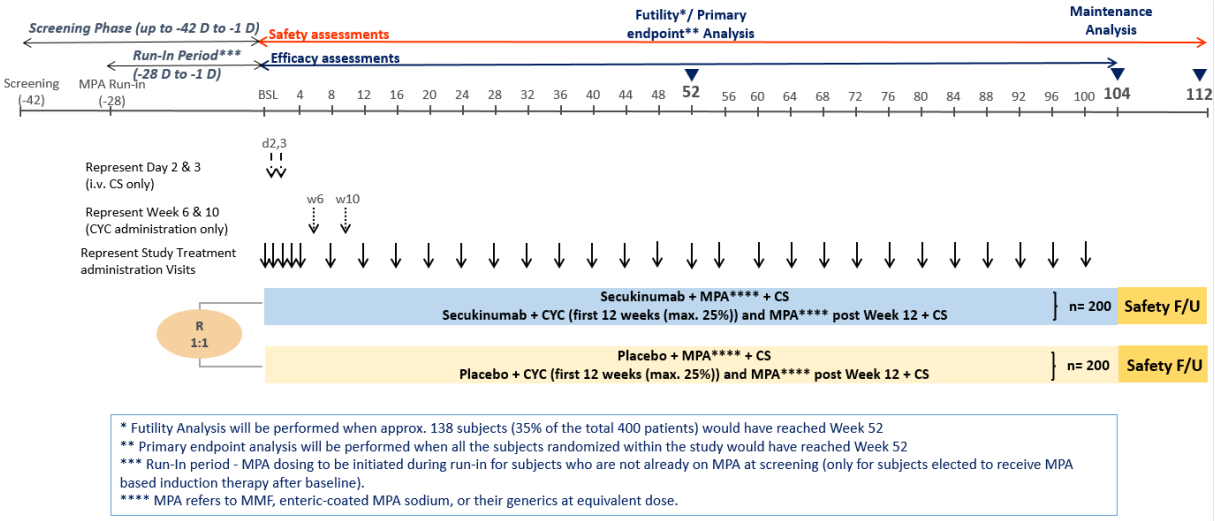
Nevertheless, since the study is terminated early due to the futile outcomes from the first interim analysis, the second interim analysis will be discarded, and the final analysis will be performed with all available patients by the time of study termination. Recruitment in this study was stopped on 26-May-2023 with 276 participants randomized. Last patient last visit (LPLV) was on 13-SEP-2023, Moreover, due to early termination of the study, only the primary efficacy analysis will be conducted with formal statistical inference. Secondary efficacy analyses will be summarized within each treatment descriptively and exploratory efficacy analyses will not be performed. An abbreviated CSR will be generated after the final database lock.

It was initially planned that unblinding of selected members of the Novartis Global Clinical Team will occur after the Week 52 database lock, original randomization of active treatment versus placebo would continue to remain blinded to all investigators, site personnel, subjects, and monitors until the final database lock and analyses are completed. Due to early termination of the trial, no unblinding of selected members of the Novartis Global Clinical Team occurred before final database lock.

The study consists of the following parts:

- Screening (up to 42 days/6 weeks)
- Run-in period (optional): For subjects who will receive MPA as SoC induction therapy as per investigator's decision and who are not already on MPA at Screening, MPA dosing will be initiated during a run-in period before Randomization (for up to 4 weeks prior to the first dose of secukinumab)
- Treatment Period: Duration of 104 weeks of treatment with secukinumab/placebo in addition to SoC treatment (with last dose given at Week 100)
- Follow-up period: Duration of 8 weeks (last visit performed 12 weeks after last dose of study medication)

Figure 1-1 Study design



Screening to Randomization (Screening and Run-in period):

A Screening period of up to 6 weeks will be used to assess subject´s eligibility and to adjust for concomitant medication(s) (-42 to -1 day). This flexible duration will provide enough time to evaluate eligibility of the subject, including renal biopsy evaluation.

Treatment period: Treatment period covers time between Randomization (Baseline) through EOT.

At Baseline, eligible subjects will be randomized in a 1:1 ratio to secukinumab 300 mg s.c. or placebo. Approximately 200 subjects were planned to be randomized to each of the two treatment arms. A blinded, weekly, s.c. secukinumab or placebo loading regimen will be administered for the first 4 weeks followed by a monthly maintenance dose in all randomized subjects thereafter. At the end of the treatment period at Week 104, the planned End of Treatment (EOT) visit will be performed.

Subjects who discontinue study treatment prematurely for any reason other than withdrawal of informed consent before Week 104 will not be considered as discontinued from the study.

Safety assessments will include physical examinations, vital signs, standard clinical laboratory evaluations, hematology (including blood coagulation assays), blood chemistry, urinalysis, adverse and serious adverse event monitoring.

1.2 Study objectives and endpoints

Table 1-1 shows the objectives and related endpoints as defined in Protocol-V01. Due to the early termination of the study: only the primary objective will be analyzed with formal statistical inference, secondary objectives will be summarized with descriptive statistics, and exploratory objectives will not be analyzed.

Table 1-1 Objectives and related endpoints

Objective(s)	Endpoint(s)
Primary objective(s)	**Endpoint(s) for primary objective(s)**
• To demonstrate that secukinumab 300 mg is superior to placebo in Complete Renal Response (CRR) rate at Week 52 in active lupus nephritis (ISN/RPS Class III or IV, with or without co-existing Class V features) patients on a background of SoC therapy.	• Proportion of participants achieving CRR at Week 52 CRR at Week 52 is an endpoint defined as meeting all of the following: • Estimated glomerular filtration rate (eGFR) ≥ 60 mL/min/1.73 m^2 or no less than 85% of Baseline • 24-hour urine protein-to-creatinine ratio (UPCR) ≤ 0.5 mg/mg • No treatment discontinuation before Week 52 • The participant did not receive more than 10 mg/day prednisone equivalent for ≥ 3 consecutive days or for ≥ 7 days in total during Week 44 through Week 52
Secondary objective(s)	**Endpoint(s) for secondary objective(s)***
• To demonstrate superiority of secukinumab compared to placebo in change from baseline in 24-hour UPCR at Week 52	• Change from Baseline in 24-hour UPCR at Week 52
• To demonstrate superiority of secukinumab compared to placebo in proportion of patients achieving partial renal response (PRR) at Week 52	• Proportion of participants achieving PRR at Week 52 defined as: • ≥ 50% reduction in 24-hour UPCR to sub-nephrotic levels (≤ 3 mg/mg) and • eGFR ≥ 60 mL/min/1.73 m2 or no less than 85% of Baseline
• To demonstrate superiority of secukinumab compared to placebo in average daily dose of oral corticosteroids administered between Week 16 and Week 52	• Average daily dose of oral corticosteroids administered between Week 16 and Week 52 compared to placebo
• To demonstrate superiority of secukinumab compared to placebo in proportion of patients achieving PRR at Week 24	• Proportion of participants achieving PRR at Week 24 defined as: • ≥ 50% reduction in 24-hour UPCR to sub-nephrotic levels (≤ 3 mg/mg) and • eGFR ≥ 60 mL/min/1.73 m2 or no less than 85% of Baseline
• To demonstrate superiority of secukinumab compared to placebo in time to achieve CRR	• Time to achieve CRR up to Week 52

Objective(s)	Endpoint(s)
• To demonstrate superiority of secukinumab compared to placebo in time to achieve PRR	• Time to achieve PRR up to Week 52
• To demonstrate superiority of secukinumab compared to placebo in time to achieve first morning void UPCR ≤ 0.5 mg/mg	• Time to achieve first morning void UPCR ≤ 0.5 mg/mg up to Week 52
• To demonstrate superiority of secukinumab compared to placebo in change in Functional Assessment of Chronic Illness Therapy – Fatigue (FACIT-Fatigue©) score at Week 52	• Improvement in FACIT-Fatigue© mean change of score from Baseline at Week 52 compared to placebo
• To demonstrate superiority of secukinumab compared to placebo in patient's health related quality of life via Medical Outcome Short Form Health Survey (SF-36 Physical Component Summary (PCS)) score at Week 52	• Improvement in SF-36 PCS mean change from Baseline at Week 52 compared to placebo
• To demonstrate superiority of secukinumab compared to placebo in change of LupusQoL (Physical Health) score at Week 52	• Improvement in LupusQoL Physical Health mean change of score from Baseline at Week 52 compared to placebo
• To evaluate the safety and tolerability of secukinumab s.c. as an add-on therapy to Standard of Care in lupus nephritis patients	• Incidence of Treatment-emergent AEs (TEAEs) / SAEs from Baseline to Week 52; vital signs and body measurements, standard chemistry and hematology up to Week 52
• To estimate the proportion of patients with maintained renal response at Week 104	• Estimate the proportion of participants with CRR at Week 104 within participants who had achieved CRR at Week 52 in the secukinumab group
• To estimate the proportion of patients with improved or maintained renal response at Week 104	• Estimate the proportion of participants with improved or maintained response (PRR or CRR) at Week 104 in participants who had achieved at least PRR at Week 52 in the secukinumab group
Exploratory objective(s)	**Endpoint(s) for exploratory objective(s)**
[redacted]	

Objective(s)	Endpoint(s)
[redacted]	

Objective(s)	Endpoint(s)
[redacted]	
• To perform exploratory pharmacogenomic analysis based on blood samples for DNA and RNA analysis (optional assessments)	• Evaluate the relationship of genetic polymorphisms and transcriptomic data with treatment response

*Intercurrent events of the secondary efficacy endpoints and the strategies to handle these are specified in Section 1.2.2.

