Official Title: A Phase III, Randomized, Double-Blind, Placebo-Controlled, Multicenter Study to Evaluate the Efficacy and Safety of Obinutuzumab in Patients with Systemic Lupus Erythematosus

NCT Number: NCT04963296

Document Date: Statistical Analysis Plan Version 2: 15-Mar-2024

STATISTICAL ANALYSIS PLAN

STUDY TITLE:	**A PHASE III, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, MULTICENTER STUDY TO EVALUATE THE EFFICACY AND SAFETY OF OBINUTUZUMAB IN PATIENTS WITH SYSTEMIC LUPUS ERYTHEMATOSUS**
STUDY NUMBER:	CA42750
STUDY NAME:	ALLEGORY
VERSION NUMBER:	2
ROCHE COMPOUND:	Obinutuzumab (RO5072759)
EUDRACT NUMBER:	2020-005760-57
IND NUMBER:	125054
NCT NUMBER:	NCT04963296
EU CT NUMBER:	2023-504774-38-00
PLAN PREPARED BY:	[redacted], M.Sc.

STATISTICAL ANALYSIS PLAN APPROVAL

SPONSOR:	F. Hoffmann-La Roche Ltd
LEGAL REGISTERED ADDRESS:	Grenzacherstrasse 124 4070 Basel, Switzerland
DATE FINAL:	See electronic date stamp on the last page of this document.

CONFIDENTIAL

This is an F. Hoffmann-La Roche Ltd document that contains confidential information. Nothing herein is to be disclosed without written consent from F. Hoffmann-La Roche Ltd.

STATISTICAL ANALYSIS PLAN VERSION HISTORY

This Statistical Analysis Plan (SAP) was developed based on Roche SAP model document Version 2, 28 February 2022.

SAP Version	Approval Date	Based on Protocol (Version, Approval Date)
2	See electronic date stamp on last page	Version 4, 29 September 2023
1	08 December 2022	Version 2, 31 March 2022

STATISTICAL ANALYSIS PLAN AMENDMENT RATIONALE

Key changes to the SAP, along with the rationales for each change, are summarized below.

Section(s)	Description of Change	Rationale for Change
1.1, 4.3.1, 4.3.2, 4.4	List of supportive secondary endpoints and exploratory endpoints has been updated	The update was made as per Protocol v4
1.2	Study schema for open-label period has been updated	The update was made as per Protocol v4
2.1	Key secondary endpoint ranking	The update was made to reflect the clinical importance of the key secondary endpoints.
2.2	Sample size increased from 200 to 300.	The update was made as per Protocol v4
3, 4.8	Definition of efficacy-evaluable set has been updated to include all randomized participants regardless of whether they received the study drug or not	The update was made to include all intent-to-treat patients in the efficacy analysis
4.1	Definitions of collected data has been included	The update was made as per Protocol v4
4.2.3	Addition of tipping point analysis	The inclusion of tipping point analysis enables the exploration of the assumptions on missing data
4.5	Safety outputs will include data until the point of rescue for the patients who receive non-corticosteroid rescue therapy.	To increase the clarity of the safety data interpretation

Additional minor changes have been made throughout to improve clarity and consistency.

TABLE OF CONTENTS

STATISTICAL ANALYSIS PLAN AMENDMENT RATIONALE 3

1. INTRODUCTION 9

1.1 Objectives, Endpoints, and Estimands 9

1.2 Study Design 16

1.2.1 Treatment Assignment and Blinding 19

1.2.1.1 Treatment Assignment 19

1.2.1.2 Blinding 19

1.2.2 Data Monitoring 20

2. STATISTICAL HYPOTHESES AND SAMPLE SIZE DETERMINATION 21

2.1 Statistical Hypotheses 21

2.2 Sample Size Determination 22

3. ANALYSIS SETS 22

4. STATISTICAL ANALYSES 23

4.1 General Considerations 23

4.2 Primary Endpoint/Estimand Analysis 24

4.2.1 Definition of Primary Endpoint/Estimand 24

4.2.2 Main Analytical Approach for Primary Endpoints 24

4.2.3 Sensitivity Analyses 27

4.2.4 Supplementary Analyses 28

4.2.4.1 Subgroup Analyses for Primary Endpoint(s) 28

4.2.4.2 Other Supplementary Analyses for Primary Endpoint/Estimand 28

4.3 Secondary Endpoints/Estimands Analyses 29

4.3.1 Key/Confirmatory Secondary Endpoints/Estimands 29

4.3.1.1 Proportion of Participants Entering the Study on Prednisone ≥ 10 mg/day who achieve Sustained Corticosteroid Control from Week 40 through Week 52 29

4.3.1.2 Time to First BILAG Flare Over 52 Weeks 29

4.3.1.3 Proportion of Participants Who Achieve SRI(6) at Week 52 30

4.3.1.4 Proportion of Participants Who Achieve BICLA at Week 52 31

4.3.1.5 Proportion of Participants Who Achieve a Sustained SRI(4) Response from Week 40 through Week 52 31
4.3.2 Supportive Secondary Efficacy Endpoints/Estimands 32
4.3.2.1 Proportion of Participants Who Achieve SRI(4) at Week 24 33
4.3.2.2 Proportion of Participants Who Achieve Clinical SRI(4) at Week 52 Excluding Laboratory SLEDAI-2K Items 33
4.3.2.3 Proportion of Participants Who Achieve SRI(4) at Week 52 on Low-Dose Corticosteroids 33
4.3.2.4 Proportion of Participants Who Achieve Lupus Low Disease Activity State at Week 52 34
4.3.2.5 Proportion of Participants Who Achieve DORIS at Week 52 34
4.3.2.6 Change in FACIT-F Scale from Baseline to Week 24 and Baseline to Week 52 34
4.3.2.7 Change in SF-36 v2 Bodily Pain Domain Scale From Baseline to Week 24 and Baseline to Week 52 34
4.3.2.8 Change in SF-36 v2 Physical Component Summary Scale from Baseline to Week 24 and Baseline to Week 52 34
4.3.2.9 Change in Active Joint Count (Swollen Plus Tender) from Baseline to Week 24 and to Week 52 34
4.3.2.10 Proportion of Participants who Achieve a ≥ 50% Reduction in Active Joint Counts at Each Study Visit 35
4.3.2.11 Proportion of Participants who Achieve a ≥ 50% Reduction in CLASI Total Activity Score at Each Study Visit 35
4.3.2.12 Proportion of Participants who Achieve Sustained Corticosteroid Control from Week 40 through Week 52 35
4.3.2.13 Cumulative Corticosteroid Use Through Week 52 35
4.3.2.14 Annualized Flare Rate Through Week 52 35
4.4 Exploratory Endpoints/Estimands Analysis 35
4.4.1 Binary Endpoint 35
4.4.2 Continuous Endpoint 36
4.4.3 Time to Event Endpoint 36
4.5 Safety Analyses 36
4.5.1 Extent of Exposure 36
4.5.2 Adverse Events 36
4.5.2.1 Laboratory Data 38

4.6 Other Analyses 38
4.6.1 Summaries of Conduct of Study 38
4.6.2 Summaries of Treatment Group Comparability/Demographics and Baseline Characteristics 38
4.6.3 Pharmacokinetic Analyses 39
4.6.4 Immunogenicity Analyses 39
4.7 Interim Analyses 40
4.7.1 Planned Interim Analyses 40
4.8 Changes to Protocol-Planned Analyses 40
5. SUPPORTING DOCUMENTATION 40
6. REFERENCES 41

LIST OF TABLES

Table 1 Primary and Key Secondary Objectives and Corresponding Estimands 9
Table 2 Supportive Secondary and Exploratory Endpoints 13
Table 3 Participant Analysis Sets 22

LIST OF FIGURES

Figure 1 Study Schema 18

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation or Term	Description
ADA	anti-drug antibody
AE	adverse event
BICLA	British Isles Lupus Assessment Group based Composite Lupus Assessment
BILAG	British Isles Lupus Assessment Group
CI	confidence interval
CLASI	Cutaneous Lupus Erythematosus Disease Area and Severity
CMH	Cochran-Mantel-Haenszel
DORIS	Definition of Remission in SLE
DRB	Data Review Board
FACIT-F	Functional Assessment of Chronic Illness Therapy-Fatigue
FCS	Fully Conditional Specification
ICE	intercurrent event
iDMC	independent Data Monitoring Committee
Ig	immunoglobulin
IV	intravenous
IxRS	interactive web response system
LOCF	last observation carried forward
LoPO	list of planned outputs
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
OCS	oral corticosteroid
OLT	open-label treatment
PD	pharmacodynamic
PGA	Physician's Global Assessment
PMM	Predictive Mean Matching
PK	pharmacokinetic
PT	preferred term
SAE	serious adverse event
SE	standard error
SF-36 v2	36 Item Short Form Survey, Version 2
SFU	study follow-up
SLE	systemic lupus erythematosus
SLEDAI-2K	Systemic Lupus Erythematosus Disease Activity Index 2000
SOC	system organ class

Abbreviation or Term	Description
SRI	Systemic Lupus Erythematosus Responder Index
VAS	Visual Analog Scale

1. INTRODUCTION

The analyses described in this Statistical Analysis Plan will supersede those specified in Protocol CA42750 for the purposes of a regulatory filing.

The purpose of this study is to assess the efficacy and safety of obinutuzumab combined with standard-of-care therapy in participants with active systemic lupus erythematosus (SLE). Currently, there are only two approved targeted therapies for the treatment of SLE, namely belimumab and anifrolumab. There remains a need for treatments with better benefit-risk profiles that improve the long-term prognosis of patients with SLE.

