Official Title: A Phase III, Randomized, Double Blind, Placebo-Controlled, Multicenter Study to Evaluate The Efficacy and Safety of Obinituzumab in Patients with ISN/RPS 2003 Class III or IV Lupus Nephritis

NCT Number: NCT04221477

Document Date: SAP Version 3: 19-July-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 ISN/RPS 2003 CLASS III OR IV LUPUS NEPHRITIS**

STUDY NUMBER: CA41705

STUDY NAME: REGENCY

VERSION NUMBER: 3

ROCHE COMPOUND(S): Obinutuzumab (RO5072759)

EUDRACT NUMBER: 2019-004034-42

IND NUMBER: 125054

NCT NUMBER: NCT04221477

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 above

STATISTICAL ANALYSIS PLAN VERSION HISTORY

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

SAP Version	Approval Date	Based on Protocol (Version, Approval Date)
3	See electronic date stamp on the last page of this document	Version 5, 7 February 2024
2	16 May 2023	Version 4, 14 March 2023
1	09 August 2021	Version 3, 23 April 2021

STATISTICAL ANALYSIS PLAN AMENDMENT RATIONALE

Key changes to the SAP, along with the rationale(s) for each change, are summarized below.

Section	Description of Change	Rationale for Change
1.1 & 4	Intercurrent events of the primary and the secondary endpoints have been updated. Study treatment discontinuation has been added as an intercurrent event that will be handled with treatment policy strategy; death has been specified as an intercurrent event that will be handled using composite strategy	In response to health authority feedback on the statistical analysis plan.
1.1.1.3	New exploratory endpoints have been added	To reflect the change in the latest protocol amendment
1.1.1.2	FACIT-F has been promoted to key secondary endpoint from supportive secondary	To reflect the change in the latest protocol amendment
2.1.1	Multiplicity adjustment method has been changed from fixed sequential to fall back method	To improve the chance of meeting the statistical significance for most of the endpoints
3.1.2	Efficacy-evaluable population has been changed to include all randomized patients regardless of whether they received study drug	In response to health authority feedback on the statistical analysis plan.
4.2.4.1	Subgroup analysis based on eGFR has been added	To assess the generalizability of the primary results in different baseline eGFR levels
4.2.4.2.2	Tipping point analysis has been added	To assess the plausibility of missing data assumptions
4.3	Intercurrent events handling strategy of the continuous secondary endpoints have been updated to treatment policy for the intercurrent events except death. Death will be handled using composite strategy	In response to health authority feedback on the statistical analysis plan.

4.3	For the FACIT-F and the continuous secondary endpoints, the missing data imputation method has been updated to multiple imputations and the statistical method have been updated to analysis of covariance (ANCOVA) from mixed models for repeated measures (MMRM)	To improve the missing data imputation technique and the analysis
4.3.1.4	The intercurrent events and missing data imputation has been updated for the death or renal related events through Week 76 endpoint	In response to health authority feedback on the statistical analysis plan.
4.6.6 & 4.6.7	Analysis strategies for post Week 76 efficacy and safety has been added	To include analysis strategies for the post primary endpoint data

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

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation or Term	Description
ADA	anti-drug antibody
AE	adverse event
AESIs	adverse events of special interest
ANCOVA	analysis of covariance
anti-dsDNA	anti–double-stranded DNA
AUC	area under the serum-concentration time curve
CCOD	clinical cutoff date
CI	confidence interval
CKD-EPI	Chronic Kidney Disease Epidemiology Collaboration
CMH	Cochrane-Mantel-Haenszel
CRR	complete renal response
DRB	Data Review Board
eCRF	electronic Case Report Form
eGFR	estimated glomerular filtration rate
EQ-5D-5L	EuroQol 5-Dimension, 5-Level Questionnaire
FACIT-F	Functional Assessment of Chronic Illness Therapy–Fatigue
FCS	fully conditional specification
iDMC	independent Data Monitoring Committee
Ig	Immunoglobulin
IRF	independent review facility
ISN	International Society of Nephrology
IV	Intravenous
IxRx	interactive voice/web-based response system
LN	lupus nephritis
LoPO	List of Planned Outputs
MedDRA	Medical Dictionary for Regulatory Activities
MCE	multivariate imputation by chained equations
mITT	modified intention-to-treat
MMF	mycophenolate mofetil
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
OLT	open-label treatment
ORR	overall renal response
PD	pharmacodynamics
PGA	Physician’s Global Assessment

PK	pharmacokinetic
PRR	partial renal response
PT	preferred term
RBCs	red blood cells
RNP	ribonucleoprotein
RPS	Renal Pathology Society
SAE	serious adverse events
SAP	Statistical Analysis Plan
SFU	study follow-up
SGA	Subject's Global Assessment
SLE	systemic lupus erythematosus
SLEDAI-2K	Systemic Lupus Erythematosus Disease Activity Index 2000
SOC	system organ class
ULN	upper limit of normal
UPCR	Urinary protein-to-creatinine ratio
VAS	visual analog scale

1. INTRODUCTION

Study CA41705 (REGENCY) is a Phase III, parallel-group, double-blind, randomized, placebo-controlled study comparing the efficacy and safety of obinutuzumab versus placebo among patients with International Society of Nephrology/Renal Pathology Society (ISN/RPS) 2003 Class III or IV lupus nephritis (LN) treated with standard-of-care therapy with mycophenolate mofetil (MMF) and corticosteroids. Lupus nephritis is the most common organ-threatening manifestation of systemic lupus erythematosus (SLE) and remains a major cause of morbidity and mortality among patients with SLE.

1.1 OBJECTIVES AND ENDPOINTS

1.1.1 Efficacy Objectives

1.1.1.1 Primary Efficacy Objective

The primary efficacy objective for this study is to evaluate the efficacy of obinutuzumab (combined treatment groups) compared with placebo on the basis of the following endpoint:

- Proportion of patients who achieve a complete renal response (CRR) at Week 76

  CRR is defined as achievement of all of the following:

  - Urinary protein-to-creatinine ratio (UPCR) $<$0.5 g/g
  - Estimated glomerular filtration rate (eGFR) $\geq$85% of baseline, as calculated using the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation
  - No occurrence of the following intercurrent events:

    Rescue therapy, treatment failure, death, or early study withdrawal

1.1.1.2 Secondary Efficacy Objective

The secondary efficacy objective for this study is to evaluate the efficacy of obinutuzumab (combined treatment groups) compared with placebo on the basis of the following endpoints.

Key Secondary Endpoints

- Proportion of patients who achieve a proteinuric response at Week 76. Proteinuric response is defined as achievement of all of the following:
  - UPCR $<$0.8 g/g
  - No occurrence of the following intercurrent events:

    Rescue therapy, treatment failure, death, or early study withdrawal
- Proportion of patients who achieve CRR with successful prednisone taper at Week 76, defined as achievement of CRR (as above) at Week 76 with the following:
  - No receipt of prednisone $>$7.5 mg/day (or equivalent) from Week 64 through Week 76

- Proportion of patients who achieve an overall renal response (ORR), defined as achievement of either CRR or partial renal response (PRR), evaluated at Week 50. PRR is defined as achievement of all of the following:
  - ≥50% reduction in UPCR from baseline
  - UPCR g/g <1 (or <3 if the baseline UPCR was ≥3)
  - eGFR ≥85% of baseline, as calculated using the CKD-EPI equation
  - No occurrence of the following intercurrent events:

    Rescue therapy, treatment failure, death, or early study withdrawal
- Proportion of patients who experience death or renal-related events through Week 76, defined as the proportion of patients with one or more of the following events:
  - Death
  - Treatment failure (see Section 4.4.4.2 of the protocol)
  - Worsening proteinuria, defined as a confirmed ≥50% increase in UPCR to a value ≥3
  - Worsening eGFR, define as a confirmed ≥30% decrease in eGFR to a value <60
- Mean change in eGFR from baseline to Week 76
- Change in Functional Assessment of Chronic Illness Therapy–Fatigue (FACIT-F) scale from baseline to Week 76

Supportive Secondary Endpoints

- Change in anti–double-stranded DNA (anti-dsDNA) titer from baseline to Week 50
- Change in C3 from baseline to Week 50
- Change in Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI2K) from baseline to Week 76
- Time to onset of CRR over the course of 76 weeks
- Proportion of patients who achieve CRR with serum creatinine criteria at Week 76. CRR with serum creatinine criteria is defined as achievement of all of the following:
  - UPCR <0.5 g/g
  - Serum creatinine ≤upper limit of normal (ULN) (as determined by the central laboratory)
  - Serum creatinine not increased from baseline by >25%
  - No occurrence of the following intercurrent events:

    Rescue therapy, treatment failure, death, or early study withdrawal

1.1.1.3 Exploratory Efficacy Objective

The exploratory efficacy objective for this study is to evaluate the efficacy of obinutuzumab (combined treatment groups) compared with placebo on the basis of the following endpoints:

- Proportion of patients who achieve the individual components of CRR at Week 76:
  - UPCR $<$0.5 g/g
  - eGFR $\geq$85% of baseline, as calculated using the CKD-EPI equation
  - No occurrence of the following intercurrent events:

    Rescue therapy, treatment failure, death, or early study withdrawal
- Proportion of patients who achieve CRR with successful prednisone taper at Week 76, defined as achievement of CRR (as above) with the following:
  - No receipt of prednisone $>$7.5 mg/day (or equivalent) from Week 52 through Week 76
- Proportion of patients who achieve CRR at Weeks 24 and 50
- Proportion of patients who achieve proteinuric response at Weeks 24 and 50
- Proportion of patients who achieve ORR at Weeks 24 and 76
- Proportion of patients who achieve CRR on the randomized, blinded therapy at Weeks 106, 132, 158, 184, and 210
- Proportion of patients who achieve proteinuric response on the randomized, blinded therapy at Weeks 106, 132, 158, 184, and 210
- Proportion of patients who achieve ORR on the randomized, blinded therapy at Weeks 106, 132, 158, 184, and 210
- Proportion of patients who achieve CRR with serum creatinine $\leq$ULN at Week 76, as defined as achievement of all of the following:
  - UPCR $<$0.5 g/g
  - Serum creatinine $\leq$ULN (as determined by the central laboratory)
  - No occurrence of the following intercurrent events:
  - Rescue therapy, treatment failure, death, or early study withdrawal
- Proportion of patients who achieve CRR at Week 76, as defined as achievement of all of the following:
  - UPCR $<$0.5 g/g
  - eGFR $\geq$85% of baseline, as calculated using the CKD-EPI equation or $\geq$60 mL/min per 1.73 m^2 of body-surface area
  - No occurrence of the following intercurrent events:

    Rescue therapy, treatment failure, death, or early study withdrawal
- Proportion of patients who achieve the CRR serum creatinine criteria at Week 76

  The CRR serum creatinine criteria require achievement of all of the following:

  - Serum creatinine ≤ULN (as determined by the central laboratory)
  - Serum creatinine not increased from baseline by >25%
  - No occurrence of the following intercurrent events:

    Rescue therapy, treatment failure, death, or early study withdrawal
- Proportion of patients who receive rescue therapy or experience treatment failure by Week 76
- Change in anti-dsDNA titer from baseline to Weeks 4, 12, 24, and 76
- Change in C3 from baseline to Weeks 4, 12, 24, and 76
- Change in C4 from baseline to Weeks 4, 12, 24, 50, and 76
- Change in UPCR from baseline to Weeks 24, 50, and 76
- eGFR slope from Week 12 to Week 76
- Time from first CRR to loss of any response, defined as a failure to meet criteria for either CRR or PRR, during blinded treatment
- Time to LN flare from Week 24, diagnosed if one of the following conditions occurred:
  - eGFR decrease > 20% compared with Week 24 in patients with UPCR >1 g/g and/or cellular casts;
  - UPCR increase (i) to >1 g/g if Week 24 UPCR was <0.2 g/g, (ii) to >2.0 g/g if Week 24 UPCR was 0.2−1 g/g, or (iii) to doubling if Week 24 UPCR was >1 g/g; or
  - Receipt of rescue therapy, except for corticosteroid-only rescue.
- Time to an unfavorable kidney outcome, defined as the first of the following events: treatment failure, serum creatinine doubling, or death
- Proportion of patients who achieve CRR including urinary sediment at Week 76, defined as achievement of CRR (as above) with urinary red blood cells (RBCs) < 10 per high-power field without RBC casts
- Cumulative corticosteroid dose from baseline to Week 76
- Change in health-related quality of life as measured by the 36-Item Short Form Survey, Version 2 (SF-36 v2) from baseline to Week 76
- Change in Physician's Global Assessment from baseline to Weeks 24, 50, and 76
- Change in Subject's Global Assessment from baseline to Weeks 24, 50, and 76

Additionally, descriptive comparisons of the obinutuzumab treatment groups will be performed on the basis of the primary and key secondary endpoints described above.

1.1.2 Safety Objective

The safety objective for this study is to evaluate the safety of obinutuzumab (combined treatment groups) compared with placebo on the basis of the following endpoints:

- Incidence and severity of adverse events (AEs), with severity determined according to National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) v5.0
- Characterization of adverse events of special interest (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

1.1.3 Pharmacokinetic Objective

The pharmacokinetic (PK) objective for this study is to characterize the obinutuzumab PK profile on the basis of the following endpoint:

- Serum concentration of obinutuzumab collected at specified timepoints

1.1.4 Immunogenicity Objectives

The immunogenicity objective for this study is to evaluate the immune response to obinutuzumab on the basis of the following endpoint:

- Prevalence of anti-drug antibodies (ADAs) at baseline and incidence of ADAs posttreatment during the study (only for patients treated with obinutuzumab)

The exploratory immunogenicity objective for this study is to evaluate potential effects of ADAs on the basis of the following measures:

- Relationship between ADA status and efficacy, safety, pharmacodynamic (PD), or PK endpoints

1.1.5 Biomarker Objective

The PD objective for this study is to characterize the obinutuzumab-induced changes in circulating B-cells based on the following endpoint:

- Total peripheral B-cell count at specified timepoints

1.1.6 Health Status Utility Objective

The exploratory health status utility objective for this study is to evaluate health status utility scores of patients treated with obinutuzumab on the basis of the following endpoints:

- Change from baseline in EuroQol 5-Dimension, 5-Level Questionnaire (EQ5D-5L) index-based and visual analog scale (VAS) scores at Weeks 0, 24, 50, and 76

1.2 STUDY DESIGN

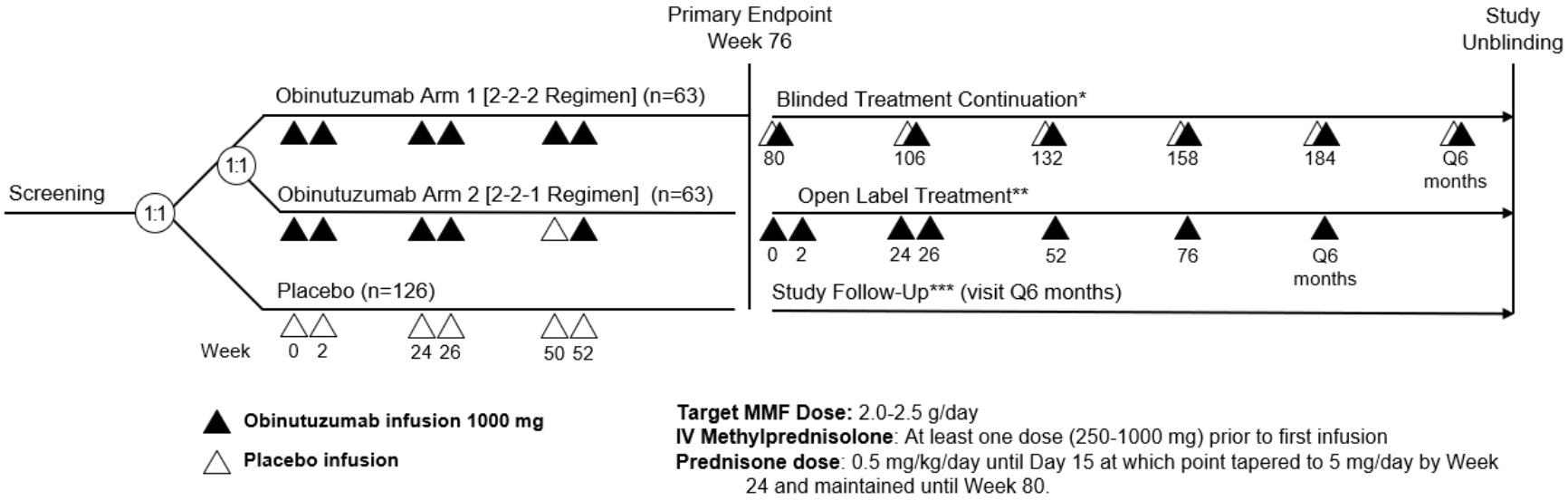
An overview of the study design is provided in Figure 1. The study consists of the following four periods: screening, blinded treatment, open-label treatment (OLT), and study follow-up (SFU).

After a screening period of up to 28 days, randomized patients will receive blinded infusions of obinutuzumab 1000 mg or placebo on Day 1 and Weeks 2, 24, 26, 50, and 52 in three treatment groups according to Figure 1. All patients will receive MMF and corticosteroids as per protocol. The primary endpoint, CRR, will be assessed at Week 76.

Eligibility for subsequent infusions after Week 76 will be determined based on the investigator's assessment of the adequacy of treatment response, need for intensification of therapy, and presence of unmanageable treatment-emergent AEs at Week 76. Patients with an adequate response at Week 76 as defined in the protocol will continue blinded infusions every 6 months. Those with an inadequate response with some improvement from baseline as detailed in the protocol will be eligible for OLT. The randomized treatment assignment will not be revealed during this process, and investigators and patients will remain blinded until study unblinding occurs.

All patients will be followed until the Week 76 assessment and for at least 12 months from the last dose of obinutuzumab or placebo. Patients with continued B-cell depletion may require additional follow-up as defined in the protocol.

Figure 1 Study Schema



MMF mycophenolate mofetil.

Note: The sample sizes shown are the planned patient recruitment numbers.

*Patients with an adequate treatment response at Week 76 continue to receive blinded infusions every 6 months starting at Week 80, until study unblinding.

**Patients with inadequate treatment response at Week 76 or with loss of response during blinded treatment after Week 80 can enter Open Label Treatment.

***Patients are followed through Week 76 and for at least 12 months from the last dose of obinutuzumab or placebo.

1.2.1 Treatment Assignment and Blinding

The investigator or the investigator's research staff will provide patient eligibility information through the interactive Web response system (IxRS) at randomization. Each patient will be randomized and assigned a unique identification number. As confirmation, the investigator will be provided with written verification of each patient's registration.

Patients will be randomized to receive obinutuzumab or placebo in a 1:1 ratio. Patients randomized to receive obinutuzumab will be further randomized to receive one of the two obinutuzumab dosing schedules in a 1:1 ratio:

- Obinutuzumab Arm 1 (2-2-2 Regimen): absolute (flat) dose of 1000 mg IV on Day 1 and Weeks 2, 24, 26, 50, and 52
- Obinutuzumab Arm 2 (2-2-1 Regimen): absolute (flat) dose of 1000 mg IV on Day 1 and Weeks 2, 24, 26, and 52

All patients will receive blinded infusions on Day 1 and Weeks 2, 24, 26, 50, and 52.

The randomization of patients into treatment and control groups will be managed by a central IxRS vendor and will be performed by stratified block design–stratified by region and race as follows:

- Region
  - United States and Canada
  - Latin America and the Caribbean
  - Other
- Race
  - Black
  - Other

These stratification factors were selected given expected differences in response by region and race. LN is clinically heterogeneous in presentation and factors such as availability and intensity of standard of care therapies; socioeconomic status; and ethnicity are known to affect a patient's response to treatment (Tesar et al. 2011). In particular, black race is associated with more aggressive disease (Korbet et al. 2007).