1.2.1 Primary estimand

The **primary clinical question of interest** to be answered in the trial is:

What is the effect of subcutaneous secukinumab 300 mg compared with placebo on achieving CRR at Week 52, in patients with active LN (ISN/RPS Class III or IV, with or without co-existing class V features), on a background of SoC therapy?

CRR is defined as

1. eGFR $\geq$ 60 mL/min/1.73m^2 or no less than 85% of Baseline

AND

2. 24-hour UPCR $\leq$ 0.5 mg/mg

This question gives rise to the primary estimand, which has the following attributes with respect to the primary question:

- **Population**: Participants with active LN (ISN/RPS Class III or IV, with or without co-existing class V features), as defined by the inclusion and exclusion criteria of the study protocol amendment 1 Section 5.
- **Treatment**: The randomized treatment (the investigational therapy subcutaneous secukinumab 300 mg or the placebo), in combination with SoC. More details about the treatment are provided in protocol amendment 1 Section 6.
- **Endpoint**: An endpoint meeting all of the following:
  - CRR at Week 52;
  - If the participant discontinues the treatment (for reasons other than "study terminated by sponsor") before Week 52, then the participant will be considered as a non-responder (composite endpoint strategy);
  - If the participant has overuse of corticosteroid (> 10 mg/day prednisone equivalent for ≥ 3 consecutive days or ≥ 7 days in total) between Week 44 and Week 52, then the participant will be considered as a non-responder (composite endpoint strategy).
  - Treatment discontinuation due to "study terminated by sponsor" will not be considered as an intercurrent event.

Summary measure: Difference in marginal response proportions of achieving CRR at Week 52, between secukinumab and placebo.

1.2.2 Secondary estimands

Due to the early termination of the study, the secondary efficacy objectives previously defined under the testing hierarchy in Protocol-V01 will be summarized descriptively using observed data. Below is the list of the secondary endpoints:

1. Change from Baseline in 24-hour UPCR at Week 52
2. Proportion of patients achieving PRR at Week 52
3. Average daily dose of oral corticosteroids administered between Week 16 and Week 52 compared to placebo
4. Proportion of patients achieving PRR at Week 24
5. Time to achieve CRR up to Week 52
6. Time to achieve PRR up to Week 52
7. Time to achieve first morning void UPCR ≤ 0.5 mg/mg up to Week 52
8. Improvement in FACIT-Fatigue© mean change from Baseline at Week 52
9. Improvement in SF-36 PCS mean change from Baseline at Week 52
10. Improvement in LupusQoL Physical Health mean change from Baseline at Week 52

In addition, the following three objectives will be assessed as secondary analyses as well:

11. To evaluate the safety and tolerability of secukinumab s.c. as an add-on therapy to Standard of Care in lupus nephritis subjects
12. To estimate the proportion of subjects with maintained renal response at Week 104

13. To estimate the proportion of subjects with improved or maintained renal response at Week 104

2 Statistical methods

2.1 Data analysis general information

This SAP guides the statistical analysis for the abbreviated CSR after all patients enrolled in the study complete their planned visits based on the study's early termination plan.

Summary statistics for continuous variables will include N, mean, standard deviation, minimum, lower quartile, median, upper quartile, maximum. Summary statistics for discrete variables will be presented in the number and percent of subjects in each category. Kaplan-Meier estimates of the probability to experience an event will be calculated from baseline along with 95% confidence intervals using Greenwood's formula. The cumulative probability to flare will be plotted against time. The default level of significance will be set to 5% (two-sided, family-wise type-I-error). For group sequential testing an alpha spending function was used to ensure control of type 1 error.

Two-sided p-value will only be provided for the primary analysis. Two-sided unadjusted confidence intervals will be displayed for the primary and select secondary analyses.

Data analyses will be presented by treatment regimen. Efficacy and safety data will be presented by the 2 treatment groups: AIN457 300 mg and placebo. The 2 treatment groups represent the regimens that subjects were randomized to for the study.

Comparative efficacy data

Comparative efficacy analyses (i.e., inferential efficacy comparisons with placebo) will focus on the primary analysis. Comparative efficacy will be performed based on the Full Analysis Set 2 (FAS2) population using the randomized treatment.

Efficacy data following study treatment discontinuation

Subjects may continue performing the study visits and assessments after permanent discontinuation of study treatment. Efficacy data for inferential analyses will be handled according the estimand definitions for the primary endpoint. For summary tables efficacy data will be presented separately after permanent treatment discontinuation.

2.2 Analysis sets

The following analysis sets will be used in this trial:

The Randomized Analysis Set (RAS) consists of all randomized subjects. Regardless of whether they actually received study medication. Subject were analyzed according to the treatment assigned at randomization.

The Full Analysis Set (FAS) will be comprised of all analyzable subjects from the randomized set to whom study treatment has been assigned. According to the intent to treat principle, subjects will be analyzed according to the treatment they have been assigned to during the

randomization procedure, but according to actual stratum. Mis-randomized subjects (mis-randomized in IRT) will be excluded from the FAS.

The Full Analysis Set 2 (FAS2) will be comprised of all analyzable subjects from the randomized set to whom study treatment has been assigned and who would have had a chance to reach 52 weeks of treatment at the study termination. According to the intent to treat principle, subjects will be analyzed according to the treatment they have been assigned to during the randomization procedure, but according to actual stratum. FAS2 will be used as the primary analysis set for the primary estimand analysis.

The Safety Set includes all subjects who received at least one dose of study treatment. Subjects will be analyzed according to the study treatment received.

[redacted]

2.2.1 Treatment groups

Efficacy and safety data for Treatment Period will be presented by the following two treatment groups.

1. AIN457 300 mg
2. Placebo

2.2.2 Subgroup of interest

Will not be done for the abbreviated CSR.

2.3 Patient disposition, demographics and other baseline characteristics

2.3.1 Patient disposition

The number of subjects screened will be presented. In addition, the reasons for screen failures will be provided. The number and percentage of subjects in the randomized set who completed the study periods and who discontinued the study or treatment prematurely (including the reason for discontinuation) will be presented for each treatment group and all subjects.

For each protocol deviation (PD), the number and percentage of subjects for whom the PD applies will be tabulated.

2.3.2 Background and demographic characteristics

The following common background and demographic variables will be summarized:

Continuous variables:

- Age
- Height

- Weight
- Body mass index (BMI) = (body weight in kilograms) / (height in meters)^2

For BMI, height and body weight, the last value prior to randomization is used. If there is no weight recorded prior to administration of study drug, BMI will be missing.

Categorical variables:

- Age categories (< 30 yr, ≥30 yr)
- Gender
- Race
- Ethnicity
- Smoking status at baseline

Baseline disease characteristics will also be summarized for the following variables:

- Patient's global assessment of disease activity, UPCR, eGFR, Serum creatinine(mg/dl), C3(mg/dl), C4(mg/dl), Anti-dsDNA (geometric mean IU/ml), Standard of care induction therapy, corticosteroids (mg/day) at screening, time since first diagnosis of SLE and LN (years), and renal biopsy LN classification

Unless otherwise specified, summary statistics will be presented for continuous variables for each treatment group and for all subjects (total) in the randomized set. The number and percentage of subjects in each category will be presented for categorical variables for each treatment group and all subjects (total) in the randomized set.

Unless otherwise specified, analyses will be based on the randomized set.