1.1 OBJECTIVES, ENDPOINTS, AND ESTIMANDS

Table 1 Primary and Key Secondary Objectives and Corresponding Estimands

Primary Objective	Estimand Definition

• To evaluate the efficacy of obinutuzumab compared with placebo	**Proportion of participants who achieve SRI(4) at Week 52**
	• Population: participants with active, autoantibody-positive SLE as defined by the IC and EC
	• Endpoint: SRI(4) at Week 52
	  • Achieving SRI(4) is defined by all of the following:
	  – Reduction from baseline of ≥4 points in the SLEDAI-2K
	  – No new systems or organs affected, as defined by ≥1 new BILAG A or ≥2 new BILAG B items compared with baseline using BILAG-2004
	  – No worsening from baseline of ≥0.30 points on a 3-point PGA-VAS
	• Treatments:
	  – Experimental: obinutuzumab 1000 mg IV infusion on Day 1 and at Weeks 2, 24, and 26
	  – Control: placebo
	• Intercurrent events:
	The following intercurrent events will be handled by composite variable strategy (i.e., considered as non-responder):
	  – Major concomitant therapy violation
	  – Treatment failure requiring rescue medication
	  – Early study discontinuation due to death, lack of efficacy, and AE

AE = adverse event; BILAG = British Isles Lupus Assessment Group; EC = exclusion criteria; IC = inclusion criteria; PGA = Physician's Global Assessment; SLE = Systemic Lupus Erythematosus; SLEDAI-2K = Systemic Lupus Erythematosus Disease Activity Index 2000; SRI = Systemic Lupus Erythematosus Responder Index; VAS = Visual Analog Scale.

Table 1 Primary and Key Secondary Objectives and Corresponding Estimands (cont.)

Primary Objective	Estimand Definition
• To evaluate the efficacy of obinutuzumab compared with placebo	The following intercurrent event will be handled by treatment policy strategy: – Treatment discontinuation without other concurrent or subsequent intercurrent events (described above) If the other intercurrent events occur concurrently with or after treatment discontinuation, the composite variable strategy will be applied (i.e., considered as non-responder). • Summary measure: difference in proportion of patients who achieve SRI(4) at Week 52
Key Secondary Objective(s)	**Estimand Definition**
• To evaluate the efficacy of obinutuzumab compared with placebo	**Proportion of participants entering the study on prednisone ≥10 mg/day (or equivalent) who achieve Sustained Corticosteroid Control from Week 40 through Week 52** • Population: participants with active, autoantibody-positive SLE as defined by the IC and EC who are on OCS ≥10 mg/day at baseline • Endpoint: Sustained Corticosteroid Control from Week 40 through Week 52, which requires all of the following: – No treatment with prednisone >7.5 mg/day (or equivalent) from Week 40 through Week 52 – No receipt of intravenous, intramuscular, or intra-articular corticosteroids from Week 40 through Week 52 • Treatments: see primary endpoint • Intercurrent events: see primary endpoint • Summary measure: difference in proportion of patients who achieve Sustained Corticosteroid Control from Week 40 through Week 52

EC = exclusion criteria; IC = inclusion criteria; OCS = oral corticosteroid; SLE = Systemic Lupus Erythematosus; SRI = Systemic Lupus Erythematosus Responder Index.

Table 1 Primary and Key Secondary Objectives and Corresponding Estimands (cont.)

Key Secondary Objective(s)	Estimand Definition
• To evaluate the efficacy of obinutuzumab compared with placebo	**Time to first BILAG flare over 52 weeks** • Population: see primary endpoint • Endpoint: time to first BILAG flare over 52 weeks post-baseline, where flare is defined as the occurrence of ≥1 new BILAG A or ≥2 new BILAG B manifestations from the previous visit • Treatments: see primary endpoint • Intercurrent events: For participants who experience treatment failure, major concomitant therapy violation, or lack of efficacy before the first BILAG flare, the participant will be considered as having a flare at the time of the event. For participants who experience early treatment discontinuation due to reasons other than lack of efficacy, the observed time to first BILAG flare before early study discontinuation or study completion will be used. For participants who experience early study discontinuation due to reasons other than lack of efficacy, the participant will be censored at the time of the event • Summary measure: hazard ratio and log rank test **Proportion of participants who achieve SRI(6) at Week 52** • Population: see primary endpoint • Endpoint: proportion of SRI(6) responders at Week 52 Achieving SRI(6) is defined as all of the following: – Reduction from baseline of ≥6 points in the SLEDAI-2K – No new systems or organs affected, as defined by ≥1 new BILAG A or ≥2 new BILAG B items compared with baseline using BILAG-2004 – No worsening from baseline of ≥0.30 points on a 3-point PGA-VAS • Treatments: see primary endpoint • Intercurrent events: see primary endpoint • Summary measure: difference in proportion of patients who achieve SRI(6) at Week 52

BILAG = British Isles Lupus Assessment Group; = Physician's Global Assessment; SLEDAI-2K = Systemic Lupus Erythematosus Disease Activity Index 2000; SRI = Systemic Lupus Erythematosus Responder Index; VAS = Visual Analog Scale.

Table 1 Primary and Key Secondary Objectives and Corresponding Estimands (cont.)

Key Secondary Objective(s)	Estimand Definition
• To evaluate the efficacy of obinutuzumab compared with placebo	**Proportion of participants who achieve BICLA at Week 52** • Population: see primary endpoint • Endpoint: proportion of participants who achieve BICLA at Week 52 Achieving BICLA is defined as all of the following: – Reduction of all baseline BILAG-2004 A items to B/C/D and baseline BILAG-2004 B items to C/D – No new systems or organs affected, as defined by ≥1 new BILAG A or ≥2 new BILAG B items compared with baseline – No net increase in SLEDAI-2K score from baseline – No worsening from baseline of ≥0.30 points on a 3-point PGA-VAS • Treatments: see primary endpoint • Intercurrent events: see primary endpoint • Summary measure: difference in proportion of patients who achieve BICLA at Week 52 **Proportion of participants who achieve Sustained SRI(4) response from Week 40 through Week 52** • Population: see primary endpoint • Endpoint: proportion of participants who achieve Sustained SRI(4) response from Week 40 through Week 52, where achieving Sustained SRI(4) response is defined by achievement of SRI(4) from Week 40 through Week 52 • Treatments: see primary endpoint • Intercurrent events: see primary endpoint • Summary measure: difference in proportion of patients who achieve sustained SRI(4) response from Week 40 through Week 52

BICLA = British Isles Lupus Assessment Group based Composite Lupus Assessment; PGA = Physician's Global Assessment; SLEDAI-2K = Systemic Lupus Erythematosus Disease Activity Index 2000; SRI = Systemic Lupus Erythematosus Responder Index; VAS = Visual Analog Scale.

Table 2 Supportive Secondary and Exploratory Endpoints

Supportive Secondary Objectives	Corresponding Endpoints
• To evaluate the efficacy of obinutuzumab compared with placebo	• Proportion of participants who achieve SRI(4) at Week 24 • Proportion of participants who achieve clinical SRI(4) at Week 52, where clinical SRI(4) is defined as SRI(4) excluding consideration of laboratory SLEDAI-2K items • Proportion of participants who achieve SRI(4) at Week 52 on low-dose corticosteroids, defined as prednisone $\leq$ 7.5 mg/day (or equivalent) • Proportion of participants who achieve LLDAS at Week 52, where LLDAS is defined as achievement of all of the following: – SLEDAI-2K $\leq$ 4 – No SLEDAI-2K activity in major organ systems (renal, CNS, cardiopulmonary, vasculitis, fever) and no BILAG hemolytic anemia or gastrointestinal activity – No new SLEDAI-2K activity compared with the previous assessment – PGA-VAS $\leq$ 1 – Current prednisone dose $\leq$ 7.5 mg/day (or equivalent) • Proportion of participants who achieve DORIS at Week 52, where DORIS is defined as achievement of all of the following: – Clinical SLEDAI-2K = 0 – PGA-VAS $<$ 0.5 – Current prednisone dose $\leq$ 5 mg/day (or equivalent) • Change in FACIT-F scale from baseline to Week 24 and from baseline to Week 52 • Change in SF-36 v2 Bodily Pain domain scale from baseline to Week 24 and from baseline to Week 52

BILAG = British Isles Lupus Assessment Group; CNS = central nervous system; DORIS = Definition of Remission in SLE; LLDAS = Lupus Low Disease Activity State; FACIT-F = Functional Assessment of Chronic Illness Therapy-Fatigue; PGA = Physician's Global Assessment; SF-36 v2 = 36-Item Short Form Survey, Version 2; SLEDAI-2K = Systemic Lupus Erythematosus Disease Activity Index 2000; SRI = Systemic Lupus Erythematosus Responder Index; VAS = Visual Analog Scale.

Table 2 Supportive Secondary and Exploratory Endpoints (cont.)