Because it is important to maintain blinding to preserve the integrity of the data collected, all laboratory studies of blood specimens, with unblinding potential, 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, PK results, specific immunoglobulin levels, or ADA results during the study, as listed below, until the primary efficacy and safety analyses through Week 76.

While PK samples must be collected from patients assigned to the comparator arm to maintain the blinding of treatment assignment, PK assay results for these patients are generally not needed for the safe conduct or proper interpretation of this study. Sponsor personnel responsible for performing PK assays will be unblinded to patients' treatment assignments to identify appropriate PK samples to be analyzed. Samples from patients who are assigned to the comparator arm will not be analyzed except by request (e.g., to evaluate a possible error in dosing).

If emergency unblinding is necessary for patient 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, he or she should contact the Medical Monitor directly. The investigator should document and provide an explanation for any premature unblinding (e.g., accidental unblinding, unblinding due to an SAE).

For regulatory reporting purposes and if required by local health authorities, the Sponsor will break the treatment code for all serious, unexpected suspected adverse reactions that are considered by the investigator or Sponsor to be related to study drug.

1.2.2 Independent Review Facility

No Independent Review Facility is planned for this study.

1.2.3 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 safety data, which will be prepared for the committee by an independent Data Coordinating Center (iDCC).

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 patient 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 (combined treatment groups) with placebo on the basis of CRR at Week 76 to demonstrate superiority of obinutuzumab over placebo. The following null and alternative hypothesis will be tested at a two-sided 0.05 significance level:

H_0: $P_{Obinutuzumb} = P_{Placebo}$ versus

H_a: $P_{Obinutuzumb} \neq P_{Placebo}$

for which the $P_{Obinutuzumb}$ and $P_{Placebo}$ refer to proportion of patients who achieve CRR for obinutuzumab (combined treatment groups) and placebo respectively.

2.1.1 Multiplicity Adjustment

Secondary endpoints will be tested to compare obinutuzumab (combined treatment groups) with placebo groups for the superiority of obinutuzumab over placebo. To control the overall type I error, a fallback method maintaining a fixed sequence for testing will be used.

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

1. Proportion of patients who achieve CRR with successful prednisone taper at Week 76
2. Proportion of patients who achieve a proteinuric response at Week 76
3. Change in eGFR from baseline to Week 76

   If the primary endpoint test is significant at two-sided alpha level 0.05, the first key secondary endpoint in the sequence, CRR with successful prednisone taper, will be tested at alpha level 0.05.

   If the CRR with successful prednisone taper is not significant, the testing stops and the endpoints after it in the sequence will be deemed non-significant.

   If the CRR with successful prednisone taper is significant, the alpha 0.05 will be split as 0.04 and 0.01 to the next two endpoints, the proteinuric response and the change in eGFR, respectively.

   If the proteinuric response endpoint is significant at 0.04 alpha level, this alpha is unused and will be passed to the change in eGFR endpoint giving a total alpha for the change in eGFR endpoint test of 0.05 (0.01+0.04). The change in eGFR endpoint test will then be performed at alpha level 0.05.

   If the proteinuric response endpoint is not significant at level 0.04, the change in eGFR endpoint will be tested at the originally reserved alpha of 0.01.

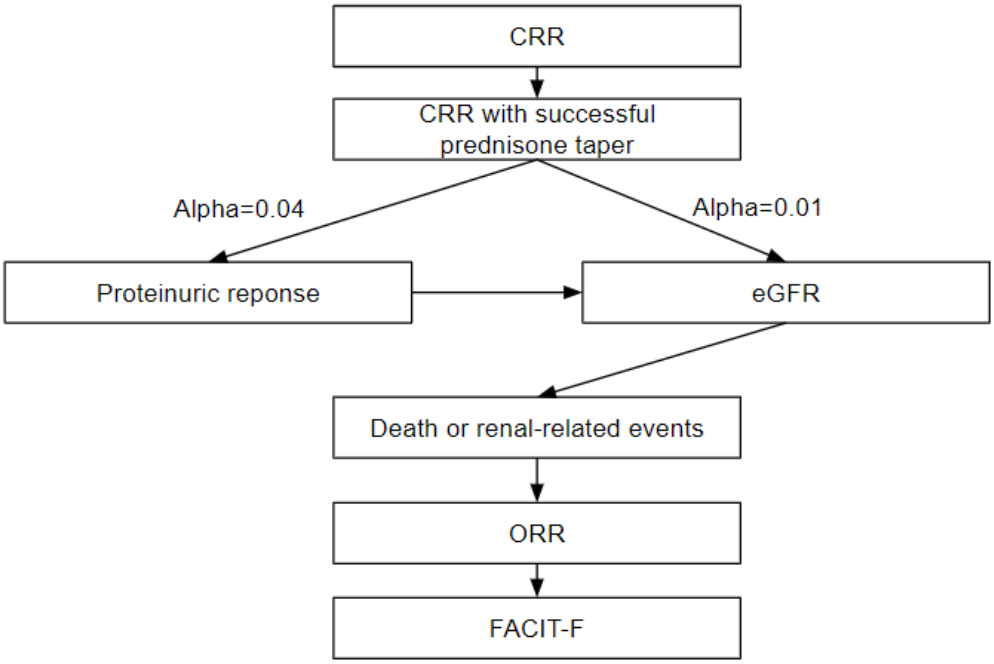
If the change in eGFR is not significant at either 0.05 or 0.01 depending on whether the proteinuric response endpoint was significant, the testing stops.

If the change in eGFR is significant, the unused alpha (either 0.05 or 0.01 depending on whether the proteinuric response endpoint was significant) will be passed to the next endpoints in the sequence and each endpoint will be tested sequentially after achieving the statistical significance on the previous endpoint. Testing stops as soon as there is a failure of an endpoint in the following sequence to show significance. The endpoints after the non-significant endpoint in the sequence will be deemed non-significant.

4. Proportion of patients who experience death or renal-related events through Week 76
5. Proportion of patients who achieve an ORR evaluated at Week 50
6. Change in FACIT-F scale from baseline to Week 76

The fallback method described above is also depicted in Figure 2.

Figure 2 Multiplicity Adjustment Using Fallback Method



2.2 SAMPLE SIZE DETERMINATION

The primary efficacy endpoint of this study is the proportion of patients who achieve CRR. Based on the Phase II NOBILITY trial, it is estimated that approximately 30% of patients with proliferative LN who are receiving MMF will achieve CRR at Week 76 and that the addition of obinutuzumab to MMF will induce an overall CRR rate of 50% at Week 76. On the basis of these assumptions, a total of 252 patients randomized to

obinutuzumab and placebo groups in a 1:1 ratio (126 patients in each of the obinutuzumab- and placebo-treated groups) stratified by region and race will yield approximately 90% power to compare the combined obinutuzumab treatment group with the placebo group at the two-sided $\alpha = 0.05$ significance level using a Cochrane-Mantel-Haenszel (CMH) test, assuming the same CRR proportions across the strata. Patients randomized to the obinutuzumab-treated group will be further randomized to obinutuzumab Arm 1 and obinutuzumab Arm 2 in a 1:1 ratio stratified by region and race.

3. ANALYSIS SETS

3.1 ANALYSIS POPULATIONS

3.1.1 Randomized Population

The randomized population will include all patients randomized into the study.

3.1.2 Efficacy-Evaluable Population

Efficacy-evaluable population will consist of all randomized patients regardless of whether they received study drug. Patients will be grouped according to randomized (assigned) treatment, rather than treatment received. Patients who received an incorrect therapy will be reported under the treatment group to which they were randomized. All efficacy analyses will be performed using the efficacy-evaluable population.

3.1.3 Safety-Evaluable Population

The safety-evaluable population is defined as patients who received any part of blinded infusion of obinutuzumab or placebo. Patients who were randomized to the study but who did not receive any part of blinded infusion of obinutuzumab or placebo will not be included in the safety population. Patients will be grouped according to the treatment that patients actually received rather than the treatment assigned. Patients who received any part of an infusion of obinutuzumab as a study treatment (excluding obinutuzumab infusion received as a rescue therapy) even if not assigned to obinutuzumab treatment group at randomization will be reported under the obinutuzumab treatment group.

All safety analyses will be performed using the safety-evaluable population.

3.1.4 Pharmacokinetic-Evaluable Population

The pharmacokinetic-evaluable population (PK population) will include all patients who have been randomized to and received any dose of obinutuzumab given as study medication, have at least one post-dose PK sample that is evaluable.

4. STATISTICAL ANALYSES

The primary analysis will be performed when all patients have completed the Week 76 visit or have withdrawn early from the study prior to Week 76, all data collected through

Week 76 are in the database, and the data have been cleaned and verified. At the time of the primary analysis, the study will be ongoing. General Consideration

Analyses of the primary and secondary efficacy endpoints will be stratified by race and region unless otherwise noted. The stratification factors as recorded in IxRS will be used.

All the endpoints will be analyzed and all outputs will presented by combined obinutuzumab and placebo treatment group unless otherwise specified. The treatment group in the document refers to combined obinutuzumab versus placebo group.

Visit windows will be continuous from the midpoint between two consecutive study visits, and will be dependent on the schedule of assessments for each variable independently. The exception will be Week 76 visit window. The upper limit of the Week 76 visit window will be 3 days prior to the next obinutuzumab or placebo infusion or 30 days beyond Week 76 whichever is shorter.

As 24-hour UPCR has been collected less frequently than the spot UPCR, the 24-hour UPCR will be windowed based on spot UPCR's schedule of assessments.

If there are multiple values in the same window, values closest to the scheduled study day will be used in the analysis. If two or more values are equidistant from the scheduled study day, the latest assessment will be used. For each patient, the day of first blinded infusion of obinutuzumab or placebo will be designated study Day 1. Each assessment will be assigned a study day calculated as:

$$\textit{date of assessment - date of first dose} + 1$$

Baseline definition: Baseline is defined as the last non-missing value prior to the receipt of blinded infusion of obinutuzumab or placebo. 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.