2.3.3 Medical history

Any condition entered on the Relevant medical history / current medical conditions CRF will be coded using the MedDRA dictionary. They will be summarized by system organ class (SOC) and preferred term (PT) of the MedDRA dictionary. Summaries for cardiovascular medical history and SLE/LN medical history will also be provided.

To establish a baseline level of cardiovascular risk, the number and percentage of subjects with pre-solicited cardiovascular risk factors will be summarized by treatment group. The number of cardiovascular risk factors that each subject has will also be summarized by treatment group. If it is unknown whether or not a subject currently or previously experienced a specific cardiovascular risk factor, it will be assumed that cardiovascular risk factor did not occur for that subject. Unless otherwise specified, analyses will be based on the randomized set.

2.4 Treatments (study treatment, rescue medication, concomitant therapies, compliance)

2.4.1 Study treatment / compliance

The analysis of study treatment data will be based on the safety set. The number of active and placebo s.c. administrations received will be presented by treatment group. The duration of exposure to study treatment will also be summarized by treatment group. In addition, the

number and percentage of subjects with cumulative exposure levels (e.g., any exposure, ≥ 4 weeks, ≥ 8 weeks, etc.) will be presented.

Duration of exposure will be defined as the time from first dose of study treatment to the time of minimum of (last dose of the treatment + 84 days) and (last visit date). For subjects who discontinue treatment, this will be the subject's last visit in the corresponding treatment period.

Duration of exposure (years) = duration of exposure (days) / 365.25

Duration of exposure (100 subject years) = duration of exposure (years) / 100

The analyses of duration of exposure described above will be done for the entire study treatment period.

2.4.2 Prior, concomitant and post therapies

Prior medications will be presented in listing. Concomitant medications will be summarized in separate tables by treatment group.

Prior medications are defined as treatments taken and stopped prior to first dose of study treatment. Any medication given at least once between the day of first dose of randomized study treatment and within 84 days after last dose will be a concomitant medication, including those which were started pre-baseline and continued into the period where study treatment is administered.

Medications will be presented in alphabetical order, by Anatomical Therapeutic Classification (ATC) codes and grouped by anatomical main group (the 1st level of the ATC codes). Tables will show the overall number and percentage of subjects receiving at least one treatment of a particular ATC code and at least one treatment in a particular anatomical main group.

Significant concomitant non-drug therapies and procedures will be summarized by primary system organ class and MedDRA preferred term.

Any non-drug therapies and procedures done between the day of first dose of study treatment and within 84 days after last dose will be defined as concomitant non-drug therapies and procedures, including those which were started pre-baseline and continued into the period where study treatment is administered.

Concomitant medication will be identified by comparing recorded or imputed start and end dates of medication taken to the reference start date.

2.4.3 Time to study treatment discontinuation

Will not be done for the abbreviated CSR.

2.5 Analysis of the primary objective

2.5.1 Primary endpoint

The primary efficacy endpoint is the complete renal response (CRR) at Week 52 with the consideration of intercurrent events, which is defined in detail at Section 1.2.1.

2.5.2 Statistical hypothesis, model, and method of analysis

The statistical hypothesis tested for the primary objective is that there is "no difference in the proportion" of patients fulfilling the response criteria at Week 52 between the secukinumab and placebo regimens.

Let p_j denote the proportion of responders at Week 52 for treatment regimens j, j=0, 1 where

- 0 corresponds to placebo regimen,
- 1 corresponds to secukinumab,

In statistical terms, H_1: $p_1 = p_0$, H_{A1}: $p_1 \neq p_0$, i.e.,

H_1: secukinumab is not different from placebo regimen with respect to CRR at Week 52.

Logistic regression model adjusting for SoC, race (White/Non-White) and Baseline First Morning Void UPCR will be used for the primary analysis. Difference in marginal response proportions with p-value and respective 95% confidence interval will be estimated from the logistic regression model (see Section 5.4.2.2).

2.5.3 Handling of intercurrent events of primary estimand

Major intercurrent events of primary estimand will be addressed with the following strategies (considering that these events occur before the assessments at Week 52):

1. Treatment discontinuation (for reasons other than "study terminated by sponsor") for any reason: non-responder (composite endpoint strategy)
2. Overuse of corticosteroid (> 10mg/day prednisone equivalent for ≥ 3 consecutive days or ≥ 7 days in total) between Week 44 and Week 52: non-responder (composite endpoint strategy)

Note that treatment discontinuation due to "study terminated by sponsor" is not considered as an intercurrent event for the primary estimand.

2.5.4 Handling of missing values not related to intercurrent events

The primary estimand analysis will be performed on FAS2. Participants in FAS2 who do not have the required data to compute CRR at Week 52 will be classified as non-responders.

2.5.5 Sensitivity analyses

Will not be done for the abbreviated CSR.

2.5.6 Supplementary analyses

A supplementary analysis of the primary endpoint will be summarizing CRR descriptively with observed data without considering any intercurrent events. This analysis will be performed using FAS population.

2.6 Analysis of secondary objectives

2.6.1 Secondary efficacy endpoints

The secondary efficacy variables and the method for adjusting for multiplicity are described below. However, due to the early termination of the study, the ten listed secondary efficacy

analyses above will no longer be in the testing hierarchy and will be summarized with descriptive statistics using observed data.

H_2: Secukinumab 300mg is not different to placebo with respect to change from Baseline in 24-hour UPCR at Week 52 compared to Baseline

H_3: Secukinumab 300mg is not different to placebo with respect to proportion of patients achieving PRR at Week 52

H_4: Secukinumab 300mg is not different to placebo with respect to average daily dose of oral corticosteroids administered between Week 16 and Week 52

H_5: Secukinumab 300mg is not different to placebo with respect to proportion of patients achieving PRR at Week 24

H_6: Secukinumab 300mg is not different to placebo with respect to time to achieve CRR,

H_7: Secukinumab 300mg is not different to placebo with respect to time to achieve PRR

H_8: Secukinumab 300mg is not different to placebo with respect to time to achieve first morning void UPCR $\leq$ 0.5 mg/mg

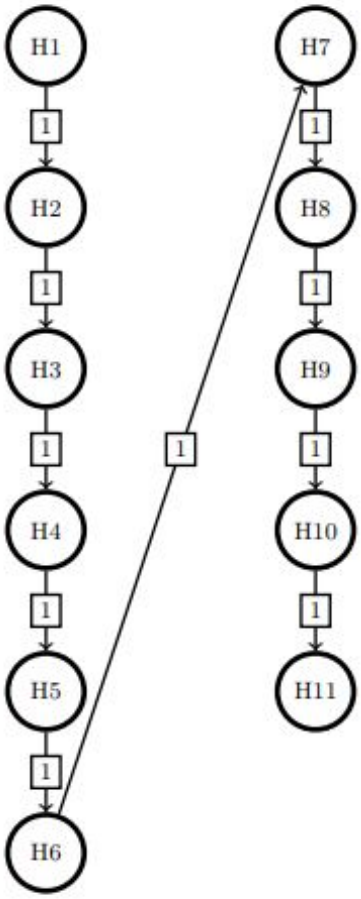
H_9: Secukinumab 300mg is not different to placebo with respect to FACIT-Fatigue©, mean change of score from Baseline at Week 52

H_{10}: Secukinumab 300mg is not different to placebo with respect to improvement in SF-36 PCS mean change from Baseline at Week 52

H_{11}: Secukinumab 300mg is not different to placebo with respect to improvement in LupusQoL physical health score mean change from Baseline at Week 52

A sequential testing strategy in combination with the group sequential testing was planned in order to control for multiplicity of testing. The graphical approach of (Bretz et al 2009) for sequentially rejective testing procedures was planned for the testing strategy:

Figure 2-1 Testing strategy



Moreover, there are two additional secondary efficacy endpoints not included in the previous testing hierarchy that will also be analyzed descriptively:

- Estimate the proportion of subjects with CRR at Week 104 within subjects who had achieved CRR at Week 52 in the secukinumab group
- Estimate the proportion of subjects with improved or maintained response (PRR or CRR) at Week 104 in subjects who had achieved at least PRR at Week 52 in the secukinumab group

Secondary efficacy variables will be analyzed using the FAS population:

- UPCR at Week 52
- PRR at Week 24 or at Week 52
- Average daily dose of oral corticosteroids administered between Week 16 and Week 52
- Time to achieve CRR: Kaplan-Meier curves will be presented for each treatment.
- Time to achieve PRR: Kaplan-Meier curves will be presented for each treatment.