Supportive Secondary Objectives	Corresponding Endpoints
• To evaluate the efficacy of obinutuzumab compared with placebo	• Change in SF-36 v2 Physical Component Summary scale from baseline to Week 24 and from baseline to Week 52 • Change in active joint count (swollen plus tender) from baseline to Week 24 and from baseline to Week 52 • Proportion of participants who achieve a ≥50% reduction in active joint counts (swollen plus tender) at each study visit, among participants with ≥4 active joints at baseline • Proportion of participants who achieve a ≥50% reduction in CLASI Total Activity Score at each study visit, among participants with CLASI Total Activity Score ≥10 at baseline • Proportion of participants who achieve Sustained Corticosteroid Control from Week 40 through Week 52, which requires all of the following: - No treatment with prednisone >5 mg/day (or equivalent) from Week 40 through Week 52 - No receipt of intravenous, intramuscular, or intra-articular corticosteroids from Week 40 through Week 52 • Cumulative corticosteroid use (in equivalent milligrams of prednisone) through Week 52 • Annualized flare rate through Week 52
• To evaluate the safety of obinutuzumab compared with placebo	• Incidence and severity of AEs, with severity determined according to NCI CTCAE v5.0 • Characterization of AESIs, including, among others, IRRs, neutropenia, infections, and thrombocytopenia • Change from baseline in targeted vital signs • Change from baseline in targeted clinical laboratory test results
• To characterize the obinutuzumab PK profile	• Serum concentration of obinutuzumab at specified timepoints

AE = adverse event; AESI = adverse event of special interest; CLASI = Cutaneous Lupus Erythematosus Disease Area and Severity; IRR = infusion-related reaction; NCI CTCAE v5.0 = National Cancer Institute Common Terminology Criteria for Adverse Events, Version 5.0; SF-36 v2 = 36-Item Short Form Survey, Version 2.

Table 2 Supportive Secondary and Exploratory Endpoints (cont.)

Supportive Secondary Objectives	Corresponding Endpoints
• To evaluate the immune response to obinutuzumab	• Prevalence of ADAs at baseline and incidence of ADAs during the study (only for participants treated with obinutuzumab)
Exploratory Objectives	**Corresponding Endpoints**
• To evaluate the efficacy of obinutuzumab compared with placebo	• Proportion of participants who achieve SRI(8) at Week 52, where achieving SRI(8) is defined as all of the following: – Reduction from baseline of ≥8 points in the SLEDAI-2K – No new systems or organs affected, as defined by ≥1 new BILAG A or ≥2 new BILAG B items compared with baseline using BILAG-2004 – No worsening from baseline of ≥0.30 points on a 3-point PGA-VAS • Proportion of participants who achieve the individual components of SRI, BICLA, LLDAS at Week 52 • Proportion of participants who experience intercurrent events by Week 52 • SLEDAI-2K change from baseline to Week 52 • Time to first SRI(4) response sustained over two consecutive scheduled visits • Proportions of participants who achieve SRI(4), clinical SRI(4), SRI(6), SRI(8), BICLA, and LLDAS at each study visit • Change in anti-dsDNA titer, C3, and C4 from baseline at each study visit • Change in PGA and SGA from baseline at each study visit • Change in SF-36 v2 Mental Component Summary scale from baseline to Week 52 • Change in SF-36 v2 Physical Functioning, Role Physical, Role Emotional, General Health, Vitality, Social Functioning, Mental Health domain scales from baseline to Week 52

ADA = anti-drug antibody; BICLA = British Isles Lupus Assessment Group-based Composite Lupus Assessment; BILAG = British Isles Lupus Assessment Group; C3 = complement component 3; C4 = complement component 4; dsDNA = double-stranded DNA; LLDAS = Lupus Low Disease Activity State; PGA = Physician's Global Assessment; SF-36 v2 = 36-Item Short Form Survey, Version 2; SGA = Subject's Global Assessment; SLEDAI-2K = Systemic Lupus Erythematosus Disease Activity Index 2000; SRI = Systemic Lupus Erythematosus Responder Index; VAS = Visual Analog Scale.

Table 2 Supportive Secondary and Exploratory Endpoints (cont.)

Exploratory Objectives	Corresponding Endpoints
• To evaluate the efficacy of obinutuzumab compared with placebo	• Proportion of participants who achieve both SRI(4) and BICLA response at Week 52 by meeting all of the following: – Reduction from baseline of ≥4 points in the SLEDAI-2K – Reduction of all baseline BILAG-2004 A items to B/C/D and baseline BILAG-2004 B items to C/D – No new systems or organs affected, as defined by ≥1 new BILAG A or ≥2 new BILAG B items compared with baseline using BILAG-2004 – No worsening from baseline of ≥0.30 points on a 3-point PGA-VAS
• To characterize the obinutuzumab-induced changes in circulating B-cells compared with placebo	• Total peripheral B-cell count via HSFC (MRB 1.1) • Proportion of participants achieving sustained B-cell depletion (<0.441 B-cells/μL of blood at Weeks 24 and 52)
• To evaluate health status utility scores of participants treated with obinutuzumab compared with placebo	• Change from baseline in EQ-5D-5L index-based and VAS scores at Weeks 0, 24, and 52

BICLA = British Isles Lupus Assessment Group-based Composite Lupus Assessment; BILAG = British Isles Lupus Assessment Group; dsDNA = double-stranded DNA; EQ-5D-5L = EuroQol 5-Dimension, 5-Level Questionnaire; HSFC = high-sensitivity flow cytometry; PGA = Physician's Global Assessment; SRI = Systemic Lupus Erythematosus Responder Index; VAS = Visual Analog Scale.

1.2 STUDY DESIGN

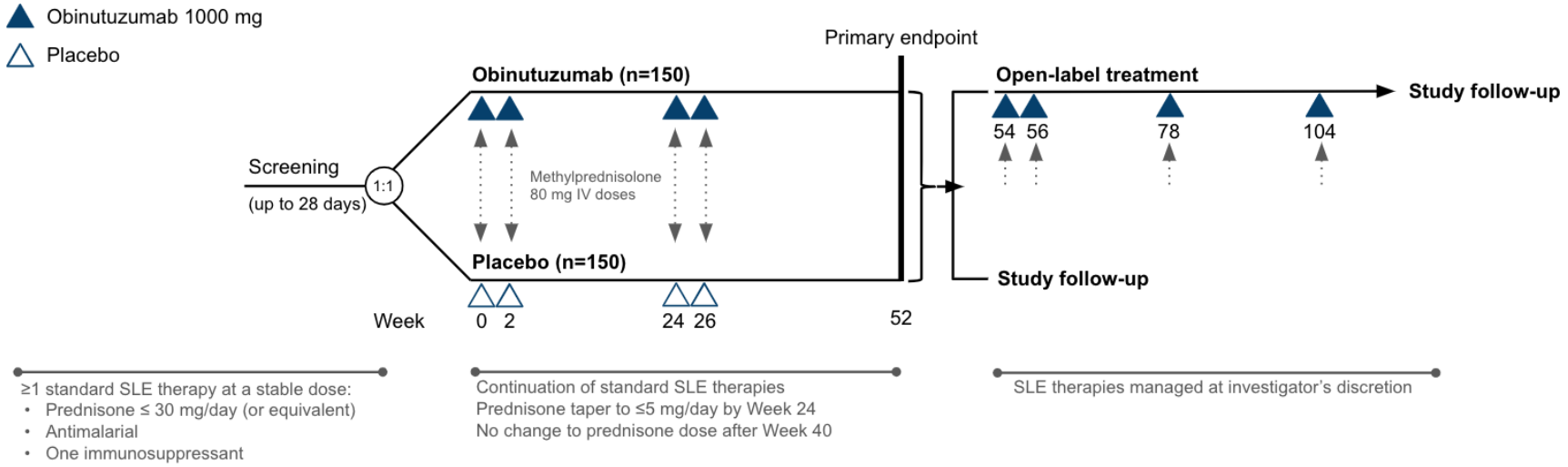
This is a parallel-group, double-blind, randomized, placebo-controlled study evaluating the efficacy and safety of obinutuzumab versus placebo among participants with active, autoantibody-positive SLE treated with standard-of-care therapy.

The study consists of the following four periods: screening, blinded treatment, open-label treatment (OLT) and study follow-up (SFU).

Consenting participants will first enter a screening period of up to 28 days to be evaluated for eligibility. Upon confirmation of eligibility, participants will be randomized in a 1:1 ratio to receive blinded infusions of obinutuzumab 1000 mg or placebo on Day 1 and at Weeks 2, 24, and 26, as shown in Figure 1. Standard SLE therapies will be continued, and a corticosteroid taper is required. The primary endpoint, Systemic Lupus Erythematosus Responder Index 4 (SRI[4]), will be assessed at Week 52. Participants who complete the blinded treatment (the Week 52 assessment) may be eligible to receive OLT with obinutuzumab. The randomized treatment assignment will not be

revealed to participants and investigators until study unblinding occurs (except if unblinding is necessary for a premature single subject in the case of a medical emergency). A 12-month SFU period from the last blinded/open-label obinutuzumab infusion will be required to monitor safety and disease activity. Participants with persistent B-cell depletion may require additional follow-up as defined in the protocol.

Figure 1 Study Schema



SLE = systemic lupus erythematosus.

1.2.1 Treatment Assignment and Blinding

1.2.1.1 Treatment Assignment

This is a parallel-group, double-blind, placebo-controlled randomized study. After initial written informed consent has been obtained, all screening procedures and assessments have been completed, and eligibility has been established for a participant, the investigator or the investigator's research staff will provide participant eligibility information through the interactive web response system (IxRS) at randomization. Each eligible participant will be randomized and assigned a unique identification number. As confirmation, the investigator will be provided with written verification of each participant's registration.

Participants will be randomized to receive obinutuzumab or placebo in a 1:1 ratio. The randomization of participants into treatment and control groups will be managed by a central IxRS vendor and will be performed by stratified permutated block design. The stratification factors include baseline disease activity oral corticosteroid (OCS) dose as follows:

- Baseline disease activity
  - Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K) score (measured at screening) ≥ 12
  - SLEDAI-2K score (measured at screening) < 12
- Baseline OCS dose
  - Prednisone ≥ 10 mg/day
  - Prednisone < 10 mg/day

These stratification factors were selected given expected differences in response by baseline disease activity and baseline OCS dose.