In general, continuous lab data (except PK data) that are reported as "<LLOQ (Lower Limit of Quantification)" will be set to LLOQ/2 and lab data that are reported as ">ULOQ (Upper Limit of Quantification)" will be set to ULOQ for the analysis. UPCR data reported as "<x.xx" will be set to x.xx.

4.1 Participant Disposition

Enrollment, study withdrawal, reasons for withdrawal from the study, major protocol deviations including major deviations of inclusion/exclusion criteria and major protocol deviations related to COVID-19 will be summarized by treatment group for the randomized population. Blinded obinutuzumab treatment discontinuation and blinded obinutuzumab treatment discontinuation reasons will be summarized by treatment group based on safety population.

4.2 PRIMARY ENDPOINT(S) ANALYSIS

4.2.1 Definition of Primary Endpoint

The primary estimand attributes are as follows:

- Population: Patients with active or active/chronic ISN/RPS 2003 Class III or IV proliferative lupus nephritis
- Endpoint (variable): CRR. A patient will be considered a responder for CRR if the following conditions are met at the Week 76 visit:
  - UPCR <0.5
  - eGFR $\geq 85\%$ of baseline, as calculated using the CKD-EPI equation
  - No occurrence of rescue therapy, treatment failure, death, or early study withdrawal.
- Treatments:
  - Experimental: Obinutuzumab 1000 mg IV infusion at Day 1 and Weeks 2, 24, 26, and either Weeks 50 and 52 or Week 52 only
  - Control: Placebo
- Intercurrent events: rescue therapy, treatment failure, study treatment discontinuation, death, or early study withdrawal prior to Week 76 (refer to Section 4.2.1.1).

  Rescue therapy, treatment failure, death and early study withdrawal are addressed in the endpoint definition and are handled under the composite strategy.

  Study treatment discontinuation will be handled using treatment policy strategy
- Summary measure: Difference in proportions at Week 76

CRR is a composite endpoint derived using several sources of information. Methods for data handling and dealing with missing data for the components of the primary endpoint are provided in Section 4.2.2. Any missing or ambiguous data will be queried, and all efforts will be made to resolve data issues prior to database lock.

4.2.1.1 Intercurrent Events

The intercurrent events include rescue therapy, treatment failure, study treatment discontinuation, death and early study withdrawal.

Rescue therapy: In some cases of worsening or severely active disease, the investigator may conclude that the patient has failed the randomized treatment regimen and requires rescue therapy.

Two categories of rescue therapies will be considered for the study:

- **Rescue therapy excluding corticosteroids**

Rescue therapies in this category include the following:

- Cyclophosphamide

- Anti-CD20 antibodies, including rituximab, ocrelizumab, ofatumumab, and obinutuzumab (if given outside the study protocol)
- Calcineurin inhibitors including ciclosporin, tacrolimus, and voclosporin
- Other investigational, biologic, or targeted therapies used for the treatment of SLE or LN

Patients who receive these rescue therapies will discontinue obinutuzumab or placebo infusions. These patients will continue to be followed according to the study schedule of activities through Week 76 and for at least 12 months from the last dose of obinutuzumab or placebo.

- **Corticosteroid only rescue**

Patients who receive high doses of corticosteroids shortly before the primary endpoint at Week 76 will be considered to have received rescue therapy. Administration of any of the following from Week 64 onward will be considered rescue:

- Methylprednisolone ≥100 mg IV or equivalent
- Prednisone ≥20 mg/day or equivalent (mean dose) over any 2-week period

Any IV or IM corticosteroids will be converted to Methylprednisolone equivalent to evaluate the Methylprednisolone ≥100 mg IV or equivalent criteria. As for evaluating the prednisone ≥20 mg/day or equivalent criteria, oral, IV, and IM corticosteroids will be converted to prednisone equivalent dose.

Methylprednisolone pre-medication given prior to obinutuzumab/placebo infusion as per protocol will be exempted from corticosteroid rescue criteria.

Patients who receive corticosteroid rescue will be treated as non-responders. Patients who receive rescue with corticosteroids alone are not required to discontinue obinutuzumab or placebo infusions.

This same rule related to corticosteroid rescue received within 12 weeks prior to endpoint evaluation will also be applied to other secondary and exploratory renal response endpoints (where applicable), including CRR, PRR, ORR, and proteinuric response at Weeks 24, 50, 76, and post Week 76 timepoints.

Treatment failure: Treatment failure is present if any of the following criteria are met:

- New end-stage renal disease (ESRD)
- Need for chronic dialysis
- Need for renal transplantation
- Clinically significant, sustained worsening in UPCR and/or eGFR from Week 24 onward that leads the investigator to conclude the patient has failed the randomized treatment regimen

- Receipt of rescue therapy, except for corticosteroid only rescue

Patients with treatment failure must discontinue obinutuzumab or placebo infusions and will be non-responders for all subsequent efficacy analyses. These patients will continue to be followed according to the study schedule of activities through Week 76 and for at least 12 months from the last dose of obinutuzumab or placebo.

Early study withdrawal: Early withdrawal from the study before Week 76 for any reason will be considered as an intercurrent event for the primary endpoint and the secondary endpoints evaluated at Week 76.

As for the secondary endpoints evaluated at visits other than Week 76, the early study withdrawal before the endpoint assessment will be considered as an intercurrent event.

Death: Death before Week 76 for any reason will be considered as an intercurrent event for the primary endpoint and the secondary endpoints evaluated at Week 76.

As for the secondary endpoints evaluated at visits other than Week 76, death before the endpoint assessment will be considered as an intercurrent event.

Study treatment discontinuation: Discontinuation from the blinded obinutuzumab infusion before Week 76 for any reason will be considered as an intercurrent event for the primary endpoint and the secondary endpoints evaluated at Week 76. Any patient who discontinues treatment should continue to undergo protocol-mandated visits.

As for the secondary endpoints evaluated at visits other than Week 76, discontinuation from the blinded obinutuzumab before the endpoint assessment will be considered as an intercurrent event.

4.2.2 Main Analytical Approach for Primary Endpoint

The proportions of patients achieving CRR in the obinutuzumab (combined treatment groups) and placebo groups will be compared using the CMH test with region and race as stratification factors. The proportion of patients achieving CRR in the obinutuzumab (combined) and placebo groups will be presented along with the adjusted difference, 95% confidence interval (CI) for the adjusted difference, and p-value. A summary of intercurrent events will be produced as well.

Missing data imputation will be performed by multiple imputations using data from patients who did not experience the intercurrent events that are handled using composite strategy, i.e., rescue therapy, treatment failure, death, and early study withdrawal. Prior to applying multiple imputations, missing 24-hour UPCR at any visit will first be imputed by Spot UPCR. Missing UPCR (when both 24-hour and spot UPCR are missing) and eGFR data for the primary endpoint CRR at Week 76 will be imputed by fully conditional specification (FCS) predicted mean matching method. This is also known as multivariate imputation by chained equations (MICE) and will follow the

method described in White et. al (2011). A set of candidate donors with 5 members is chosen for the predictive mean matching method. Separate imputation models will be used for each treatment arm. Only patients who did not experience the intercurrent events that are handled using the composite strategy, i.e., rescue therapy, treatment failure, death, and early study withdrawal, will be included in the imputation model.

UPCR and eGFR will be imputed using a single imputation model. The imputation model will specify UPCR at Week 76 as the dependent variable and will include eGFR at Week 76 as an auxiliary variable to be imputed as part of the MICE procedure. UPCR and eGFR at Week 24, 36, 50, 64 and 76 will be used in the analysis. As 24-hour UPCR is not collected at Weeks 36 and 64, spot UPCR will be used in the models for these time points.

The model will contain the following variables (in order of being imputed):

- Baseline UPCR
- Baseline eGFR
- Log-serum Albumin at Baseline
- Region (stratification factor)
- Race (stratification factor)
- Sex
- Age
- BMI
- Week 24 UPCR
- Week 24 eGFR
- Week 36 UPCR
- Week 36 eGFR
- Week 50 UPCR
- Week 50 eGFR
- Week 64 UPCR
- Week 64 eGFR
- Log-serum Albumin at Week 76
- Week 76 eGFR
- Week 76 UPCR

The imputation will be implemented using the three standard steps to generate inference from imputed data: imputation step, analysis step, and pooling step. The number of imputations will be set to 100.

1. **Imputation step:** The missing data will be filled in 100 times to generate 100 imputed datasets for the patients who did not experience the intercurrent events that are handled using composite strategy, i.e., rescue therapy, treatment failure, death, and early study withdrawal. A number of burn-in iterations of 100 and a seed of 5678234 will be used. CRR will be derived for each dataset from UPCR and eGFR according to the primary endpoint definition in Section 4.2.1.
2. Patients who experienced the intercurrent events, i.e., rescue therapy, treatment failure, death, and early study withdrawal, will be added in each of the datasets to obtain full datasets including all the patients in the analysis population. These patients will be set non-responder.
3. **Analysis step:** Each of the 100 imputed datasets will be analyzed using the CMH test with region and race as stratification factors.
4. **Pooling step:** The results from the 100 imputed datasets will be combined for inference following the methodology developed by Rubin (Rubin 1987).

A Wilson-Hilferty transformation will be applied to the CMH test statistic for normalization (Ratitch et al. 2013).

In order to pool the adjusted difference in proportions (i.e. common risk difference) calculated for multiple imputed datasets, the required standard error will be obtained from stratified Newcombe CI. The adjusted difference in proportions and the Newcombe CI will be calculated using Mantel-Haenszel weights. The standard error from the Newcombe CI will be calculated as follows:

Standard Error=(Newcombe Upper Limit-Newcombe Lower Limit)/Z
Where Z=3.92 for 95% confidence intervals (Higgins et al. 2021).