- Time to achieve first morning void UPCR ≤ 0.5 mg/mg: Kaplan-Meier curves will be presented for each treatment.
- FACIT-Fatigue© change from Baseline at Week 52
- SF-36 PCS change from Baseline at Week 52
- LupusQoL physical health score change from Baseline at Week 52
- Maintained renal response at Week 104: A patient with maintained renal response at Week 104 is one who achieves CRR at Week 104 after already achieving CRR at Week 52. Response rate with 95% confidence interval will be presented for the secukinumab group.
- Improved or maintained renal response at Week 104: A patient with improved or maintained renal response at Week 104 is one who achieves CRR or PRR at Week 104 after achieving at least PRR at Week 52. Response rate with 95% confidence interval will be presented for the secukinumab group.

2.6.2 Handling of intercurrent events of secondary estimands

Not applicable.

2.6.3 Handling of missing values not related to intercurrent events

Binary, continuous, and time to event variables will be analyzed using observed data.

2.6.4 Sensitivity analyses

Will not be done for the abbreviated CSR.

2.6.5 Supplementary analyses

Will not be done for the abbreviated CSR.

2.7 Analysis of exploratory objectives

Exploratory analyses will not be performed for the abbreviated CSR due to the early termination of the study.

2.8 Safety analyses

Summary will be performed for the entire treatment period. The analyses of the follow-up period will be limited to summaries for treatment-emergent adverse events, serious adverse events and risks based on adverse events.

Safety analyses will be performed on treatment received or actual treatment as described below:

The actual treatment or treatment received for summaries of safety data will differ to the treatment assigned at randomization only if a subject received the wrong treatment during the entire study.

2.8.1 Adverse events (AEs)

The crude incidence of treatment emergent adverse events (i.e., events started after the first dose of study treatment or events present prior to the first dose of study treatment but increased in severity based on preferred term and on or before last dose date + 84 days) will be summarized by primary system organ class and preferred term. Confidence intervals for the crude rate will be derived as described in Section 5.4.5. In addition, exposure time-adjusted incidence rates including 95% confidence intervals will be provided for the entire treatment period (see Section 5.4.6).

Adverse events reported will be presented in descending frequency according to its incidence in the secukinumab group starting from the most common event. Summaries (crude incidences only) will also be presented for AEs by severity and for study treatment related AEs. If a particular AE 'severity' is missing, this variable will be listed as missing and treated as missing in summaries. If a subject reported more than one adverse event with the same preferred term, the adverse event with the greatest severity will be presented. If a subject reported more than one adverse event within the same primary system organ class, the subject will be counted only once with the greatest severity at the system organ class level, where applicable.

Separate summaries will be provided for adverse events suspected to be related to study drug, deaths, serious adverse events, and adverse events leading to discontinuation and adverse events requiring concomitant medication.

The crude incidence rate will be provided without treatment information. For SAEs occurred during screening a listing will be prepared for all subjects screened including screening failures.

An overview of the safety analyses which will be performed for treatment emergent AEs, labs and vital signs for each analysis period is described in Table 2-1.

Table 2-1 Overview of analyses on some safety endpoints

Analysis period	AEs & SAEs & Safety topics of interest AEs	AEs by severity	Study drug related AEs	Notables for (vitals), lab criteria
Entire Treatment	• crude incidence • exposure time adjusted incidence	• crude incidence	• crude incidence	• crude incidence

Safety topics of interest included as the safety risks as defined in the Risk Management Plan (RMP) as well as those that are not listed in the RMP but are deemed important and relevant; these topics are also defined in the Program Case Retrieval Sheet.

The crude incidence and exposure-adjusted incidence rates for the safety topics of interest AEs will be summarized. In addition, separate listings will be provided for SAEs..

Algorithms for date imputations will be provided in Programming Datasets Specifications.

2.8.2 Laboratory data

The summary of lab data will only include treatment emergent data, which are defined as those lab assessments after the first dose of study treatment and on or before last dose + 84 days.

Reported laboratory assessments with either a less than or greater than sign ("<" or ">") will be used for analysis after removal of the sign and conversion to standard unit. These laboratory data will be displayed in listings using the standard unit with the reported sign ("<" or ">").

The summary of laboratory evaluations will be presented for three groups of laboratory tests (hematology, chemistry and urinalysis).

For urinalysis, frequency table will be presented.

Descriptive summary statistics for the change from baseline to each study visit will be presented. These descriptive summaries will be presented by laboratory test and treatment group. Change from baseline will only be summarized for subjects with both baseline and post baseline values and will be calculated as:

change from baseline = post baseline value – baseline value

The following laboratory parameters will be analyzed with respect to numerical Common Terminology Criteria for Adverse Events (CTCAE) grades, given in Table 2-2: hemoglobin, platelets, white blood cell count, neutrophils, lymphocytes, creatinine, total bilirubin (TBL), gamma-glutamyl transferase (GGT), alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), glucose, cholesterol, triglycerides (TG).

These summaries will be split into hematology and chemistry.

Table 2-2 CTCAE grades for laboratory parameters to be analyzed

CTCAE v4.0 Term	Grade 1	Grade 2	Grade 3	Grade 4
HGB decreased (Anemia)	<LLN – 100 g/L	<100 – 80 g/L	<80 g/L	See note below
Platelet count decreased	<LLN – 75.0 x10e9 /L	<75.0 - 50.0 x10e9 /L	<50.0 – 25.0 x10e9 /L	<25.0 x 10e9 /L
White blood cell decreased	<LLN - 3.0 x 10e9 /L	<3.0 - 2.0 x 10e9 /L	<2.0 - 1.0 x 10e9 /L	<1.0 x 10e9 /L
Neutrophil count decreased	<LLN - 1.5 x 10e9 /L	<1.5 - 1.0 x 10e9 /L	<1.0 - 0.5 x 10e9 /L	<0.5 x 10e9 /L
Lymphocyte count decreased	<LLN - 0.8 x 10e9/L	<0.8 - 0.5 x 10e9 /L	<0.5 - 0.2 x 10e9 /L	<0.2 x 10e9 /L
Creatinine increased*	>1 - 1.5 x baseline; >ULN - 1.5 x ULN	>1.5 - 3.0 x baseline; >1.5 - 3.0 x ULN	>3.0 baseline; >3.0 - 6.0 x ULN	>6.0 x ULN
TBL increased	>ULN - 1.5 x ULN	>1.5 - 3.0 x ULN	>3.0 - 10.0 x ULN	>10.0 x ULN
GGT increased	>ULN - 2.5 x ULN	>2.5 - 5.0 x ULN	>5.0 - 20.0 x ULN	>20.0 x ULN
ALT increased	>ULN - 3.0 x ULN	>3.0 - 5.0 x ULN	>5.0 - 20.0 x ULN	>20.0 x ULN
AST increased	>ULN - 3.0 x ULN	>3.0 - 5.0 x ULN	>5.0 - 20.0 x ULN	>20.0 x ULN
ALP increased	>ULN - 2.5 x ULN	>2.5 - 5.0 x ULN	>5.0 - 20.0 x ULN	>20.0 x ULN
Glucose increased (Hyperglycemia)	; >ULN - 8.9 mmol/L	>8.9 - 13.9 mmol/L	>13.9 - 27.8 mmol/L	>27.8 mmol/L
Glucose decreased (Hypoglycemia)	<LLN - 3.0 mmol/L	<3.0 - 2.2 mmol/L	<2.2 - 1.7 mmol/L	<1.7 mmol/L
Cholesterol high	>ULN - 7.75 mmol/L	>7.75 - 10.34 mmol/L	>10.34 - 12.92 mmol/L	>12.92 mmol/L
Hypertriglyceridemia	1.71 - 3.42 mmol/L	>3.42 - 5.7mmol/L	>5.7 - 11.4 mmol/L	>11.4 mmol/L

Note: Grade 4 Hemoglobin events are defined as life-threatening anemia events and will not be displayed in the table, as a numerical range is not provided in the CTCAE.
*Note: for “creatinine increased” the baseline criteria do not apply.

Shift tables will be presented comparing baseline laboratory result (CTCAE grade) with the worst results (expressed in CTCAE grade) during the treatment phase (either initial or entire) analyzed. Of note, baseline will be defined as last assessment prior to first dosing in initial treatment phase. Subjects with abnormal laboratory values will be listed and values outside the normal ranges will be flagged.