Achievement of SRI(4) includes an absolute reduction of four points or more in the SLEDAI-2K score from baseline, which may represent a small or large relative reduction in a participant's disease activity depending on their baseline total score. Associations have previously been reported between baseline SLEDAI-2K score and the achievement of SRI(4) response (Furie et al. 2014; Gatto et al. 2020). Participants receiving a high dose of corticosteroids at baseline may be at increased risk for flares after protocol-required tapering of dose, in comparison to participants who had a lower or nonexistent corticosteroid dose.

1.2.1.2 Blinding

Study site personnel, investigators, participants, and the Sponsor will be blinded to treatment assignment during the study. Assessment of data collected with the potential to unblind treatment assignment, such as all laboratory studies of blood specimens, will be performed by a central laboratory. Therefore, site personnel and the Sponsor's staff involved with the conduct of the study will not receive unblinded data related to

peripheral B-cell counts, pharmacokinetic (PK) results, specific immunoglobulin levels (immunoglobulin [Ig]M, IgA), or anti-drug antibody (ADA) results until the primary efficacy and safety analyses through Week 52 have been completed.

While PK and immunogenicity samples must be collected from participants assigned to the comparator arm to maintain the blinding of treatment assignment, PK and ADA assay results for these participants are generally not needed for the safe conduct or proper interpretation of the study data. Laboratories responsible for performing study drug PK and ADA assays will be unblinded to participant treatment assignments to identify appropriate samples for analysis. Pharmacokinetic samples from participants assigned to the comparator arm will not be analyzed for study drug PK concentration except by request (e.g., to evaluate a possible error in dosing). Baseline immunogenicity samples will be analyzed for all participants. Post-baseline immunogenicity samples from participants assigned to the comparator arm will not be analyzed for ADAs except by request.

If emergency unblinding is necessary for participant management (e.g., in the case of a serious adverse event [SAE] for which patient management might be affected by knowledge of treatment assignment), the investigator will be able to break the treatment code by contacting the IxRS. Treatment codes should not be broken except in emergency situations. If the investigator wishes to know the identity of the study treatment for any other reason (non-emergency unblinding), they should contact the Medical Monitor directly. If the Medical Monitor agrees to patient unblinding, the investigator will be able to break the treatment code by contacting the IxRS. The investigator should document and provide an explanation for any premature unblinding.

As per health authority reporting requirements, the Sponsor's Drug Safety representative will break the treatment code for all serious, unexpected suspected adverse reactions that are considered by the investigator or Sponsor to be related to an investigational medicinal product. The participant may continue to receive treatment, and the investigator, participant, and Sponsor personnel, with the exception of the Drug Safety representative and personnel who must have access to participant treatment assignments to fulfill their roles (as defined above), will remain blinded to treatment assignment.

1.2.2 Data Monitoring

An independent Data Monitoring Committee (iDMC) will be used to monitor study data on an ongoing basis. The iDMC will meet approximately every three to six months to evaluate unblinded study data, which will be prepared for the committee by an independent Data Coordinating Center.

In addition to the regularly scheduled safety reviews, an unscheduled review of the data may be performed at the request of the iDMC or study team, on the basis of a perceived concern for participant safety.

The Sponsor's study team will have no direct contact with the iDMC during the trial conduct except for the scheduled open sessions of the iDMC. The iDMC will provide recommendations to the Sponsor's Data Review Board (DRB) chair as described in the iDMC charter. On behalf of the Sponsor, the DRB chair will accept or reject the recommendations.

Further details can be found in the iDMC Charter.

2. STATISTICAL HYPOTHESES AND SAMPLE SIZE DETERMINATION

2.1 STATISTICAL HYPOTHESES

The primary efficacy analysis for this trial will compare obinutuzumab with placebo on the basis of SRI(4) achievement at Week 52 to demonstrate superiority of obinutuzumab over placebo. The following null and alternative hypothesis will be tested at a two-sided $\alpha = 0.05$ significance level:

$$H_0: P_{Obinutuzumb} = P_{Placebo}, \text{ versus } H_a: P_{Obinutuzumb} \neq P_{Placebo}$$

for which $P_{Obinutuzumb}$ and $P_{Placebo}$ refer to proportion of participants who achieve SRI(4) for obinutuzumab and placebo, respectively.

Multiplicity Adjustment

Primary and key secondary endpoints will be tested to compare obinutuzumab with placebo groups for the superiority of obinutuzumab over placebo. Type I error will be controlled for multiplicity using a fixed sequence method (Tamhane et al. 2010). The primary and key secondary endpoints will be tested in the order shown below at a two-sided $\alpha = 0.05$ significance level, proceeding sequentially from the primary endpoint and testing each endpoint after achieving statistical significance on the previous endpoint.

The sequence for endpoint testing will be the primary endpoint, followed by key secondary endpoints in the following manner:

1. Proportion of participants who achieve SRI(4) at Week 52
2. Proportion of participants who achieve British Isles Lupus Assessment Group based Composite Lupus Assessment (BICLA) at Week 52
3. Proportion of participants who achieve an OCS dose reduction by at least 25% to ≤ 7.5 mg/day at Week 40, which is maintained through Week 52 in the subgroup of patients with baseline OCS ≥ 10 mg/day
4. Proportion of participants who achieve Sustained SRI(4) response from Week 40 through Week 52
5. Proportion of participants who achieve SRI(6) at Week 52
6. Time to first British Isles Lupus Assessment Group (BILAG) flare over 52 weeks

2.2 SAMPLE SIZE DETERMINATION

The primary efficacy endpoint of this study is the proportion of participants who achieve SRI(4). Based upon previous studies, it is estimated that approximately 35% of SLE participants who are receiving standard therapies will achieve SRI(4) at Week 52, and that the addition of obinutuzumab to standard therapies will induce an overall SRI(4) rate of 55% at Week 52. On the basis of these assumptions, a total of 300 participants randomized to obinutuzumab and placebo groups in a 1:1 ratio (150 participants in each of the obinutuzumab and placebo groups), stratified by baseline disease activity and OCS dose, will yield over 90% power to detect a 20% treatment effect at the two-sided $\alpha = 0.05$ significance level using a Cochran-Mantel-Haenszel (CMH) test, assuming the same SRI(4) proportions across the strata. In addition, a sample size of 300 will achieve a minimal detectable difference of at least 11% in the key secondary endpoints, assuming placebo response rates up to 35%.

3. ANALYSIS SETS

The participant analysis sets for the purposes of analyses are defined in Table 3.

Table 3 Participant Analysis Sets

Participant Analysis Set	Description
Full analysis	The full analysis set will include all participants who are randomized into the study.
Efficacy-evaluable	The efficacy-evaluable set will include all randomized participants. Participants will be grouped according to randomized (assigned) treatment, rather than treatment received. Participants who received an incorrect therapy will be reported under the treatment arm to which they were randomized. All efficacy analyses will be performed using the efficacy-evaluable set.
Safety-evaluable	The safety-evaluable set will include all randomized participants who receive any amount of study drug. Participants who were randomized but did not receive any part of an infusion of study drug will not be included in the safety population. Participants will be grouped and analyzed according to the treatment they actually received. All safety analyses will be performed using the safety-evaluable set.
PK-evaluable	The PK evaluable set will include all participants in the efficacy-evaluable set who have received any dose of obinutuzumab given as study drug, have at least one post-dose PK sample that is evaluable, and have no major protocol deviations that would impact the PK results. All PK analyses will be performed using the PK evaluable set.
ADA-evaluable	The ADA evaluable analysis set will include all participants with at least one ADA assessment. Participants will be grouped according to the treatment received or, if no treatment is received prior to study discontinuation, according to the treatment assigned.

ADA = anti-drug antibody; PK = pharmacokinetic.

4. STATISTICAL ANALYSES

The primary analysis will be performed when all participants have completed the Week 52 visit or have withdrawn early from the study, all data collected through Week 52 are in the database, and the data have been cleaned and verified. A database snapshot will be taken for the primary analysis as the study will be ongoing.

The final analysis will be performed when all participants have either completed all the study follow-up or withdrawn early from the study, all data from the study are in the database, and the database is locked.

4.1 GENERAL CONSIDERATIONS

All efficacy analyses will be performed in the efficacy-evaluable population, unless otherwise specified. Participants will be analyzed according to the treatment assigned at randomization.

All safety analyses will be performed in the safety-evaluable population, unless otherwise specified. Patients will be analyzed according to the treatment they actually received.

Analyses of the primary and secondary efficacy endpoints will be stratified by baseline disease activity and OCS dose, unless otherwise noted. The stratification factors as recorded in IxRS will be used in the analyses.

In addition to specific analyses that are detailed in the following sections, collected data will be defined as summarized:

- For each participant, the first day of study treatment will be designated Study Day 1. Each assessment will be assigned a study day calculated as follows: Date of Assessment − Date of Randomization + 1.
- Baseline is defined as the last non-missing value prior to the receipt of study treatment. Typically, baseline will be the pre-dose Day 1 assessment, but may include screening results if the pre-dose Day 1 result is not available or missing.
- Intercurrent events include the following:
  - Major concomitant therapy violation
  - Treatment failure and receiving rescue medication
  - Early study discontinuation due to death, lack of efficacy, and adverse event (AE)
  - Treatment discontinuation

Handling of intercurrent events and missing values will be discussed in the following sections.