4.2.3 Sensitivity Analyses for Primary Endpoint

A sensitivity analysis of the primary endpoint will be performed using the following missing data imputation technique:

eGFR: Missing eGFR at Week 76 will be imputed from the eGFR at Week 64.

UPCR: If the 24-hour UPCR at Week 76 is missing, then it will be imputed in the following order:

- Spot UPCR at Week 76
- Spot UPCR at Week 64

If any of the components for CRR are still missing after the above imputation rules are applied, then the patient will be set to non-responder.

The populations, intercurrent events, and handling of intercurrent events will be the same as in the primary analysis.

4.2.4 Supplementary Analyses for Primary Endpoint

4.2.4.1 Subgroup Analyses for Primary Endpoint

The generalizability of CRR results when comparing obinutuzumab group to placebo group will be investigated by estimating the treatment effect in subgroups based on the following baseline factors:

- Region (United States and Canada vs. Latin America and the Caribbean vs. other)
- Race (black vs. other)
- Sex
- UPCR (≥3 vs. <3)
- anti-dsDNA (>120 IU/mL vs. ≤120 IU/mL)
- C3 (<0.9 g/L vs. ≥0.9 g/L)
- C4 (<0.1 g/L vs. ≥0.1 g/L)
- Class III versus Class IV LN
- Concomitant Class V LN
- Prior history of LN (Yes/No)
- eGFR (<30 vs. 30–<60 vs. 60–<90 vs. ≥90 mL/min/1.73 m2)

Summaries of CRR by these subgroups will be provided in forest plots.

4.2.4.2 Other Supplementary Analyses for Primary Endpoint

4.2.4.2.1 Analysis with Early Study Withdrawals Being Handled as Missing Data

An analysis will be performed with a different strategy to handle early study withdrawals. Early study withdrawal will not be considered as an intercurrent event. Patients who withdraw early from the study will be considered as having missing data after their withdrawal. Statistical analysis and missing data imputation will be performed using the approach described in Section 4.2.2.

The estimand attributes for the analysis are as follows:

- Population: Patients with active or active/chronic ISN/RPS 2003 Class III or IV proliferative lupus nephritis
- Endpoint (variable): CRR. A patient will be considered a responder for CRR if the following conditions are met at the Week 76 visit:
  - UPCR <0.5
  - EGFR ≥85% of baseline, as calculated using the CKD-EPI equation
  - No occurrence of rescue therapy, treatment failure, or death prior to Week 76
- Treatments:

- Experimental: Obinutuzumab 1000 mg IV infusion at Day 1 and Weeks 2, 24, 26, and either Weeks 50 and 52 or Week 52 only
- Control: Placebo

- Intercurrent events: Rescue therapy, treatment failure, study treatment discontinuation, or death prior to Week 76. Rescue therapy, treatment failure, and death are addressed in the endpoint (variable) definition and are handled under the composite strategy.

  Study treatment discontinuation will be handled using treatment policy strategy.
- Summary measure: Difference in proportions at Week 76

4.2.4.2.2 Treatment Policy Strategy

Primary endpoint will be analyzed using treatment policy strategy to handle the intercurrent events: rescue therapy, treatment failure, and study treatment discontinuation. Death will be handled using composite strategy. Patients who withdraw early from the study will be considered as having missing data after their withdrawal. Missing data will be imputed by multiple imputations using the same approach considered in the primary analysis (Section 4.2.2). Only patients who did not experience death will be included in the imputation model.

Tipping Point Analysis

A tipping point analysis will be performed to explore the plausibility of the missing data assumptions (i.e. Missing At Random [MAR]). The TP 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 estimand of the supplementary analysis based on the treatment policy strategy above.

The response values for participants with missing CRR will be imputed deterministically, 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 TP 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 $P\geq0.05$) will be interpreted as the tipping point.

4.3 SECONDARY ENDPOINTS ANALYSES

All the secondary endpoints will be compared between obinutuzumab (combined treatment groups) and placebo groups. Proportions by treatment groups will be presented for binary endpoints along with difference in proportions between treatment groups, two-sided 95% CI for the difference and p-value.

Estimand attributes for the binary secondary endpoints are defined as follows unless otherwise specified:

- Population: patients with active or active/chronic ISN/RPS 2003 Class III or IV proliferative lupus nephritis.
- Endpoint (variable): Binary secondary endpoint of interest. Patients will be considered non-responder if rescue therapy, treatment failure, death, or early study withdrawal prior to endpoint assessment occur.
- Treatments:

  Experimental: Obinutuzumab 1000 mg IV infusion at Day 1 and Weeks 2, 24, 26, and either Weeks 50 and 52 or Week 52 only

  Control: Placebo
- Intercurrent events: Rescue therapy, treatment failure, study treatment discontinuation, death, or early study withdrawal prior to endpoint assessment. Rescue therapy, treatment failure, death and early study withdrawal are addressed in the endpoint (variable) definition and are handled under the composite strategy.

  Study treatment discontinuation will be handled using treatment policy strategy.
- Summary measure: Difference in proportions at a specified visit of interest

Intercurrent events described in Section 4.2.1.1 will be handled in the same approach as the primary endpoint unless otherwise specified. Missing data imputation will be similar to the primary endpoint and discussed in the next sections for each of the binary endpoint.

The continuous endpoints (change from baseline) will be summarized using descriptive statistics. A 95% CI of adjusted mean difference between the treatment groups (combined obinutuzumab vs placebo) will be presented along with a p-value. Estimand attributes for the continuous secondary endpoints are defined as follows:

- Population: patients with active or active/chronic ISN/RPS 2003 Class III or IV proliferative LN
- Endpoint (variable): Change from baseline for the variable of interest
- Treatments:

  Experimental: Obinutuzumab 1000 mg IV infusion at Day 1 and Weeks 2, 24, 26, and either Weeks 50 and 52 or Week 52 only

  Control: Placebo

- Intercurrent events: Rescue therapy, treatment failure, study treatment discontinuation, or death prior to endpoint assessment. Rescue therapy, treatment failure and study treatment discontinuation will be handled using treatment policy strategy.

  Death will be handled using composite strategy. Data after death will be imputed by a pre-specified single fixed value specified in the next sections for each of the continuous endpoints.
- Summary measure: Difference in adjusted means at a specified visit of interest

The intercurrent events are described in Section 4.2.1.1.

4.3.1 Key Secondary Endpoints

4.3.1.1 Proportion of Patients who Achieve a Proteinuric Response at Week 76

The proportion of patients who achieve a proteinuric response at Week 76 will be analyzed using a CMH test with region and race as stratification factors. Proteinuric response is defined as achievement of all of the following:

- UPCR <0.8
- No occurrence of the following intercurrent events:
  - Rescue therapy, treatment failure, death, or early study withdrawal

Missing data: Data obtained from the central laboratory only will be used. UPCR will be imputed by multiple imputations using the model specified in Section 4.2.2. The same imputed datasets produced during the analysis of CRR will be used for the analysis of proteinuric response. The analysis will follow the same methods described in Section 4.2.2.

4.3.1.2 Proportion of Patients who Achieve CRR with Successful Prednisone Taper at Week 76

The proportion of patients who achieve CRR with successful prednisone taper at Week 76 is defined as achievement of CRR (defined as Section 4.2.1) at Week 76 with the following:

- No receipt of prednisone >7.5 mg/day (or equivalent) from Week 64 through Week 76

The endpoint will be analyzed using a CMH test with region and race as stratification factors.

Only oral corticosteroid will be considered for evaluating the no receipt of prednisone >7.5 mg/day (or equivalent) criteria and converted to prednisone equivalent dose when applicable.

Data Handling for Oral Corticosteroids

The study team will monitor oral corticosteroids on an ongoing basis for data quality and to ensure all dose adjustments are being correctly captured. All efforts will be made to resolve missing or ambiguous data prior to database lock. In the case that data queries cannot be resolved, data will be handled in a conservative manner as described in the sections below.

The intervals for each dose of oral corticosteroid captured during the taper are expected to be continuous and non-overlapping. Patients who are adequately controlled by their blinded treatment may discontinue oral corticosteroids altogether, but may re-start treatment as required to treat any increases in disease activity. With this in mind, the use of available data will be maximized and incorporated into the imputation of missing or partial dates as follows:

Missing Term	Imputation Method
Start Date	
Missing Day	Latest of the following: • Previous End Date+1 • First day of the month
Missing Month/Year	Previous End Date+1
End Date	
Missing Day	Earliest of the following: • Next Start Date-1 • Last day of the month
Missing Month/Year	Next Start Date-1

Within two consecutive dosing intervals, if the end date of the first dose interval is missing or partially missing and the start date of the second dose interval is missing or partially missing, a midpoint approach using the start date of the first interval and the end date of the second interval will be used to assign a cut-off end date, providing these dates are available.

A missing or 'ongoing' final end date will be imputed with clinical cutoff date (CCOD) or the date of withdrawal from the study in case a patient withdraw early from the study. A missing initial start date will be imputed with Day 1 providing the end date is after Day 1, otherwise no imputation will be made.

If there is a start and/or end date, but the total daily dose is missing then the dose will be imputed using the dose from the previous dose.

In the event of overlapping dose intervals, the patient will be assumed to be taking the higher dose for the maximum amount of time possible such that the total daily dose will never be double counted.

Missing UPCR and eGFR Data

Missing UPCR and eGFR data will be imputed by multiple imputations using the model specified in Section 4.2.2. The same imputed datasets produced during the analysis of CRR (described in Section 4.2.2) will be used for this analysis. The CRR with successful prednisone taper at Week 76 will be derived for each datasets. The analysis will follow the same methods described in Section 4.2.2.

4.3.1.3 Proportion of Patients who Achieve an ORR, Defined as Achievement of Either CRR or PRR, Evaluated at Week 50

The proportion of patients who achieve an ORR, defined as achievement of either CRR or PRR, evaluated at Week 50 will be analyzed using a CMH test with region and race as stratification factors. The definition of CRR is explained in Section 4.2.1. PRR is defined as achievement of all of the following:

- ≥50% reduction in UPCR from baseline
- UPCR <1 (or <3 if the baseline UPCR was ≥3)
- eGFR ≥85% of baseline, as calculated using the CKD-EPI equation
- No occurrence of the following intercurrent events:
  - Rescue therapy, treatment failure, death, or early study withdrawal

Intercurrent events are described in Section 4.2.1.1.