Summaries for newly occurring or worsening clinically notable lipid abnormalities will also be provided cumulatively for each of the following parameters and categories:

- HDL:
  - <=LLN
  - <0.8 x LLN
- LDL, cholesterol, triglycerides:
  - >=ULN
  - >1.5 x ULN
  - >2.5 x ULN

Newly occurring or worsening liver enzyme abnormalities will also be summarized based on the event criteria given in Table 2-3:

Table 2-3 Liver-related events

Parameter	Criterion
ALT	>3xULN; >5xULN; >8xULN;>10xULN, >20xULN
AST	>3xULN; >5xULN; >8xULN >10xULN; >20xULN
ALT or AST	>3xULN; >5xULN; >8xULN >10xULN; >20xULN
TBL	>1.5xULN, >2xULN, >3xULN,
ALP	>2xULN, >3xULN. >5xULN
ALT or AST & TBL	ALT or AST>3xULN & TBL >2xULN; ALT or AST >5xULN & TBL >2xULN; ALT or AST >8xULN & TBL >2xULN; ALT or AST >10xULN & TBL >2xULN
ALP & TBL	ALP >3xULN & TBL >2xULN ALP >5xULN & TBL >2xULN
ALT or AST & TBL & ALP	ALT or AST>3xULN & TBL >2xULN & ALP <2xULN **(Hy's Law**) Note: elevated ALP may suggest obstruction as a consequence of gall bladder or bile duct disease; ALP may also be increased in malignancy. FDA therefore terms Hy's Law cases as indicators of *pure hepatocellular injury*. This does not mean that cases of ALT or AST >3xULN & TBL >2xULN & ALP ≥2xULN may not result in severe DILI.

Notes:
In studies which enroll subjects with pre-existing liver disease, baseline LFT may be increased above ULN; in such a case it is meaningful to add the condition "and worse than baseline" to the abnormality criteria

For a combined criterion to be fulfilled, all conditions have to be fulfilled on the same visit. The criteria are not mutually exclusive, e.g. a subject with ALT = 6.42xULN is counted for ALT >3xULN and ALT>5x ULN.

Individual subject data listings will be provided for subjects with abnormal laboratory data. Data of subjects with newly occurring or worsening liver enzyme abnormalities will be listed in an additional listing.

Boxplots over time will be presented for selected laboratory parameters (neutrophils, liver and lipid parameters).

2.8.3 Other safety data

2.8.3.1 Vital signs

The summary of vital signs will only include treatment emergent data, which are defined as those vital sign measurements after the first dose of study treatment and on or before last dose + 84 days.

Analysis in vital sign measurement using descriptive summary statistics for the change from baseline for each post-baseline visit will be performed. These descriptive summaries will be presented by vital sign and treatment group. Change from baseline will only be summarized for subjects with both baseline and post-baseline values and will be calculated as:

change from baseline = post-baseline value – baseline value

The number and percentage of subjects with newly occurring notable vital signs will be presented. Criteria for notable vital sign abnormalities are provided in Table 2-4:

Table 2-4 Criteria for notable vital sign abnormalities

Vital sign (unit)	Notable abnormalities
Systolic blood pressure (mmHg)	>= 140 mmHg or < 90 mmHg
Diastolic blood pressure (mmHg)	>=90 mmHg or <60 mmHg
Pulse (bpm)	> 100 bpm or <60 bpm

2.11 Patient reported outcomes (PROs)

Health-related Quality of Life assessments will be evaluated based on FAS unless otherwise specified.

SF-36

The following variables will be evaluated:

- SF-36 domain scores (based on a scale of 0-100)
- SF-36 PCS and MCS scores (norm-based scores)
- SF-36 PCS and MCS responder (improvement of ≥ 2.5 points, (Lubeck 2004))

For the change in SF-36 summary scores (PCS and MCS), summary statistics will be provided using observed data for each treatment regimen.

The SF-36 domain scores will be summarized by treatment.

FACIT-Fatigue

For the change in FACIT-Fatigue scores, summary statistics of observed data by visit and change from Baseline in FACIT will be provided for each treatment.

LupusQoL

For the change in LupusQoL domain scores, summary statistics of observed data by visit and change from Baseline will be provided for each treatment.

2.12 Biomarkers

The analysis of biomarker data will be data-driven and, hence, not part of the Clinical Study Report (CSR).

2.13 Other Exploratory analyses

Not applicable.

2.14 Interim analysis

Two interim analyses were planned for the study. However, since the futility analysis from the first interim analysis shows futile outcomes, the study is terminated early and the second interim analysis will not be conducted.

2.14.1 First interim analysis

The first interim analysis will be performed when approximately 138 participants (approximately 35% of the total 400 participants) are expected to complete 52 weeks of treatment (including those who discontinue the treatment early).

Futility Analysis

A futility analysis will be conducted at the first interim analysis. A Go/No-Go decision will be taken at this futility analysis based on the predictive probability of achieving statistical significance for the primary estimand. Futility stopping rules will be defined in the DMC charter and DMC SAP.

[redacted]

2.14.2 Second interim analysis

Based on the group sequential design applied in the study, the second interim analysis is planned when approximately 308 participants (approximately 77 % of the total 400 participants) complete 52 weeks of treatment (including those who discontinue the treatment early). The results from this interim analysis will support the decision-making concerning the current clinical study, i.e., to continue or to stop the trial based on efficacy (superiority or futility) and/or safety findings.

Superiority analysis

A Lan-De Mets alpha spending function (Lan and DeMets 1983) with Pocock type stopping boundary (Pocock 1977) (as implemented in the software East 6.5) will be used to maintain the overall type-I error rate for the primary and secondary endpoints at Week 52.

Based on the choice of α-spending function described above, the efficacy boundary in terms of p-value scale at the superoprity analysis is calculated as p=0.042 for a two-sided test. The observed (i.e., nominal) p-value has to be smaller than 0.042 to conclude superior efficacy at the superiority interim analysis. If a hypothesis is not rejected at the superiority interim analysis, it will be tested again at the Week 52 final analysis. The efficacy boundary for the final analysis in terms of p-value scale is p=0.024.

The exact rejecting boundaries will be calculated after the exact number of participants in each treatment arm is available.

A sequential testing hierarchy will be used to test the secondary hypotheses. The secondary hypotheses will be tested at the superiority analysis only if the primary hypothesis is rejected. This guarantees the 5% overall level of significance for the primary and secondary hypotheses (Glimm 2010). For Week 52 secondary endpoints the same rejecting boundaries will be used as for the primary endpoint.

Futility analysis

A second futility analysis will also be conducted at the time of the second interim analysis. Futility stopping rules will be specified in the DMC charter.

3 Sample size calculation

3.1.1 Primary endpoint(s)

The total planned sample size is 400 participants. Based on literature (Mysler et al 2013), (Rovin et al 2012), (Furie et al 2020), (Rovin et al 2021), the control response rate was assumed to be approximately 25%. Assuming 15% treatment difference with 40% response rate for secukinumab, the power for rejecting the null hypothesis for primary endpoint (CRR) is 87%. The assumptions are summarized in Table 3-1 below.

Table 3-1 Power for primary endpoint

Analysis	Control response rate	Secukinumab response rate	Sample size	Cumulative power for primary endpoint
Superiority Interim analysis	25%	40%	308	79%
Final analysis	25%	40%	400	87%

3.1.2 Secondary endpoint(s)

Using published data for control (Mysler et al 2013), (Rovin et al 2012), (Deng et al 2018) (Wallace et al 2017), (McElhone et al 2016) and assuming given treatment effects for secukinumab, a summary for power for secondary efficacy parameters is shown in Table 3-2 for binary endpoints, Table 3-3 for continuous endpoints and Table 3-4 for time to event endpoints. Power for the secondary endpoints was calculated with the group sequential design and the total cumulative power, to reject either at the superiority interim or final analysis, is presented.