4.2 PRIMARY ENDPOINT/ESTIMAND ANALYSIS

4.2.1 Definition of Primary Endpoint/Estimand

The primary endpoint is the proportion of participants who achieve SRI(4) at Week 52. The estimand for the primary efficacy endpoint is defined as follows:

- Population: participants with active, autoantibody-positive SLE as defined by the inclusion criteria and exclusion criteria
- Primary endpoint (variable): SRI(4) at Week 52

  SRI(4) is defined as all of the following:

  – Reduction from baseline of $\geq$ 4 points in the SLEDAI-2K

  – No new systems or organs affected, as defined by $\geq$ 1 new BILAG A or $\geq$ 2 new BILAG B items compared with baseline using BILAG-2004

  – No worsening from baseline of $\geq$ 0.30 points on a 3-point Physician's Global Assessment- Visual Analog Scale (PGA-VAS)
- Treatments:

  – Experimental: obinutuzumab 1000 mg intravenous (IV) infusion on Day 1 and at Weeks 2, 24, and 26

  – Control: placebo
- Intercurrent events

  The following intercurrent events will be handled by composite variable strategy (i.e., considered as non-responder):

  – Major concomitant therapy violation

  – Treatment failure requiring rescue medication

  – Early study discontinuation due to death, lack of efficacy, and AE

  The following intercurrent event will be handled by treatment policy strategy:

  – Treatment discontinuation without other concurrent or subsequent intercurrent events (described above)

  – If the other intercurrent events occur concurrently with or after treatment discontinuation, the composite variable strategy will be applied (i.e., considered as non-responder).
- Summary measure: difference in proportion of patients who achieve SRI(4) at Week 52

4.2.2 Main Analytical Approach for Primary Endpoints

The proportion of participants achieving SRI(4) across obinutuzumab and placebo groups will be compared using a CMH test stratified by:

- SLEDAI-2K score at screening: SLEDAI-2K score $\geq$ 12 vs. $<$ 12
- Baseline OCS dose: prednisone or equivalent $\geq$ 10 mg/day vs. $<$ 10 mg/day

The proportion of participants achieving SRI(4) in obinutuzumab and placebo arms will be presented along with the difference, 95% confidence interval (CI) for the difference calculated from the pooled analysis result from multiple imputation, and p-value. The hypothesis test will be conducted at a two-sided $\alpha = 0.05$ significance level.

SRI(4) is a composite endpoint derived using several sources of information. Any missing or ambiguous data will be queried and all efforts will be made to resolve data issues prior to the primary analysis snapshot and final database lock.

Missing Data Imputation

Missing data of SRI(4) will be imputed using multiple imputation at a component level. The SLEDAI-2K total score will be imputed as a continuous variable. The criteria as to whether or not there are ≥ 1 new BILAG Category A or ≥ 2 new BILAG Category B items compared with baseline (referred as binary BILAG response hereafter) will be imputed as a binary variable. Physician's Global Assessment-Visual Analog Scale score will be imputed as a continuous variable.

Missing data in SLEDAI-2K total score and PGA-VAS score at Week 52 will be imputed using a Fully Conditional Specification (FCS) Predictive Mean Matching (PMM) approach. Missing data in binary response of BILAG-2004 will be imputed using an FCS-logistic approach. Fully Conditional Specification is also known as multivariate imputation by chained equations and will follow the method described in White et al. 2011. A set of candidate donors with 5 members is chosen for the PMM method. Separate imputation models will be used for each treatment arm. The handling strategy for intercurrent events will be implemented after imputation within each imputed dataset before analysis and pooling. SRI(4) will be derived for each participant in each imputed dataset according to the primary endpoint definition in Section 4.2.1.

The model for imputation was developed based on a Phase IIb SLE study ATHOS (GA30044) data. Justification for the imputation model is described below and follows the guidance by van Buuren et al. 1999 and White et al. 2011. The list of auxiliary variables was chosen based on clinical knowledge of factors that are known to be associated with SRI(4) measurement. The model will contain the following variables:

- Stratification factors: SLEDAI-2K measured at screening and OCS dose at baseline
- Baseline demographics: age, race, sex, and duration of disease
- Measurements of disease activities: SLEDAI-2K score at all visits, binary BILAG response at all visits and PGA-VAS score at all visits

The detailed imputation method for each of the SRI(4) component are listed below:

- SLEDAI-2K total score: The imputation model will be fit with observed values of SLEDAI-2K total score as the dependent variable, and SLEDAI-2K total score at each prior timepoint, PGA-VAS at each prior timepoint, baseline stratification factors, and baseline demographics as covariates.

- PGA-VAS: Similar to SLEDAI-2K total score, the imputation model will be fit with observed values of PGA-VAS as the dependent variable, and PGA-VAS at each prior timepoint, SLEDAI-2K total score at each prior timepoint, baseline stratification factors, and baseline demographics as covariates.
- Binary BILAG response: The imputation model will be fit with binary BILAG response as the dependent variable, and binary BILAG response at each prior timepoint, baseline stratification factors, and baseline demographics as covariates.

At baseline, the missing values for SLEDAI-2K, PGA-VAS and binary response for BILAG-2004 will be imputed from the latest assessment during screening visit. The imputation model will be specified as follows:

- $\mathrm{SLEDAI_j} = \beta_0 + \sum_{r=0}^{j-1} \beta_{r+1} \times \mathrm{SLEDAI_r} + \sum_{r=0}^{j-1} \beta_{r+j+1} \times \mathrm{PGA_r} + \beta_{2j+1} \times \text{screening SLEDAI}[\geq 12 \text{ vs.} < 12] + \beta_{2J+2} \times \text{baseline OCS}[\geq 10 \text{ mg/day vs.} < 10\text{mg/day}] + \beta_{2J+3} \times \text{age} + \beta_{2j+4} \times \text{race}[\text{black vs. non} - \text{black}] + \beta_{2j+5} \times \text{sex} + \beta_{2j+6} \times \text{duration of disease}$
- $\mathrm{PGA_j} = \beta_0 + \sum_{r=0}^{j-1} \beta_{r+1} \times \mathrm{PGA_r} + \sum_{r=0}^{j-1} \beta_{r+j+1} \times \mathrm{SLEDAI_r} + \beta_{2j+1} \times \text{screening SLEDAI}[\geq 12 \text{ vs.} < 12] + \beta_{2J+2} \times \text{baseline OCS}[\geq 10 \text{ mg/day vs.} < 10\text{mg/day}] + \beta_{2J+3} \times \text{age} + \beta_{2j+4} \times \text{race}[\text{black vs. non} - \text{black}] + \beta_{2j+5} \times \text{sex} + \beta_{2j+6} \times \text{duration of disease}$
- $\mathrm{BILAG_j} = \beta_0 + \sum_{r=0}^{j-1} \beta_{r+1} \times \mathrm{BILAG_r} + \beta_{j+1} \times \text{screening SLEDAI}[\geq 12 \text{ vs.} < 12] + \beta_{J+2} \times \text{baseline OCS}[\geq 10 \text{ mg/day vs.} < 10\text{mg/day}] + \beta_{J+3} \times \text{age} + \beta_{j+4} \times \text{race}[\text{black vs. non} - \text{black}] + \beta_{j+5} \times \text{sex} + \beta_{j+6} \times \text{duration of disease}$

where $j = 1,2,\ldots,13$ such that $\mathrm{SLEDAI\text{-}2K_j}$, $\mathrm{PGA_j}$ and binary BILAG response at Week $j \times 4$.

If quasi-complete separation pattern is observed in the binary BILAG response imputation, augmentation will be applied to ensure maximum likelihood estimates for the FCS logistic model exist. If all observed binary BILAG response are identical at a particular visit, the missing value will be imputed by the observed BILAG response.

A number of burn-in iterations of 100 and a random seed of 321 will be used with 100 imputed datasets. Each of the 100 complete datasets will be analyzed using the CMH test specified in Section 4.2.2. A Wilson-Hilferty transformation will be applied to the CMH test statistics for normalization (Ratitch et al. 2013). Stratified Newcombe CIs will be obtained for the difference in proportions (i.e., common risk difference), estimated from the CMH test. The standard errors (SE) associated with the stratified Newcombe confidence intervals for the difference in proportions from the CMH test are required in order to pool analyses using Rubin's rules; these will be calculated as follows:

$$\mathrm{SE} = (\text{Newcombe upper limit} - \text{Newcombe lower limit})/Z$$

where $Z = 3.92$ for 95% CIs (Higgins et al. 2021).