Missing data: Data obtained from the central laboratory only will be used.

Multiple imputations will be used for the missing data imputation using the model specified in Section 4.2.2. The analysis will follow the same methods described in Section 4.2.2.

4.3.1.4 Proportion of Patients who Experience Death or Renal-related Events through Week 76

Proportion of patients who experience death or renal-related events through Week 76, defined as the proportion of patients with one or more of the following events:

- Death
- Treatment failure (see Section 4.2.1.1)
- Worsening proteinuria, defined as a confirmed ≥50% increase in UPCR to a value ≥3
- Worsening eGFR, define as a confirmed ≥30% decrease in eGFR to a value <60

Worsening proteinuria and eGFR criteria will be evaluated by calculating the specified increase or decrease from the previous visit.

The death or renal related event will occur if any of the above events occur at any visit through Week 76. The proportion of patients who experience death or renal-related

events through Week 76 will be analyzed using a CMH test with region and race as stratification factors.

Estimand attributes for the endpoint are defined as follows:

- Population: patients with active or active/chronic ISN/RPS 2003 Class III or IV proliferative LN.
- Endpoint (variable): Death or renal-related events as defined above.
- Treatments:

  Experimental: Obinutuzumab 1000 mg IV infusion at Day 1 and Weeks 2, 24, 26, and either Weeks 50 and 52 or Week 52 only

  Control: Placebo
- Intercurrent events: Early study withdrawal prior to Week 76 due to lack of efficacy. It will be handled using composite strategy. That is, the early study withdrawals due to lack of efficacy will be considered as events.
- Summary measure: Difference in proportions at Week 76

Early study withdrawals due to reasons other than lack of efficacy will be considered as having missing data.

The missing UPCR and eGFR values will be imputed by the multiple imputation model specified for the primary endpoint in Section 4.2.2 using data from the patients who did not experience death, treatment failure or early study withdrawal due to lack of efficacy.

The death or renal related event endpoint for patients who withdrew early from the study for reasons other than lack of efficacy will then be evaluated based on only the following criteria using the imputed UPCR and eGFR values:

- Worsening proteinuria, defined as a confirmed ≥50% increase in UPCR to a value ≥3
- Worsening eGFR, define as a confirmed ≥30% decrease in eGFR to a value <60

The death or treatment failure will not be imputed.

Worsening proteinuria will be assessed based on 24-hour UPCR which is collected at Week 24, 50 and 76 as per the protocol.

Imputation of eGFR will be made only for the Weeks 24, 36, 50, 64, and 76 as these visits are included in the imputation model (Section 4.2.2). If eGFR values worsened at other missing visits (Weeks 2, 4, 12, 26, and 52) that are not imputed, the worsening due to disease activity should be captured at the next observed or imputed visits. Study withdrawals who have not achieved the event following imputation of their UPCR or eGFR data will be assumed to have not had the event.

The multiple imputation parameters, analysis and pooling steps will be similar to the methods specified in the Section 4.2.2.

4.3.1.5 Mean Change in eGFR from Baseline to Week 76

Change in eGFR from baseline at Week 76 will be analyzed using Analysis of Covariance (ANCOVA) model. The response variable in the model will be change in eGFR from baseline at Week 76. Treatment group (combined obinutuzumab vs. placebo), baseline eGFR and the stratification factors race and region will be included in the model as independent variables. The adjusted difference between the treatment group with 95% CI and p-value will be presented.

The eGFR values for patients who died will be imputed with 0 (composite strategy). All other intercurrent events (specified in Section 4.3) will be handled using treatment policy strategy.

Missing data will be imputed by multiple imputations using FCS predicted mean matching method. A set of candidate donors with 5 members is chosen for the predictive mean matching method. Only patients who did not experience death will be included in the imputation model. Separate imputation models will be used for each treatment group.

The models will contain the following variables (in order of being imputed):

- Baseline eGFR
- Log-serum Albumin at Baseline
- Region (stratification factor)
- Race (stratification factor)
- Sex
- Age
- BMI
- Week 24 eGFR
- Week 36 eGFR
- Week 50 eGFR
- Week 64 eGFR
- Log-serum Albumin at Week 76
- Week 76 eGFR

The imputation will be implemented using the following three standard steps to generate inference from the imputed data. The number of imputations will be set of 100.

1. **Imputation step:**
    - The missing data will be imputed 100 times to generate 100 imputed datasets for the patients who did not experience death.
    - Patients who experienced death will be added in each of the datasets to obtain full datasets including all the patients in the analysis population.
    - A number of burn-in iterations of 100 and a seed of 5678234 will be used.
2. **Analysis step:** Each of the 100 imputed datasets will be analyzed using ANCOVA model
3. **Pooling step:** The results from the 100 imputed datasets will be combined for inference following the methodology developed by Rubin (Rubin 1987).

4.3.1.6 Change in FACIT-F scale from Baseline to Week 76

Change in FACIT-F scale from baseline at Week 76 will be analyzed using ANCOVA model. The response variable in the model will be change in FACIT-F scale from baseline at Week 76. Treatment group, baseline FACIT-F scale and the stratification factors race and region will be included in the model as independent variables. The adjusted difference between the treatment group with 95% CI and p-value will be presented.

The FACIT-F scale for patients who died will be imputed with 0 (composite strategy). All other intercurrent events (specified in Section 4.3) will be handled using treatment policy strategy.

Missing data will be imputed by multiple imputations. Only patients who did not experience death will be included in the imputation model. Multiple imputation methods will be the same as described in Section 4.3.1.5.

The multiple imputation models will contain the following variables (in order of being imputed):

- Baseline FACIT-F score
- Region (stratification factor)
- Race (stratification factor)
- Sex
- Age
- Week 24 FACIT-F score
- Week 50 FACIT-F score
- Week 76 FACIT-F score

4.3.2 Supportive Secondary Endpoints

4.3.2.1 Change in Anti-dsDNA Titer from Baseline to Week 50

Change in anti-dsDNA from baseline at Week 50 will be analyzed with ANCOVA model using the same methods described in Section 4.3.1.5. Anti-dsDNA data will be log-transformed before analyzing. The response variable in the model will be change in log-transformed anti-dsDNA from baseline. Treatment group, baseline log-transformed anti-dsDNA and the stratification factors race and region will be included in the model as independent variables. The adjusted difference between the treatment group with 95% CI and p-value will be presented.

The anti-dsDNA values for patients who died will be imputed with upper limit of quantification (ULOQ, 890 IU/mL) of the analytical method used. All other intercurrent events (specified in Section 4.3) will be handled using treatment policy strategy.

Missing data will be imputed by multiple imputations using the same methods described in Section 4.3.1.5.

The models will contain the following variables (in order of being imputed):

- Baseline log-anti-dsDNA
- Region (stratification factor)
- Race (stratification factor)
- Sex
- Age
- Week 4 log-anti-dsDNA
- Week 12 log-anti-dsDNA
- Week 24 log-anti-dsDNA
- Week 50 log-anti-dsDNA
- Week 76 log-anti-dsDNA

Inclusion of Week 76 value in the model has minimal impact on the Week 50 imputation. The Week 76 anti-dsDNA has been added in the imputation model to facilitate the imputation and analysis of anti-dsDNA at Week 76 which is an exploratory endpoint. Also, to facilitate the analysis of anti-dsDNA at other visits, patients who experienced death will be included in the imputation model, but imputed values after death will be replaced by the ULOQ.

4.3.2.2 Change in C3 from Baseline to Week 50

Change in C3 from baseline at Week 50 will be analyzed using the same method described in Section 4.3.2.1. The response variable in the model will be change in C3 from baseline. Treatment group, baseline C3 and the stratification factors race and region will be included in the model as independent variables.

The C3 values for patients who died will be imputed with LLOQ/2 where LLOQ being the lower limit of quantification. The LLOQ at the central lab is set for C3 at 0.100 g/L. All other intercurrent events (specified in Section 4.3) will be handled using treatment policy strategy.

Missing data will be imputed by multiple imputations using the same methods described in Section 4.3.1.5.

The multiple imputation models will contain the following variables (in order of being imputed):

- Baseline C3
- Region (stratification factor)
- Race (stratification factor)
- Sex
- Age
- Week 4 C3
- Week 12 C3
- Week 24 C3
- Week 50 C3
- Week 76 C3

Inclusion of Week 76 value in the model has minimal impact on the Week 50 imputation. The Week 76 C3 has been added in the imputation model to facilitate the imputation and analysis of C3 at Week 76 which is an exploratory endpoint. Also, to facilitate the analysis of C3 at other visit, patients who experienced death will be included in the imputation model, but imputed values after death will be replaced by the LLOQ/2.

4.3.2.3 Change in SLEDAI-2K from Baseline to Week 76

Change in SLEDAI-2K score from baseline at Week 76 will be analyzed using the same methods as described in Section 4.3.1.5 and Section 4.3.1.

The SLEDAI-2K for patients who died will be imputed with 105 which is the maximum possible SLEDAI-2K value. All other intercurrent events (specified in Section 4.3) will be handled using treatment policy strategy.

The multiple imputation models will contain the following variables (in order of being imputed):

- Baseline SLEDAI-2K
- Region (stratification factor)
- Race (stratification factor)

- Sex
- Age
- Week 24 SLEDAI-2K
- Week 50 SLEDAI-2K
- Week 76 SLEDAI-2K

4.3.2.4 Time to Onset of CRR Over the Course of 76 Weeks

Time to onset of CRR is defined as the time from randomization to the achievement of CRR as defined in Section 4.2.1 for the first time. Based on the schedule of activities, 24-hour UPCR was measured at baseline, Week 24, 50, and 76; Serum creatinine was measured at baseline, Week 2, 4, 12, 24, 26, 36, 50, 52, 64, and 76, through the 76 weeks of treatment visit. The time to onset of CRR over the course of 76 weeks will be either on Week 24, 50, or 76, depending on the timing when CRR is achieved.