Table 3-2 Summary of power for binary secondary endpoints

Endpoint	Response rate		Power
	Secukinumab 300 mg (N = 200)	Placebo (N = 200)	
PRR at Week 52	75%	60%	87%
PRR at Week 24	75%	60%	87%

Table 3-3 Summary of power or continuous endpoints

Endpoint	Mean values		Common standard deviation	Power
	Secukinumab 300 mg (N = 200)	Placebo (N = 200)		
Change from baseline in UPCR at Week 52	-3.5	-2.68	2.69	83%
Average daily dose of corticosteroids between Week 16 and Week 52	5	5.52	1.67	84%
FACIT-Fatigue© change from baseline at Week 52	7.0	2.82	10.0	98%

Endpoint	Mean values		Common standard	Power
SF-36 PCS change from baseline at Week 52	6.1	3.1	8.0	95%
LupusQoL Physical Health change from baseline at Week 52	9.6	5.6	14	78%

Table 3-4 Summary of power for time to event endpoints

Endpoint	Event rates and hazard ratios		Power
	Placebo (N = 200) Event rate at Week 52	Hazard Ratio Secukinumab (N = 200)	
Time to CRR	30%	1.67	85%
Time to PRR	60%	1.67	98%
Time to UPCR ≤ 0.5 mg/mg	35%	1.67	89%

4 Change to protocol specified analyses

The following significant changes from Protocol-V01 are implemented in this SAP:

1. Terminate the study early due to the futile outcomes from the first interim analysis.
2. Change the Clinical Study Report type from full to abbreviated.
3. Remove the second interim analysis.
4. Indicate the CRO [redacted] will carry out the final analysis.
5. Exclude treatment discontinuation due to "study terminated by sponsor" as an intercurrent event for the primary objective.
6. Update FAS2 as patients with chance to reach Week 52 and use it as the analysis set for the primary analysis in the final analysis.
7. Remove subgroup analyses for the primary and secondary endpoints.
8. Remove the sensitivity analyses for the primary objective.
9. Only have the observed "no intercurrent event" analysis for the supplementary analysis of the primary objective.
10. Change hierarchical secondary efficacy analyses from formal comparisons to descriptive summaries.
11. Remove the intercurrent events and their strategies for the secondary efficacy objectives. Secondary efficacy objectives will be summarized with observed data.

12. Remove the sensitivity and supplementary analyses for the secondary objectives.
13. Remove all exploratory analyses.
14. Remove the initial period (Week 1 - 52) and keep the entire treatment period safety analyses.
15. Remove some safety analyses.
16. Add safety topics of interest AEs for safety analyses.
[redacted].
18. Indicate biomarker analyses are not part of the CSR.

5 Appendix

5.1 Description of additional efficacy variables

Estimated glomerular filtration rate (eGFR)

The glomerular filtration rate will be estimated using the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation (Martínez-Martínez et al 2013) based on subject gender, age (years) and serum creatinine (mg/dL).

Central laboratory serum creatinine values will be used for all renal function data analysis.

5.2 Description of health-related quality of life variables

SF-36

The Short Form Health Survey (SF-36) is a widely used and extensively studied instrument to measure health-related quality of life among healthy subjects and patients with acute and chronic conditions. It consists of eight subscales (domains) that can be scored individually: Physical Functioning, Role-Physical, Bodily Pain, General Health, Vitality, Social Functioning, Role- Emotional, and Mental Health. Two overall summary scores, the Physical Component Summary (PCS) and the Mental Component Summary (MCS) also can be computed. The SF-36 has proven useful in monitoring general and specific populations, comparing the relative burden of different disease, differentiating the health benefits produced by different treatments, and in screening individual patients. The eight domains are based on a scale from 0-100 while PCS and MCS are norm-based scores with a mean of 50 and a standard deviation of 10.

Quality metric uses weighted maximum likelihood estimation, a modified version of item response theory (IRT) to estimate scale scores when a respondent is missing multiple items. The PCS summary score measure requires scores for seven scales, one of which must be the PF scale and the MCS score also requires scores for seven scales, one of which must be the MH scale. Only one item is needed for each of the multi-item domains.

FACIT-Fatigue

The FACIT-Fatigue© is a 13-item questionnaire that assesses self-reported fatigue and its impact upon daily activities and function. The purpose of FACIT-Fatigue in this study is to

assess the impact of treatment on fatigue and daily activities and function.in patients with axSpA.

Subjects respond to each item on a 5-point Likert-type scale (0 = not at all; 1 = a little bit; 2 = somewhat; 3 = quite a bit; 4 = very much) based on their experience of fatigue during the past 2 weeks. The scale score is computed by summing the item scores, after reversing those items that are worded in the negative direction. Numbering the questions from 1 to 13, it is evident that questions 7 and 8 are worded in the positive direction (4 indicates a desirable response) and all other questions are worded in the negative directions (4 indicates an undesirable response). Thus, it is necessary to reverse the responses for questions 7 and 8 (i.e. original response of 0 gets mapped to 4, 1=3, 2=2, 3=1, and 4=0) for scoring purposes.

When there are missing item scores, the scale score was computed by summing the non-missing item scores, multiplying by 13 (the total number of items in the scale) and dividing by the number of non-missing items (i.e. normalizing the results). The latter rule applied only when at least half of the items (seven or more) are non-missing.

FACIT Fatigue scale score range from 0 to 52, where higher scores represent less fatigue (Cella D et al. 1993).

Lupus Quality of Life (LupusQoL)

The LupusQoL is a disease-specific, 34-item, self-report questionnaire designed to measure the health-related quality of life (HRQoL) of subjects with SLE within 8 domains (i.e., physical health (8 items), emotional health (6 items), body image (5 items), pain (3 items), planning (3 items), fatigue (4 items), intimate relationships (2 items), and burden to others (3 items). Responses are based on a 5-point Likert scale where 0 (all of the time) to 4 (never) (Yazdany 2011), (RWS Life).

Each domain of the LupusQoL is scored separately

5.3 Visit Windows

Baseline and post-baseline definitions

In general, a *baseline* value refers to the last measurement made prior to administration of the first dose of study treatment. A *post-baseline* value refers to a measurement taken after the first dose of study treatment.

Analysis visit windows

Analysis visit windows will be used for the data that is summarized by visit; they are based on the study visit and evaluation schedule and comprise a set of days around the nominal visit day. For any assessment, there are protocol defined scheduled visits around which analysis visit windows were created to cover the complete range of days within the study.

When visit windows are used, all visits will be re-aligned, i.e., they will be mapped into one of the visit windows. E.g., if the *Week 4* visit of a subject is delayed and occurs on Day 46 instead of on Day 29, it will be re-aligned to visit window *Week 8*. In the case of major deviations from the visit schedule, or due to unscheduled visits, several assessments of a subject may fall in a

particular visit window (either scheduled or unscheduled). Statistical approaches to handle multiple assessments in a given visit window are specified below.

For lab/vital signs, follow-up (F/U) visit is excluded from analysis visit mapping window. Only assessments that come as F/U nominal visit will be directly assigned as analysis F/U visit. Other assessments that are beyond the last on-treatment visit window (W104) or after nominal F/U visit date won't be mapped to any analysis visit. F/U visit will not be included in the summary tables by visit.

Of note, subjects are allowed to have gaps in visits.

Table 5-1 Analysis visit windows

Analysis Visit		Target Day	Analysis Visit Window	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10	Group 11
Baseline		1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1
Week 1		8	2-11	2-11									2-11	
Week 2		15	12-18	12-18									12-18	
Week 3		22	19-25	19-25									19-25	
Week 4		29	26-43	26-43					2-57			2-57	26-43	
Week 8		57	44-71	44-71	2-71	2-71							44-64	
Week 12		85	72-99	72-99	72-127	72-127	2-127		58-99	2-127	2-127	58-127	65-106	65-148
Week 16		113	100-127	100-127					100-141				107-127	
Week 20		141	128-155	128-155									128-148	
Week 24		169	156-183	156-183	128-267	128-211	128-211	2-267	142-211	128-267	128-267	128-267	149-190	149-232
Week 28		197	184-211	184-211									191-211	
Week 32		225	212-239	212-239									212-232	
Week 36		253	240-267	240-267		212-309	212-309		212-309				233-274	233-344
Week 40		281	268-295	268-295									275-295	
Week 44		309	296-323	296-323									296-323	
Week 48		337	324-351	324-351									324-344	
Week 52		365	352-379	352-379	268-449	310-449	310-449	268-547	310-547	310-547	268-449	268-449	345-386	345-512
Week 56		393	380-407	380-407									387-407	
Week 60		421	408-435	408-435									408-435	
Week 64		449	436-463	436-463									436-463	
Week 68		477	464-491	464-491									464-491	

Analysis Visit		Target Day	Analysis Visit Window	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10	Group 11
Week 72		505	492-519	492-519									492-512	
Week 76		533	520-547	520-547	450-631	450-631	450-631				450-631	450-631	513-554	513-708
Week 80		561	548-575	548-575									555-575	
Week 84		589	576-603	576-603									576-603	
Week 88		617	604-631	604-631									604-631	
Week 92		645	632-659	632-659									632-659	
Week 96		673	660-687	660-687									660-687	
Week 100		701	688-715	688-715									688-708	
Week 104		729	≥716	≥716	≥632	≥632	≥632	≥548	≥548	≥548	≥632	≥632	≥709	≥709
	Group 1: Physical Examination, Body Weight, Hematology, Clinical Chemistry (exclude eGFR), Coagulation Panel, Urinalysis,Vital Signs Group 2: Lipids Group 3: hsCRP, ESR Group 4: Anti-dsDNA antibody and ANA, [redacted] FACIT-Fatigue, SF-36, LupusQoL [redacted] Group 8: Serum collection for optional biomarkers Group 9: Urine collection for optional biomarkers Group 10: eGFR Group 11: 24-hr urine collection If the participant has received treatment from the extension study, then the assessment values will not be included in the visits of this core study.													