4.2.3 Sensitivity Analyses

The following sensitivity analysis will be performed for the primary efficacy endpoint to evaluate the robustness of the primary analysis finding:

a) Test assumption of correlation between SLEDAI-2K and PGA

The same estimand and analysis method as defined in the primary analysis (Section 4.2.1 and Section 4.2.2) will be used, with the exception of multiple imputation model. To test the assumption of dependency between the SRI(4) components for the same patient, the imputation model will be modified as follows:

- $\text{SLEDAI}_j = \beta_0 + \sum_{r=0}^{j-1} \beta_{r+1} \times \text{SLEDAI}_r + \beta_{j+1} \times \text{screening SLEDAI}[\geq 12 \text{ vs.} < 12] + \beta_{J+2} \times \text{baseline OCS}[\geq 10 \text{ mg/day vs.} < 10\text{mg/day}] + \beta_{J+3} \times \text{age} + \beta_{j+4} \times \text{race}[\text{black vs. non} - \text{black}] + \beta_{j+5} \times \text{sex} + \beta_{j+6} \times \text{duration of disease}$
- $\text{PGA}_j = \beta_0 + \sum_{r=0}^{j-1} \beta_{r+1} \times \text{PGA}_r + \beta_{j+1} \times \text{screening SLEDAI}[\geq 12 \text{ vs.} < 12] + \beta_{J+2} \times \text{baseline OCS}[\geq 10 \text{ mg/day vs.} < 10\text{mg/day}] + \beta_{J+3} \times \text{age} + \beta_{j+4} \times \text{race}[\text{black vs. non} - \text{black}] + \beta_{j+5} \times \text{sex} + \beta_{j+6} \times \text{duration of disease}$
- $\text{BILAG}_j = \beta_0 + \sum_{r=0}^{j-1} \beta_{r+1} \times \text{BILAG}_r + \beta_{j+1} \times \text{screening SLEDAI}[\geq 12 \text{ vs.} < 12] + \beta_{J+2} \times \text{baseline OCS}[\geq 10 \text{ mg/day vs.} < 10\text{mg/day}] + \beta_{J+3} \times \text{age} + \beta_{j+4} \times \text{race}[\text{black vs. non} - \text{black}] + \beta_{j+5} \times \text{sex} + \beta_{j+6} \times \text{duration of disease}$

where $j = 1,2,\ldots,13$ such that SLEDAI-2K$_j$, PGA$_j$ and binary BILAG response at Week $j \times 4$.

b) Test the impact of missing data by last observation carried forward approach

A sensitivity analysis will be performed on the estimand of primary endpoint (Section 4.2.1), where last observation carried forward (LOCF) will be implemented for the imputation of any missing values in SLEDAI-2K, BILAG-2004, and PGA-VAS at questionnaire item level, if only a single (non-consecutive) visit had missing data. If missing data is found in two or more consecutive visits, only the first missing value will be imputed. For SLEDAI-2K, missing data will be imputed at an item level. For BILAG domains that require laboratory measurements, if a laboratory is missing, both the laboratory value and SLE-attribution status will be LOCF from the same previous visit. For BILAG domains that do not require laboratory measurements, BILAG algorithm will be applied to the existing data in determining appropriate BILAG score for the visit. If BILAG score is still not calculable, the domain score will be LOCF. Missing PGA score will be imputed by LOCF method. Participants with missing data after the imputation will be considered as non-responders.

c) Tipping point analysis

The tipping point analysis will vary assumptions about missing outcomes for the two treatment arms independently, to explore scenarios under which there is no longer evidence of treatment effect. This analysis will target the supplementary analysis based on treatment policy strategy (i.e. include all observed data regardless of treatment adherence). Detail of the supplementary analysis is outlined in Section 4.2.4.2 below.

For this sensitivity analysis, the response values for participants with missing SRI-4 will be imputed deterministically after the handling of intercurrent event (ICE), exploring all possible responder/non-responder combinations across treatment arms. In each unique scenario, a Pearson's chi-squared test will be implemented to assess the treatment effect given the imputed response values in each respective scenario. The stratification factors used in the main analysis will be disregarded and an unstratified analysis will be carried out within the tipping point analysis. The only foreseen impact of carrying out an unstratified analysis is that conclusions will be slightly more conservative.

For each unique scenario, a corresponding p-value from the Pearson's chi-squared test will be obtained, providing a result of statistical significance, at a level of 0.05. These results will be plotted on a grid, with the x- and y-axes representing the number of participants who have been imputed as responders for the placebo and obinutuzumab arms respectively. The region on the produced plot where the conclusion changes (significant to non-significant, $p < 0.05$ to $p \geq 0.05$) will be interpreted as the tipping point.

4.2.4 Supplementary Analyses

4.2.4.1 Subgroup Analyses for Primary Endpoint(s)

The generalizability of SRI(4) response results when comparing obinutuzumab group to placebo group will be investigated by estimating the treatment effect in the following sub-populations:

- Baseline OCS dose ($\geq$ 10mg/day or $<$ 10mg/day)
- Baseline disease activity (screening SLEDAI-2K score $\geq$ 12 or $<$ 12)
- Race (Black vs. other)
- Ethnicity (Hispanic or Latino vs. not Hispanic or Latino)
- Baseline immunosuppressant (received vs. not received)
- Baseline urinary protein-to-creatinine ratio ($\geq$ 0.5 g/g vs. $<$ 0.5 g/g)
- Baseline complement proteins (low level vs. normal)
- Baseline anti–double-stranded DNA (high level vs. normal)

Summaries of SRI(4) response by these subgroups will be provided in tables and forest plots.

4.2.4.2 Other Supplementary Analyses for Primary Endpoint/Estimand

The following supplementary analyses will be performed for the primary efficacy endpoint to provide further understanding of the treatment effect:

Analysis Using Treatment Policy Strategy for Intercurrent Events

The main analysis method will be the same as the primary analysis (Section 4.2.1 and Section 4.2.2), except for the intercurrent event handling strategies. Early study discontinuation due to death, lack of efficacy, and AE will still be handled by composite variable strategy. However, all other intercurrent events will be handled by treatment

policy strategy, where all observed values will be used regardless of the occurrence of major concomitant medication violation, treatment failure requiring rescue medication, or treatment discontinuation.

4.3 SECONDARY ENDPOINTS/ESTIMANDS ANALYSES

4.3.1 Key/Confirmatory Secondary Endpoints/Estimands

The definition of the following key secondary endpoints is specified in Table 1. Details of the estimand attributes and analysis methods for these endpoints are listed below.

4.3.1.1 Proportion of Participants Entering the Study on Prednisone ≥10 mg/Day Who Achieve Sustained Corticosteroid Control from Week 40 through Week 52

Sustained OCS control is defined as a dose reduction by at least 25% to ≤7.5 mg/day at Week 40, which is maintained through Week 52. Sustained OCS control from Week 40 through Week 52 requires all of the following:

- No daily treatment with prednisone >7.5mg/day (or equivalent) from Week 40 through Week 52
- No receipt of IV, intramuscular, or intra-articular corticosteroids from Week 40 through Week 52

The estimand attributes of the endpoint are defined as follows:

- Population: participants with active, autoantibody-positive SLE as defined by the inclusion criteria and exclusion criteria who are on OCS ≥10 mg/day at baseline
- Endpoint (variable): proportion of participants who achieve Sustained Corticosteroid Control from Week 40 through Week 52
- Intercurrent events: see primary endpoint in Section 4.2.1
- Summary measure: difference in proportion of patients who achieve Sustained Corticosteroid Control from Week 40 through Week 52

The analysis methods are defined as follows:

The proportion of participants who achieve Sustained Corticosteroid Control from Week 40 through Week 52 among those who on a baseline dose of 10 mg or more per day, will be analyzed using the method specified in Section 4.2.2, with only baseline disease activity (SLEDAI-2K score at screening ≥12 vs. <12) as stratification factor.

4.3.1.2 Time to First BILAG Flare Over 52 Weeks

BILAG flare is defined as the occurrence of ≥1 new BILAG Category A or ≥2 new BILAG Category B manifestations from the previous visit.

The estimand attributes of the endpoint are defined as follows:

- Population: see primary endpoint in Section 4.2.1
- Endpoint (variable): time to first BILAG flare from randomization to Week 52

- Treatments: see primary endpoint in Section 4.2.1
- Intercurrent events: For participants who experience treatment failure, major concomitant therapy violation, or lack of efficacy before the first BILAG flare, the participant will be considered as having a flare at the time of the event.

  For participants who experience early treatment discontinuation due to reasons other than lack of efficacy, the observed time to first BILAG flare before early study discontinuation or study completion will be used.

  For participants who experience early study discontinuation due to reasons other than lack of efficacy, the participant will be censored at the time of the event
- Summary measure: hazard ratio and log rank test

The analysis methods are defined as follows:

Time to first BILAG flare will be derived as date of first BILAG flare minus date of randomization. Time to first BILAG flare will be summarized using Kaplan-Meier curves including the number of subjects at risk at each visit. A stratified log-rank test will be used to compare the treatment arms over the 52-week treatment period, adjusting for baseline disease activity and OCS dose. The hazard ratio and its 95% CI will be estimated using a Cox regression model stratified by the stratification factors. Ties in model fitting process will be handled by Efron 1977 estimation.

4.3.1.3 Proportion of Participants Who Achieve SRI(6) at Week 52

The estimand attributes of the endpoint are defined as follows:

- Population: see primary endpoint in Section 4.2.1
- Endpoint (variable): proportion of SRI(6) responders at Week 52
- Treatments: see primary endpoint in Section 4.2.1
- Intercurrent events: see primary endpoint in Section 4.2.1
- Summary measure: difference in proportion of patients who achieve SRI(6) at Week 52

The analysis methods defined as follows:

The proportion of participants achieving SRI(6) at Week 52 will be analyzed using the method described in Section 4.2.2. Both the baseline disease activity (SLEDAI-2K score at screening) and OCS dose will be used in the stratified CMH analysis.

Imputation rules for missing data will follow the same approach described for the primary endpoint (Section 4.2.2).

4.3.1.4 Proportion of Participants Who Achieve BICLA at Week 52

The estimand attributes of the endpoint are defined as follows:

- Population: see primary endpoint in Section 4.2.1
- Endpoint (variable): proportion of participants who achieve BICLA at Week 52
- Treatments: see primary endpoint in Section 4.2.1
- Intercurrent events: see primary endpoint in Section 4.2.1
- Summary measure: difference in proportion of patients who achieve BICLA at Week 52

The analysis methods are defined as follows:

The proportion of participants who achieve BICLA at Week 52 will be analyzed using the method described in Section 4.2.2. Both the baseline disease activity (SLEDAI-2K score at screening) and OCS dose will be used in the stratified CMH analysis. Imputation rules for missing data will follow the same approach described for the primary endpoint (Section 4.2.2).