Data will be censored as follows:

- Patients who experience the intercurrent events of rescue therapy, treatment failure, death or early study withdrawal before achieving CRR will be censored at Week 76
- Patients completing the 76-week treatment period who do not experience CRR will be censored at Week 76

Time to CRR will be summarized using Kaplan-Meier curves. A stratified log-rank test will be used to compare the treatment groups (combined obinutuzumab vs. placebo) over the 76 weeks treatment period, adjusting for race and region. Median time to event (when estimable), and their 95% CIs, and a p-value from the log-rank test will be presented. The hazard ratio and its 95% CI will be estimated using a Cox regression model stratified by stratification factors. In fitting the Cox model, ties will be handled with the approximate likelihood method of Efron (1977).

Estimand attributes:

- Population: Patients with active or active/chronic ISN/RPS 2003 Class III or IV proliferative LN
- Endpoint (variable): Time to CRR
- Treatments:
  - Experimental: Obinutuzumab 1000 mg IV infusion at Day 1 and Weeks 2, 24, 26, and either Weeks 50 and 52 or Week 52 only
  - Control: Placebo
- Intercurrent events: Rescue therapy, treatment failure, study treatment discontinuation, death, or early study withdrawal.

Patients who experience rescue therapy, treatment failure, death, or early study withdrawal before achieving CRR will be censored at Week 76

Study treatment discontinuation will be handled using the treatment policy strategy. Patients who discontinued study treatment will be followed for CRR.

- Summary measure: Log rank test and hazard ratio

Intercurrent events are described in Section 4.2.1.1 in details.

4.3.2.5 Proportion of Patients who Achieve CRR with Serum Creatinine Criteria at Week 76

The proportion of patients who achieve CRR with serum creatinine criteria at Week 76 will be analyzed using a CMH test with region and race as stratification factors. CRR with serum creatinine criteria is defined as achievement of all of the following:

- UPCR <0.5
- Serum creatinine $\leq$ULN (as determined by the central laboratory)
- Serum creatinine not increased from baseline by $>25\%$
- No occurrence of the following intercurrent events:
  - Rescue therapy, treatment failure, death or early study withdrawal

Intercurrent events are described in Section 4.2.1.1.

Missing data: Data obtained from the central laboratory only will be used. Missing UPCR and serum creatinine data will be imputed by multiple imputations using the model specified in Section 4.2.2 replacing eGFR by serum creatinine in the model. The analysis will follow the same methods described in Section 4.2.2.

4.4 EXPLORATORY ENDPOINTS ANALYSIS

Exploratory endpoints detailed in Section 1.1.1.3 may also be analyzed. Binary endpoints will use similar statistical methods used for the primary endpoint (see Section 4.2.2) where appropriate. Missing data will be imputed using multiple imputations in the same approach as described in Section 4.2.2, where applicable.

For CRR including urinary sediment, RBC, and RBC cast will be imputed from the previous scheduled visit by single imputation similar to the approach described in Section 4.2.3. If RBC cast is missing at the previous visit, it will be imputed as non RBC cast. If RBC is missing at the previous visit, the patient will be considered as non-responder. UPCR and eGFR will be imputed from previous visits as described in Section 4.2.3.

Continuous endpoints will be summarized using descriptive statistics or analyzed using similar statistical methods used for the continuous endpoints in Section 4.2.3, where appropriate.

eGFR slope through Week 76 will be analyzed using mixed effect model with random intercept and random slope.

Time-to-event endpoints will be analyzed similarly to the time to onset of CRR (Section 4.3.2.4), where appropriate. Patients who do not experience the events of interest will be censored at the last known date of not experiencing the events.

Additionally, descriptive comparisons of the obinutuzumab treatment groups (obinutuzumab Arm 1 [2-2-2 Regimen] and obinutuzumab Arm 2 [2-2-1 Regimen]) will be performed on the basis of the primary and the key secondary endpoints.

Further analysis details with post Week 76 blinded treatment period data are provided in Section 4.6.6.

4.5 SAFETY ANALYSES

The safety analyses will be performed on the safety-evaluable population as defined in Section 3.1.3 with data collected in the Week 76 treatment period for each patient. Safety outputs will include data until the point of rescue for the patients who receive rescue therapy (excluding corticosteroid rescue). Limited key safety outputs will be produced for post-rescue data.

Safety data will be summarized by the obinutuzumab (combined treatment groups) and placebo groups. Additionally, combined obinutuzumab treatment group will be further broken down to obinutuzumab arm 1 (2-2-2 regimen) and obinutuzumab arm 2 (2-2-1 regimen) for selected key safety data up to Week 76.

Additional analysis including post Week 76 data collected till CCOD will be produced. The post Week 76 analysis details are provided in Section 4.6.7.

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 (eCRF). The number of blinded obinutuzumab infusions and treatment duration will be summarized. Obinutuzumab treatment duration will be calculated as difference between date of last dose and date of first dose plus one day.

MMF exposure will be summarized using treatment duration and total cumulative dose.

4.5.2 Adverse Events

All verbatim AEs terms will be coded using the latest version of Medical Dictionary for Regulatory Activities (MedDRA).

For safety analyses, unless otherwise specified, only treatment-emergent AEs will be included in the analyses. Treatment-emergent AEs will be defined as any new AEs reported or any worsening of an existing condition on or after the infusion of Obinutuzumab or placebo post randomization (excluding AEs reported in the eCRF as occurred prior to the first study drug administration) through the completion of the study

or until a patient discontinues study prematurely. 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.

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 groups within each SOC and PT. 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 1 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 population. Each table will also present the total number of AEs reported where multiple occurrences of the same AE in an individual are counted separately.

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

- All AEs
- SAEs
- AEs by relationship to obinutuzumab/placebo
- AEs by relationship to MMF
- AEs leading to blinded obinutuzumab discontinuation
- AEs leading to MMF discontinuation
- AEs by NCI CTCAE grade
- Infections
- Serious Infections
- Adverse events of special interest (AESIs) as defined in Section 5.2.3 of the protocol
  - Infusion Related Reactions (IRRs)
  - Grade 3–5 infections
  - Any hepatitis B reactivation or Progressive Multifocal Leukoencephalopathy (PML)
  - Drug-related neutropenia
  - Drug-related thrombocytopenia
  - Gastrointestinal perforations
  - Worsening of pre-existing cardiac conditions
  - Cases of potential drug-induced liver injury that include an elevated ALT or AST in combination with either an elevated bilirubin or clinical jaundice, as defined by Hy's Law

  - Suspected transmission of an infectious agent by the study drug
- AEs leading to study discontinuation
- AEs leading to death
- AEs leading to blinded obinutuzumab dose interruption
- AEs leading to MMF dose modification or interruption
- COVID-19 AEs

For key safety events, the AE rates per 100 patient-years exposure will be calculated with 95% CIs by treatment groups, using all occurrences of adverse events. The rate of AEs per 100 patient years is calculated as:

(Total Number of AEs/Total Number of Patient Years at Risk)×100

All summaries and listings of AEs will be listed in the List of Planned Outputs (LoPO).

4.5.3 Laboratory Data

4.5.3.1 Safety Laboratory Parameters

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 lab assessments, including shift tables from baseline to highest grade post-baseline.

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

4.5.3.2 Immunoglobulins

The absolute values and change from baseline will be summarized by treatment group and visit for total Immunoglobulin (Ig), IgA, IgG, and IgM.

The proportion of patients with a value of IgA, IgG, and IgM which is less than the lower limit of normal per visit will be produced.

4.5.3.3 Lymphocyte Populations

The absolute values and change from baseline will be summarized by treatment group and visit for $CD3^+$ T-cells, $CD3^+CD8^+$ cytotoxic T-cells, $CD3^+CD4^+$ helper T-cells, and $CD16^+CD56^+$ natural killer cells.

4.5.4 Vital Signs

Vital signs are measured at each scheduled visit in accordance with the schedule of assessments (see protocol): baseline, Week 2, Week 4, Week 24, Week 26, Week 50, Week 52, and Week 76. At infusion visits, vital signs should be taken pre-infusion, then every 15 minutes for 1 hour, then every 30 minutes until the end of the infusion, and

within 30 minutes post-infusion. Additional readings may be obtained at the discretion of the investigator.

Changes in vital signs from baseline will be summarized by visit.

4.6 OTHER ANALYSES

4.6.1 Summaries of Treatment Group Comparability/Demographics and Baseline Characteristics

Summaries by obinutuzumab (combined treatment groups) and placebo groups 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:

- Age (years)
- Sex
- Weight (kg)
- Race
- Ethnicity
- Region
- Serum Creatinine
- eGFR
- UPCR
  - 24-hour UPCR
  - Proportion of patients with UPCR ≥ 3
- anti-dsDNA ($>$120 IU/mL vs $\leq$120 IU/mL)
- C3 ($<$0.9 g/L vs. $\geq$0.9 g/L)
- C4 ($<$0.1 g/L vs. $\geq$0.1 g/L)
- Serum albumin
- LN classification
  - Proportion of patients with Class III
  - Proportion of patients with Class IV
  - Proportion of patients with concomitant Class V
- Duration of LN (months)
- Prior history of LN
- SLEDAI-2K

4.6.2 Pharmacokinetic Analyses

Serum obinutuzumab concentrations and observed PK estimates by non-compartmental analysis (NCA) will be summarized (mean, minimum, maximum, SD and geometric mean and geomean CV) and reported in this study. In addition, non-linear mixed effects modeling will be used to analyze the dose-concentration-time data of obinutuzumab. PK data will be used to refine a previously developed population PK (PopPK) model describing obinutuzumab PK following IV administration to LN patients including the effect of major covariates on the main PK parameters (e.g., clearance). Derivation of individual measures of exposure such as area under the serum-concentration time curve (AUC) and maximum observed concentration (C_{max}) will depend on the final model used for the analysis. The PopPK 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 PD analyses may be conducted as appropriate. Results of the population PK and exposure-response/PD analyses may be reported separately.