The mapping described above applies to all visits (not just scheduled visits). Repeat and/or unscheduled visits (which will be numbered in the database according to new NCDS standards) will be mapped for analysis purposes in the same way as scheduled visits. This leaves the possibility, then, for multiple measurements within an analysis window. The following conventions will be used to determine the appropriate measurement to be summarized in the event of multiple measurements within a visit window.

Table 5-2 Rules for flagging variables

Timing of measurement	Type of data	Rule
Baseline	All data	The last measurement made prior to administration of the first dose of study treatment – note this may include measurements taken on the day of randomization (e.g., lab). Baseline assessments scheduled for and captured on Day 1 will be considered baseline measurements regardless of the time of assessment. If a patient did not receive any dose of study treatment then the randomization date will be used. Only the date part will be considered if there is only one assessment on Day 1 but if there are multiple assessments on Day 1, then the following rules will apply: 1. If time of assessment exist, • select the last available measurement prior to the reference start date/time considering time • if no measurement prior to the reference start date/time then considering time select the earliest measurement post reference start date/time 2. If time of assessment does not exist the measurement from the lowest CRF visit number will be used.
Post-baseline efficacy	All data	The measurement closest to the target day will be used. In the event two measurements are taken equally apart (e.g., 1 day before target date and 1 day after) the first one will be used. Cases where the same parameter is recorded more than once on the same date will be handled as follows: • If time of completion exists the earliest measurement will be use • If time does not exist the measurement from the lowest CRF visit number will be used
Post-baseline safety	Summary visit information (e.g., lab, etc.)	The measurement closest to the target day will be used. In the event two measurements are taken equally apart (e.g., 1 day before target date and 1 day after) the first one will be used. Cases where the same parameter is recorded more than once on the same date will be handled as follows: • If time of completion exists the earliest measurement will be used • If time does not exist the measurement from the lowest CRF visit number will be used • If CRF visit number is the same the average value will be used
Post-baseline safety	Notable abnormalities (e.g., lab)	The most extreme measurement in the window will be used. Note this means a patient can have a notably high and notably low measurement within a window

5.4 Statistical methodology and assumptions

5.4.1 Analysis of continuous data

5.4.1.1 Summary statistics for continuous data

Summary statistics (including N, mean, standard deviation, minimum, lower quartile, median, upper quartile, maximum) will be provided for continuous data by visit and treatment group.

5.4.2 Analysis of binary and categorical data

5.4.2.1 Summary statistics for binary and categorical data

Summary statistics for discrete variables will be presented in contingency tables and will include count and frequency in each category. If applicable, confidence intervals will be derived as well based on the score method including continuity correction (Newcombe 1998):

With z as (1-α/2)-quantile of the standard normal distribution (SAS: z = probit(1-α/2), n as total number of subjects (i.e. number of subjects in the denominator), p as estimated crude incidence (number of subjects with event / n) and q = 1-p

Then the lower limit is

$$L = 100 \times \max\left(0, \frac{2np + z^2 - 1 - z\sqrt{z^2 - 2 - 1/n + 4p(nq+1)}}{2(n+z^2)}\right)$$

and the upper limit is

$$U = 100 \times \min\left(1, \frac{2np + z^2 + 1 + z\sqrt{z^2 + 2 - 1/n + 4p(nq-1)}}{2(n+z^2)}\right).$$

In addition, if L > p then L = p and if U < p then U = p.

For binary response variables the placebo-adjusted response rates including 95% confidence interval will be derived.

SAS code for risk difference:

```
proc freq data=;
    tables response* treatment / riskdiff;
run;
```

Note the response value should be sorted with '1' ahead of '0'.

5.4.2.2 Logistic regression

Certain binary outcome variables, e.g., response outcomes, will be evaluated using a logistic regression model with treatment, stratification factor (SoC), and race (Black vs. Non-Black) as factors and baseline score (if applicable. e.g. baseline UPCR) as covariates. The marginal standardization method will be used to calculate the mean response rate in each treatment group as well as their difference. This method uses the same fitted logistic model, but involves using the model to predict, for each patient in the study, the mean outcome assuming assignment to each particular treatment group in turn, assuming each patient's observed values for the other baseline covariates (i.e., disease condition and weight). Averaging these predictions for each treatment group provides the estimate of the mean response rate for each treatment group. Then

the difference will be derived based on the estimated mean response rates comparing secukinumab i.v. regimen vs placebo.

The macro Margins (Predictive margins and average marginal effects) will be used.

SAS code example as the following,

%Margins(data = *mydata*,

class = *treatment strata*,

response = *response*,

roptions = event='1',

model = *treatment strata*,

dist = binomial,

margins = *treatment*,

options = cl diff)

For cases where the convergence status indicates that the model did not reach appropriate convergence (conv_status is not 0), no risk difference or p-value will be presented from that model.

However, if the issue relates to the primary timepoint of Week 52 then the following steps will be followed:

1. Change race categorization from “Black vs. Non-Black” to “White vs. Non-White”. If there are still issues, perform step 2.
2. Remove race from the model. If there are still issues, perform step 3.
3. Remove FMV baseline score from the model. If there are still issues, perform step 4.
4. Remove strata from the model. If there are still issues, perform step 5.
5. Use Fisher’s exact test as described below.

When Fisher’s exact test is applied, only a p-value for a test of equal response in the two groups will be presented.

ods output fishersexact=*fisher*;

proc freq data=*mydata*;

by *visit*;

table *treatment***response* / fisher;

run;

Input dataset should only contain data from the two treatment groups to be compared.

5.4.3 Kaplan-Meier analysis

The SAS procedure PROC LIFETEST is used to obtain Kaplan-Meier estimates of the proportion of patients with disease flare.

• The Kaplan-Meier estimate for the cumulative Failure (1-survival) rate along with the corresponding 95% confidence interval, using Greenwood's formula to compute the estimate of the standard error, at specific time points,

•median (95% confidence interval) and the estimate of the event rate at the end of the Treatment Period 2 (Week 52) will be provided.

The SAS procedure PROC LIFETEST will be used to obtain Kaplan-Meier estimates of the time to event variables.

The Kaplan-Meier estimate will be derived along with the corresponding 95% confidence interval, using Greenwood's formula to compute the estimate and the standard error, at specific time points,

5.4.4 Imputation methods

Missing values for the primary objective's main analysis will be imputed with non-responder imputation.

5.4.5 Crude incidence and related risk estimates

5.4.5.1 Crude incidence and 100*(1-α)% confidence interval

For n subjects, each at risk to experience a certain event with probability π, the crude incidence is estimated as $p=x/n$, where x is the number of subjects with the event.

Absolute and relative frequencies will be displayed as well as 95% confidence interval for the relative frequency based on the score method including continuity correction (Newcombe 1998).

With z as $(1-\alpha/2)$-quantile of the standard normal distribution (SAS: z = PROBIT(1-α/2), n as total number of subjects (i.e. number of subjects in the denominator), and p as estimated crude incidence (number of subjects with event / n) it is $q=1-p$.

Then the lower limit is

$$L = \max\left(0, \frac{2np + z^2 - 1 - z\sqrt{z^2 - 2 - 1/n + 4p(nq+1)}}{2(n+z^2)}\right)$$

and the upper limit is

$$U = \min\left(1, \frac{2np + z^2 + 1 + z\sqrt{z^2 + 2 - 1/n + 4p(nq-1)}}{2(n+z^2)}\right).$$

In addition, if L > p then L = p and if U < p then U = p.