4.3.1.5 Proportion of Participants Who Achieve a Sustained SRI(4) Response from Week 40 through Week 52

The estimand attributes of the endpoint are defined as follows:

- Population: see primary endpoint in Section 4.2.1
- Endpoint (variable): proportion of participants who achieve Sustained SRI(4) response from Week 40 through Week 52
- Treatments: see primary endpoint in Section 4.2.1
- Intercurrent events: see primary endpoint in Section 4.2.1
- Summary measure: difference in proportion of patients who achieve Sustained SRI(4) response from Week 40 through Week 52

The analysis methods are as follows:

The proportion of participants who achieve Sustained SRI(4) from Week 40 through Week 52 will be analyzed using the method described in Section 4.2.2. Both the baseline disease activity (SLEDAI-2K score at screening) and baseline OCS dose will be used in the stratified CMH analysis.

Imputation rules for missing data will follow the same approach described for the primary endpoint (Section 4.2.2), except that a binary response indicating whether a sustained SRI(4) is achieved will be estimated for each participant.

4.3.2 Supportive Secondary Efficacy Endpoints/Estimands

The definition of the following supportive secondary efficacy endpoints are specified in Table 2. Details of the estimand attributes and analysis methods for these endpoints are listed below.

The estimand attributes for the binary secondary endpoints are defined as follows:

- Population: participants with active, autoantibody-positive SLE as defined by the inclusion criteria and exclusion criteria
- Endpoint (variable): binary secondary endpoint of interest
- Treatments:
  - Experimental: obinutuzumab 1000 mg IV infusion on Day 1 and at Weeks 2, 24, and 26
  - Control: placebo
- Intercurrent events:

  The following intercurrent events will be handled by composite variable strategy (i.e., considered as non-responder):
  - Major concomitant therapy violation
  - Treatment failure requiring rescue medication
  - Early study discontinuation due to death, lack of efficacy, and AE

  The following intercurrent event will be handled by treatment policy strategy:
  - Treatment discontinuation without other concurrent or subsequent intercurrent events (described above)

    If the other intercurrent events occur concurrently with or after treatment discontinuation, the composite variable strategy will be applied (i.e., considered as non-responder).
- Summary measure: difference in proportion of patients at a specific visit of interest

The estimand attributes for the continuous secondary endpoints are defined as follows (unless otherwise specified):

- Population: participants with active, autoantibody-positive SLE as defined by the inclusion criteria and exclusion criteria
- Endpoint (variable): change from baseline for the variable of interest
- Treatments:
  - Experimental: obinutuzumab 1000 mg IV infusion on Day 1 and at Weeks 2, 24, and 26
  - Control: placebo

- Intercurrent events:

  Data after the first occurrence of the following intercurrent event will be imputed using the worst post-baseline value from the following assessment: the last value available prior to the occurrence of the first ICE and all values available after the occurrence of the ICE. The ICE handled by this method include:

  – Major concomitant therapy violation
  – Treatment failure requiring rescue medication
  – Early study discontinuation due to death, lack of efficacy, and AE

  The following intercurrent event will be handled by treatment policy strategy:

  – Treatment discontinuation without other concurrent or subsequent intercurrent events (described above)

    If the other intercurrent events occur concurrently with or after treatment discontinuation, the participant will be imputed using the worst post-baseline value mentioned above.

- Summary measure: difference in adjusted means at a specific visit of interest

Main analytical approaches and imputation methods will be listed below.

4.3.2.1 Proportion of Participants Who Achieve SRI(4) at Week 24

The proportion of participants who achieve SRI(4) at Week 24 will be analyzed using the method described in Section 4.2.2. Imputation rules for missing data will follow LOCF approach at an item level similar to the method outlined in sensitivity analysis for primary endpoint specified in Section 4.2.3 (b).

4.3.2.2 Proportion of Participants Who Achieve Clinical SRI(4) at Week 52 Excluding Laboratory SLEDAI-2K Items

The proportion of participants who achieve clinical SRI(4) (where clinical SRI(4) is defined as SRI(4) excluding consideration of laboratory SLEDAI-2K items) at Week 52 will be analyzed using the method described in Section 4.2.2. Missing data will be imputed using LOCF approach at an item level similar to the method outlined in sensitivity analysis for primary endpoint specified in Section 4.2.3(b).

4.3.2.3 Proportion of Participants Who Achieve SRI(4) at Week 52 on Low-Dose Corticosteroids

The proportion of participants who achieve SRI(4) at Week 52 on low-dose corticosteroids (defined as prednisone $\leq$7.5 mg/day or equivalent) will be analyzed using the method described in Section 4.2.2. Missing data will be imputed using LOCF approach at an item level similar to the method outlined in sensitivity analysis for primary endpoint specified in Section 4.2.3(b).

4.3.2.4 Proportion of Participants Who Achieve Lupus Low Disease Activity State at Week 52

The proportion of participants who achieve Lupus Low Disease Activity State at Week 52 will be analyzed using the method described in Section 4.2.2. Missing data will be imputed using LOCF approach at an item level similar to the method outlined in sensitivity analysis for primary endpoint specified in Section 4.2.3(b).

4.3.2.5 Proportion of Participants Who Achieve DORIS at Week 52

The proportion of participants who achieve Definition of Remission in SLE (DORIS) at Week 52 will be analyzed using the method described in Section 4.2.2. Imputation rules for missing clinical SLEDAI-2K and PGA-VAS score will follow the LOCF approach at an item level similar to the method outlined in sensitivity analysis for primary endpoint specified in Section 4.2.3(b).

4.3.2.6 Change in FACIT-F Scale from Baseline to Week 24 and Baseline to Week 52

Change in Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F) scale from baseline to Week 24 and from baseline to Week 52 will be analyzed using analysis of covariance model. The dependent variable will be the change from baseline at Week 24 or at Week 52, and the independent variables in the model will be baseline value and stratification factors. Missing data in FACIT-F scale will be imputed using LOCF approach.

4.3.2.7 Change in SF-36 v2 Bodily Pain Domain Scale From Baseline to Week 24 and Baseline to Week 52

Change in 36 Item Short Form Survey, Version 2 (SF-36 v2) Bodily Pain domain scale from baseline to Week 24 and from baseline to Week 52 will be analyzed using the method described in Section 4.3.2.5. Missing data in SF-36 v2 Bodily Pain domain scale will be imputed using LOCF approach.

4.3.2.8 Change in SF-36 v2 Physical Component Summary Scale from Baseline to Week 24 and Baseline to Week 52

Change in 36-Item Short Form Survey, Version 2 (SF-36 v2) Physical Component Summary scale from baseline to Week 24 and from baseline to Week 52 will be analyzed using the method described in Section 4.3.2.5. Missing data in SF-36 v2 Physical Component Summary scale will be imputed using LOCF approach.

4.3.2.9 Change in Active Joint Count (Swollen Plus Tender) from Baseline to Week 24 and to Week 52

Change in active joint count from baseline to Week 24 and from baseline to Week 52 will be analyzed using the method described in Section 4.3.2.6. Missing data in active joint count will be imputed using LOCF approach.

4.3.2.10 Proportion of Participants who Achieve a ≥50% Reduction in Active Joint Counts at Each Study Visit

Proportion of participants who achieve a ≥50% reduction in active joint counts (among participants with ≥4 active joints at baseline) will be analyzed at each visit using the method described in Section 4.2.2. Missing data in active joint count will be imputed using LOCF approach.

4.3.2.11 Proportion of Participants who Achieve a ≥50% Reduction in CLASI Total Activity Score at Each Study Visit

Proportion of participants who achieve a ≥50% reduction in Cutaneous Lupus Erythematosus Disease Area and Severity (CLASI) total activity score (among participants with CLASI total activity score ≥10 at baseline) will be analyzed at each visit using the method described in Section 4.2.2. Missing data in CLASI total activity score will be imputed using LOCF approach.

4.3.2.12 Proportion of Participants Who Achieve Sustained Corticosteroid Control from Week 40 through Week 52

Proportion of participants who achieve Sustained Corticosteroid Control from Week 40 through Week 52 will be analyzed using the method described in Section 4.2.2. Corticosteroid dose will be reported as prednisone or equivalent.

4.3.2.13 Cumulative Corticosteroid Use through Week 52

Cumulative corticosteroid dose will be summarized using a Wilcoxon Mann-Whitney test, assuming that the data is non-normally distributed. If data is found to follow a normal distribution by Shapiro-Wilk test, a t-test will be used to compare the difference in two treatment groups.

4.3.2.14 Annualized Flare Rate through Week 52

The annualized flare rate through Week 52 will be compared between treatment and placebo group using a negative binomial model. The number of flares over the 52-week blinded treatment period will be the response variable. Treatment arm and the stratification factors will be included as covariates. Follow-up time (log-transformed) will be used as an offset variable. The treatment coefficient and the 95% CI will be presented.

4.4 EXPLORATORY ENDPOINTS/ESTIMANDS ANALYSIS

The definition of the following exploratory efficacy endpoints are specified in Table 2. Descriptive summary will be provided using the main analytical approaches and imputation methods listed below.

4.4.1 Binary Endpoint

Unless otherwise noted, binary endpoints will be analyzed using the same method detailed in Section 4.2.2. Data will also be summarized using counts and percentages

by treatment group. Missing data for SLEDAI-2K, BILAG-2004, PGA-VAS at different timepoints not preceded by an intercurrent event will be imputed using LOCF method.

4.4.2 Continuous Endpoint

Unless otherwise noted, change from baseline will be analyzed using the same method detailed in Section 4.3.2.5. Data will also be summarized using mean, standard deviation, median, and range by treatment group. Missing data in continuous endpoints not preceded by an intercurrent event will be imputed using LOCF method.

4.4.3 Time to Event Endpoint

Time to event endpoints will be summarized using the method described in Section 4.3.1. Patients who do not experience the events of interest will be censored at the study completion date or early discontinuation date, whichever comes first.

4.5 SAFETY ANALYSES

The safety analyses will be performed on the safety-evaluable set as defined in Section 3, unless otherwise specified. Safety data collected during the blinded treatment period will be presented in the primary Clinical Study Report. Data collected in OLT and SFU period may be presented separately. Safety outputs will include data until the point of rescue for the patients who receive non-corticosteroid rescue therapy. Additional key safety outputs will be produced for post-rescue data.

Safety data will be summarized by obinutuzumab and placebo groups according to the treatment patients actually received.

4.5.1 Extent of Exposure

A summary of exposure to study treatment will be produced using the dose recorded on the electronic Case Report Form. The number of blinded obinutuzumab infusions, treatment duration, and total cumulative dose of obinutuzumab over the 52-week treatment period will be summarized. Obinutuzumab treatment duration will be calculated as difference between date of last dose and date of first dose plus one day.

4.5.2 Adverse Events

All verbatim AE terms will be mapped to Medical Dictionary for Regulatory Activities thesaurus terms, and AE severity will be graded according to National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE), Version 5.0.

Only treatment-emergent AEs will be included in safety analyses, unless otherwise specified. Treatment-emergent AEs will be defined as any newly reported AE or any worsening of an existing condition that occurred on or after the first dose of obinutuzumab or placebo infusion from the time of the participant's randomization through the completion of the study or until early discontinuation.

Adverse events with missing onset date will be considered to be treatment-emergent. Adverse events with partially missing onset date will also be included as treatment-emergent when the month (if it was recorded) and the year occur on or later than the month and year of the study treatment start date.

All AEs will be summarized by system organ class (SOC) and preferred term (PT), and will be presented in order of descending frequency summed across the treatment arms within each SOC and PT. For events of varying severity, the highest grade will be used in the summaries. For each treatment group, the incidence count for each AE PT will be defined as the number of patients reporting at least one treatment-emergent occurrence of the event (multiple occurrences of the same AE in one patient will be counted only once). The proportion of patients with an AE will be calculated as the incidence count divided by the total number of patients in the set. Each table will also present the total number of AEs reported where multiple occurrences of the same AE in an individual are counted separately.

Relevant laboratory, vital signs (pulse rate, respiratory rate, blood pressure, and temperature), and ECG data will be displayed by time, with grades identified where appropriate. A shift table of selected laboratory tests will be used to summarize the baseline and maximum post-baseline severity grade. Changes in vital signs and ECGs will be summarized.

The following will be summarized, and/or listings will be produced where required:

- All AEs
- SAEs
- AEs by relationship to obinutuzumab/placebo
- AEs by relationship to corticosteroids
- AEs leading to death
- AEs by common terminology criteria grade
- Infections
- Serious infections
- AEs leading to blinded obinutuzumab discontinuation
- AEs leading to study withdrawal
- Adverse event of special interest as follows:
  - Cases of potential drug-induced liver injury as defined by Hy's Law
  - Suspected transmission of an infectious agent by a study treatment
  - Infusion-related reactions
  - Grade ≥ 3 infections
  - Any hepatitis B reactivation

- Progressive multifocal leukoencephalopathy
- Study drug-related neutropenia
- Study drug-related thrombocytopenia
- Gastrointestinal perforations
- Worsening of pre-existing cardiac conditions

- Disseminated intravascular coagulation

All summaries and listings of AEs will be specified in the list of planned outputs (LoPO).

4.5.2.1 Laboratory Data

Summary statistics for the absolute and change from baseline values will be produced by visit and treatment group for laboratory parameters.

Laboratory abnormalities will be summarized by visit and treatment group. Summaries by NCI CTCAE grade will also be produced for certain laboratory assessments, including shift tables from baseline to highest grade post-baseline.

Positive pregnancy results will be listed, as identified using the serum test results.

All summaries and listings of laboratory data will be based on the safety populations and specified in the LoPO.

4.6 OTHER ANALYSES

4.6.1 Summaries of Conduct of Study

Enrollment, treatment discontinuation, and early discontinuation from the study will be summarized; reasons for study treatment discontinuation and withdrawal from study will be tabulated; and major protocol deviations, including major deviations with regard to the inclusion and exclusion criteria, will be compiled by treatment arm for the randomized population. A summary of screened participants and reasons for screen failure will be provided.

4.6.2 Summaries of Treatment Group Comparability/Demographics and Baseline Characteristics

Summaries by treatment arm according to the randomization for clinically important demographic and baseline characteristics will be provided. The variables will be summarized using means, standard deviations, medians, and ranges for continuous variables and proportions for categorical variables.

The baseline summary will include the following:

- Age (years)
- Sex
- Weight (kg)

- Race
- Ethnicity
- Region
- Screening SLEDAI-2K score (≥ 12 vs. < 12)
- Baseline OCS dose (≥ 10 mg/day vs. < 10 mg/day)
- Baseline SLEDAI-2K score
- Baseline VAS score
- Baseline BILAG-2004
- Duration of SLE
- Baseline treatment for SLE

4.6.3 Pharmacokinetic Analyses

Serum obinutuzumab concentrations will be summarized (mean, minimum, maximum, SD, and geometric mean) and reported in this study. In addition, non-linear mixed effects modelling will be used to analyze the dose concentration-time data of obinutuzumab. Pharmacokinetic data will be used to refine a previously developed population PK model describing obinutuzumab PK following IV administration to lupus nephritis patients, including the effect of major covariates on the main parameters (e.g., clearance). Derivation of individual measures of exposure such as area under the concentration-time curve and maximum observed concentration will depend on the final model used for the analysis.

The population PK analysis will be preceded by exploratory graphical evaluations of the data and will be used to assess the relationships between obinutuzumab exposure and efficacy, and obinutuzumab exposure and safety. Relevant observed relationships between exposure and safety parameters and exposure and efficacy may be further characterized using different approaches such as logistic regression analyses.

Additional PK and pharmacodynamics (PD) analyses may be conducted as appropriate. Results of the population PK and exposure-response/PD analyses may be reported separately.

4.6.4 Immunogenicity Analyses

The immunogenicity analysis population will consist of all participants with at least one ADA assessment. Participants will be grouped according to treatment received or, if no treatment is received prior to study discontinuation, according to treatment assigned.

The number and proportion of ADA-positive and ADA-negative participants at baseline (baseline prevalence) and after drug administration (post-baseline incidence) will be summarized by study treatment group.

Participants are considered to be ADA-positive at post-baseline incidence if one of the following definitions apply:

- They are ADA-negative or have missing data at baseline but develop an ADA response following study drug exposure (treatment-induced ADA response)

  Or

- They are ADA-positive at baseline and the titer of one or more post-baseline samples with a titer $\geq$0.60 titer unit greater than the titer of the baseline sample (treatment-enhanced ADA response).

Participants are considered to be ADA-negative at post-baseline incidence if one of the following definitions apply:

- They are ADA-negative or have missing data at baseline and all post-baseline samples are negative

  Or

- They are ADA-positive at baseline but do not have any post-baseline samples with a titer $\geq$0.60 titer unit greater than the titer of the baseline sample (treatment unaffected).

The relationship between ADA status and safety, efficacy, PK, and biomarker endpoints may be analyzed and reported via descriptive statistics.

4.7 INTERIM ANALYSES

4.7.1 Planned Interim Analyses

No interim analyses are planned.

4.8 CHANGES TO PROTOCOL-PLANNED ANALYSES

The following changes have been made to protocol-planned analyses:

- In Section 3, the efficacy-evaluable set has been updated to include all randomized patients regardless whether they received study drug or not.
- Duration of SLE and baseline treatment for SLE have been added to the summary of baseline characteristics (Section 4.6.2).

5. SUPPORTING DOCUMENTATION

This section is not applicable, since there is no additional supporting document.

6. REFERENCES

Efron B. The efficiency of Cox's likelihood function of censored data. Journal of the American Statistical Association. 1977; 72, 557–65.

Furie R, Nicholls K, Cheng TT, et al. Efficacy and safety of abatacept in lupus nephritis: a twelve-month, randomized, double-blind study. Arthritis Rheumatol 2014;66:379–89.

Gatto M, Saccon F, Zen M, et al. Early disease and low baseline damage as predictors of response to belimumab in patients with systemic lupus erythematosus in a real-life setting. Arthritis Rheumatol 2020;72:1314–24.

Higgins JPT, Li, Deeks JJ. Chapter 6: Choosing effect measures and computing estimates of effect. In: Higgins JPT, Thomas J and Chandler J et al. Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021). Cochrane, 2021. Available from www.training.cochrane.org/handbook.

Ratitch B, Lipkovich I, O'kelly M. Combining Analysis Results from Multiply Imputed Categorical Data. PharmaSUG 2013 - Paper SP03.

Tamhane AC, Dmitrienka A. Analysis of Multiple Endpoints in Clinical Trials. In: Dmitrienko A, Tamhame AC, Bretz F, editors. Multiple Testing Problems in Pharmaceutical Sciences. Boca Raton: Chapman and Hall/CRC, 2010:131–64.

van Buuren S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. Stat Med 1999;18:681–94.

White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med 2011;30:377–99.

Signature Page for Statistical Analysis Plan - CA42750 - Published
System identifier: RIM-CLIN-524610

Approval Task	[redacted] Company Signatory 15-Mar-2024 14:08:17 GMT+0000