If appropriate, an exact 100*(1-α)% confidence interval (Clopper-Pearson, 1934) will be obtained by using the SAS procedure PROC FREQ with the EXACT BINOMIAL statement.

However, the confidence interval derived via the score method including continuity correction will be the default in safety analyses.

5.4.5.2 Odds ratio and 100*(1-α)% confidence interval

For an investigational drug group with n_1 subjects at risk, independent from the control group (e.g., placebo or comparator) with n_0 subjects at risk, of whom x_1 and x_0 experience a certain event with probability π_1 and π_0 respectively, the odds ratio is estimated as

$\frac{p_1/(1-p_1)}{p_0/(1-p_0)}$ with $p_1 = x_1/n_1$ and $p_0 = x_0/n_0$. A conditional exact 100*(1-α)% confidence interval can be obtained by using the SAS procedure PROC FREQ with statement EXACT OR. However, to be able to adjust for covariates odds ratios will primarily be obtained from PROC LOGISTIC.

5.4.5.3 Risk difference and 100*(1-α)% confidence interval

For an investigational drug group with n_1 subjects at risk, independent from the control group (e.g., placebo or comparator) with n_0 subjects at risk, of whom x_1 and x_0 experience a certain event, the risk difference is estimated as p_1-p_0 with $p_1 = x_1/n_1$ and $p_0 = x_0/n_0$.

Exact unconditional confidence limits for the risk difference can be obtained with SAS procedure PROC FREQ and option RISKDIFF in the TABLES statement, specifying the RISKDIFF option also in the EXACT statement.

5.4.6 Exposure adjusted incidence rate and related risk estimates

5.4.6.1 Exposure adjusted incidence rate and 100*(1-α)% confidence interval

It will be assumed that for each of n subjects in a clinical trial the time t_j (j=1,...,n) to the first occurrence of a certain event is observed, or if the event was not experienced, the (censored) time to the end of the observation period. The sequence of first occurrences of an event will be modeled to follow approximately a Poisson process with constant intensity θ. The rate parameter θ will be estimated as $\lambda = D/T$, where $T = \sum_{j=1}^{n} t_j$ and D is the number of subjects with at least one event. Conditionally on T, an exact 100*(1-α)% confidence interval for a Poisson variable with parameter θT and observed value D can be obtained based on (Garwood 1936), from which an exact 100*(1-α)% confidence interval for D/T will be derived as follows (Sahai 1993; Ulm 1990):

Lower confidence limit $L = \frac{0.5c_{\alpha/2,2D}}{T}$ for D>0, 0 otherwise,

Upper confidence limit $U = \frac{0.5c_{1-\alpha/2,2D+2}}{T}$

Where $c_{\alpha,k}$ is the αth quantile of the Chi-square distribution with k degrees of freedom. The example below shows how this should be handled for cases where subjects switch treatment. In particular for summarizing 'Any AIN' as a group, one should take into consideration the sequence of treatments while calculating exposure time for subjects. a

Table 5-3 Examples for calculating exposure time for incidence rates (IR)

1st treatment / total exposure time	2nd treatment / total exposure time	AE event onset (in days from study start)	Exposure for IR
Placebo / 100 days	AIN457 150 mg / 200 days	Day 50 (during 1st treatment) Day 110 (10 days into 2nd treatment)	Placebo: 50 days AIN457 150 mg: 10 days Any AIN: 10 days

6 Reference

Bretz F, Maurer W, Brannath W, et al. (2009) A graphical approach to sequentially rejective multiple test procedures. Stat Med; 28(4):586-604.

Bruin G, Loesche C, Nyirady J, et al (2017) Population Pharmacokinetic Modeling of Secukinumab in Patients With Moderate to Severe Psoriasis. J Clin Pharmacol; 57(7):876-85.

Cella DF, Tulsky DS, Gray G et al. (1993) The Functional Assessment of Cancer Therapy (FACT) scale: Development and validation of the general measure. Journal of Clinical Oncology; 11(3):570-579.

Clopper CJ, Pearson ES (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika, 26; 404–413.

Deng J, Luo L, Zhu L, et al (2018) Multitarget therapy versus intravenous cyclophosphamide in the induction treatment of lupus nephritis: a metaanalysis of randomized controlled trials. Turk J Med Sci; 48(5):901-10.

FDA Guidance for Industry: Population Pharmacokinetics p. https://www.fda.gov/media/71364/download.

Furie R, Rovin BH, Houssiau F, et al (2020) Two-Year, Randomized, Controlled Trial of Belimumab in Lupus Nephritis. N Engl J Med; 383(12):1117-28.

Garwood F (1936). Fiducial limits for the Poisson distribution. Biometrika, 46; 441–453.

Glimm, E., Maurer, W., and Bretz, F. (2010) Hierarchical testing of multiple endpoints in group-sequential trials. Statistics in Medicine 29, 219-228

Heinze G (2006). A comparative investigation of methods for logistic regression with separated or nearly separated data. Statistics in Medicine; 25:4216–4226

Koch GG, Tangen CM, Jung JW, et al. (1998) Issues for covariance analysis of dichotomous and ordered categorical data from randomized clinical trials and non-parametric strategies for addressing them. Statistics in Medicine; 17:1863-92.

Lan KK, DeMets DL (1983) Discrete sequential boundaries for clinical trials. Biometrika; 70:659-63.

Levin A, Stevens PE, Bilous RW, et al (2013) 'Kidney disease: Improving global outcomes (KDIGO) CKD work group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease' Kidney International Supplements; 3(1):1-150.

Lubeck, DP. (2004) Patient-reported outcomes and their role in the assessment of rheumatoid arthritis. Pharmacoeconomics 22(1): 27-38.

Martínez-Martínez MU, Mandeville P, Llamzares-Azuara L et al (2013) CKD-EPI is the most reliable equation to estimate renal function in patients with systemic lupus erythematosus. Nefrologia; 33(1):99-106.

McElhone K, Abbott J, Sutton C, et al (2016) Sensitivity to Change and Minimal Important Differences of the LupusQoL in Patients With Systemic Lupus Erythematosus. Arthritis Care Res (Hoboken); 68(10):1505-13.

Mysler EF, Spindler AJ, Guzman R, et al (2013) Efficacy and safety of ocrelizumab in active proliferative lupus nephritis: results from a randomized, double-blind, phase III study. Arthritis Rheum; 65(9):2368-79.

Newcombe, RG. (1998) Two-sided confidence intervals for the single proportion: comparison of seven methods. Statistics in Medicine; 17: 857-872.

Pocock SJ (1977) Group Sequential Methods in the Design and Analysis of Clinical Trials. Biometrika; 64(2):191–9.

Predictive margins and average marginal effects: http://support.sas.com/kb/63/038.html

Permutt T, Li F(2017). Trimmed means for symptom trials with dropouts. Pharmaceutical statistics.;16(1):20-8.

RWS Life Sciences LupusQoL: http://www.corptransinc.com/sites/lupusqol/instrument-information.

Rovin BH, Furie R, Latinis K, et al (2012) Efficacy and safety of rituximab in patients with active proliferative lupus nephritis: the Lupus Nephritis Assessment with Rituximab study. Arthritis Rheum; 64(4):1215-26.

Rovin BH, Teng YKO, Ginzler EM, et al (2021) Efficacy and safety of voclosporin versus placebo for lupus nephritis (AURORA 1): a double-blind, randomised, multicentre, placebo-controlled, phase 3 trial. Lancet; 397:2070-80.

Rubin, DB (1987). Multiple imputation for nonresponse in surveys. John Wiley & Sons.

Sahai H, Khurshid Anwer (1993). Confidence intervals for the mean of a poisson distribution: a review. Biom J, 35 (7); 857-867

Ulm K (1990). A simple method to calculate the confidence interval of a standard mortality ratio. American Journal of Epidemiology, 131(2); 373-375

Wallace DJ, Strand V, Merrill JT, et al (2017) Efficacy and safety of an interleukin 6 monoclonal antibody for the treatment of systemic lupus erythematosus: a phase II dose-ranging randomised controlled trial. Ann. Rheum. Dis; 76(3):534-42.

Yazdany J (2011) Health-related quality of life measurement in adult systemic lupus erythematosus: Lupus Quality of Life (LupusQoL), Systemic Lupus Erythematosus-Specific Quality of Life Questionnaire (SLEQOL), and Systemic Lupus Erythematosus Quality of Life Questionnaire (L-QoL). Arthritis Care Res (Hoboken); 63(11):S413-9.