4.6.3 Immunogenicity Analyses

The immunogenicity analysis will be based on the safety population.

The numbers and proportions of ADA-positive patients and ADA-negative patients at baseline (baseline prevalence) and after drug administration (post baseline incidence) will be summarized by treatment group. When determining post baseline incidence, patients are considered to be ADA positive if they are ADA negative or have missing data at baseline but develop an ADA response following study drug exposure (treatment-induced ADA response), or if they are ADA positive at baseline and the titer of one or more post baseline samples is at least $\geq$4-fold greater than the titer of the baseline sample (treatment-enhanced ADA response). Patients are considered to be ADA negative if they are ADA negative or have missing data at baseline and all post baseline samples are negative, or if they are ADA positive at baseline but do not have any post baseline samples with a titer that is $\geq$4-fold greater than the titer of the baseline sample (treatment unaffected).

Additional analyses to assess the relationship between ADA status and safety, efficacy, PK and biomarker endpoints may be conducted as appropriate. Results of the analyses may be reported separately.

4.6.4 Biomarker Analyses

All PD analyses will be carried out using the safety-evaluable analysis population. The primary PD marker will be peripheral CD19-positive B cell counts. Total CD19-positive B-cell counts via MRB 1.1 and TBNK flow cytometry will be summarized graphically and

descriptively over time by treatment group. The absolute and percent change from baseline for each marker will be computed at each sampling timepoint. Additionally, the number and proportion of patients demonstrating B cell depletion per protocol definition (as defined in Section 5.1.1.7 of the protocol) will be summarized at each timepoint. For MRB1.1 based assessments, B cell depletion will be defined as $<$0.441 cells/μL (LLOQ of central lab assay). Counts of individual B cell subsets via BCP2.2 will be summarized graphically and descriptively over time by treatment group.] Additionally, the absolute and percent change from baseline for individual B cell subsets will be computed at each sampling timepoint.

Additional exploratory analyses may be performed to provide evidence of PD activity in circulation, urine and/or in tissues. These potential PD markers of B cell activity, inflammation, may include but are not limited to, B cell activation factor, anti-dsDNA, C3 and C4 complement. Percent and absolute change from baseline at each sampling timepoint for the above markers may be summarized graphically and descriptively as appropriate by treatment group.

Further exploratory analyses evaluating the possible relationships between biomarkers and clinical responses, safety, PK, or other biomarker endpoints maybe performed and will be defined in a separate biomarker analyses plan. Results of the biomarker analyses may be reported separately.

4.6.5 Health Status Utility Analyses

Change from baseline in EQ-5D-5L VAS scores will be summarized at Week 76. The questions will be summarized using proportions at Week 76.

4.6.6 Post Week 76 Efficacy Analyses

The post-Week 76 efficacy analysis will be performed for the patients who meet the adequate response criteria as assessed by investigator at Week 76 and continue blinded treatment after Week 76. An adequate response is present if all of the following criteria are met (REGENCY, Protocol Section 3.1.2):

- UPCR $<$0.8 g/g or $\geq$50% reduction in UPCR from baseline to subnephrotic levels ($<$3 g/g)
- No deterioration in renal function from baseline (eGFR $\geq$85% of baseline)
- No need for high-dose corticosteroids
- No receipt of rescue therapy or treatment failure

The primary and the key secondary endpoints will be summarized post-Week 76 at Week 106, 132, 158, 184, and 210, by obinutuzumab and placebo groups for the post Week 76 efficacy analysis. This will enable a supportive, exploratory assessment of maintenance of response, acknowledging the limitations of small sample sizes (particularly at later visits), and potential baseline imbalances between groups.

Proportions by treatment groups will be presented for binary endpoints along with difference in proportions between treatment groups, two-sided 95% CI for the difference.

The estimand attributes for the binary endpoints are defined as follows unless otherwise specified:

- Population: Patients with active or active/chronic ISN/RPS 2003 Class III or IV proliferative LN who meet the adequate response criteria at Week 76.
- Endpoint (variable): The primary and key secondary endpoints at Week 106, 132, 158, 184, and 210
- Treatments:
  - Experimental: Obinutuzumab 1000 mg IV infusion start at Day 1 and Week 2, 24, 26, and either Weeks 50 and 52 or Week 52 only, Week 80 and every 6 months thereafter until the study unblinded to investigator and patients
  - Control: Placebo
- Intercurrent events: rescue therapy, treatment failure, study treatment discontinuation, death, early study withdrawal, entering into OLT or entering into SFU prior to endpoint assessment.

Rescue therapy, treatment failure, death, early study withdrawal and entering into the OLT will be handled using a composite strategy.

Study treatment discontinuation will be handled using a treatment policy strategy.

Entering into SFU will be handled using a hypothetical strategy targeting an effect that would occur in the hypothetical scenario in which no patient entered into SFU.

- Summary measure: Difference in proportions at a specified visit of interest. The analysis will only include patients who would reach to the specified visit at the time of CCOD (patients who moved to OLT or SFU period or withdraw from the study prior to endpoint assessment, but would otherwise reach to the assessment visit by CCOD will be included).

Last Observation Carry Forward (LOCF) will be used as the missing data imputation method. Twenty-four-hour UPCR will be imputed in the following order:

1. Spot UPCR at the same visit
2. 24-hour UPCR at the previous visit
3. Spot UPCR at the previous visit

The endpoint proportion of patients who achieve CRR with successful prednisone taper at post-Week 76 weeks will be evaluated as achievement of CRR at the endpoint assessment visit with the following:

- No receipt of prednisone >7.5 mg/day (or equivalent) from 12 week prior to the endpoint assessment visit

The endpoint proportion of patients who experience death or renal-related events through the post-Week 76 visits will be evaluated from the beginning of the study.

The change from baseline for eGFR and FACIT-F post Week 76 will be summarized using descriptive statistics. A 95% CI of mean difference between the treatment groups will be presented.

Estimand attributes for the change from baseline for eGFR and FACIT-F are defined as follows:

- Population: Patients with active or active/chronic ISN/RPS 2003 Class III or IV proliferative LN who meet the adequate response criteria at Week 76.
- Endpoint (variable): Change from baseline for eGFR and FACIT-F at Week 106, 132, 158, 184, and 210.
- Treatments:
  - Experimental: Obinutuzumab 1000 mg IV infusion start at Day 1 and Week 2, 24, 26, and either Weeks 50 and 52 or Week 52 only, Week 80 for every 6 months until the study unblinded to investigator and patients
  - Control: Placebo
- Intercurrent events: rescue therapy, treatment failure, study treatment discontinuation, death, entering into OLT or entering into SFU prior to endpoint assessment. The treatment policy strategy will be used for all the intercurrent events with the exception of death. Data after death will be imputed with 0.

It is noted that the "Clinically significant, sustained worsening in UPCR and/or eGFR from Week 24 onward that leads the investigator to conclude the patient has failed the randomized treatment regimen" criteria of the treatment failure was not evaluated post Week 76.

- Summary measure: Difference in means at a specified visit of interest.

LOCF will be used as missing data imputation method.

The analysis will be based on post Week 76 blinded treatment efficacy-evaluable population as defined below.

Post Week 76 blinded treatment efficacy-evaluable population: Post Week 76 blinded treatment efficacy-evaluable population will consist of patients who continue blinded treatment period post Week 76 and reach at the specified endpoint assessment visits by CCOD. Patients enter the blinded treatment period, but later moved to OLT or SFU period or withdraw from the study prior to endpoint assessment, but would otherwise reach to the assessment visit by CCOD will be included. Patients will be grouped according to randomized (assigned) treatment, rather than treatment received. Patients

who received an incorrect therapy will be reported under the treatment group to which they were randomized.

4.6.7 All Exposure Safety Analyses including Post-Week 76 Data

In addition to the safety analyses detailed in Section 4.5, an All Exposure safety analysis will be performed including post Week 76 data from both blinded treatment and Open-Label Treatment (OLT) period.

The obinutuzumab arm of the All Exposure analysis will comprise of safety data from patients that received any dose of obinutuzumab. Patients that receive any dose of obinutuzumab, whether during the blinded treatment period or post Week 76 via OLT, will contribute safety data from their first dose of obinutuzumab until their last dose +6 months, the CCOD, study withdrawal or receipt of rescue therapy (whichever occurs first):

- Patients randomized to obinutuzumab at Day 1 who achieve adequate or inadequate response at the Week 76 primary analysis will continue to receive obinutuzumab (either blinded or open label). These patients will therefore contribute safety data to the obinutuzumab all exposure arm from their first dose at Day 1 following randomization.
- Patients randomized to placebo at Day 1 who achieve inadequate treatment response at Week 76 will switch to open label obinutuzumab treatment. These patients will therefore contribute safety data to the obinutuzumab all exposure arm following their first dose after switching to obinutuzumab treatment in the OLT period of the trial. Patients that switch from placebo to obinutuzumab will also be re-baselined as of their first dose of obinutuzumab during the OLT period.

The placebo arm of the All Exposure analysis comprises safety data from patients that received any dose of placebo. Patients that received any dose of placebo during the blinded treatment period including post Week 76, will contribute safety data from their first dose of placebo until their last dose+6 months, the CCOD, receipt of obinutuzumab - 1 day or study withdrawal (whichever occurs first):

- Patients randomized to placebo at Day 1 who achieve adequate response at the Week 76 primary analysis will continue to receive placebo during continued blinded treatment period. These patients will therefore contribute safety data to the placebo all exposure arm from their first dose following randomization, until their last placebo dose+6 months, the CCOD or study withdrawal (whichever occurs first)
- Patients randomized to placebo at Day 1 who achieve inadequate response at the Week 76 primary analysis will switch to obinutuzumab treatment during OLT period. These patients will therefore contribute safety data to the placebo all exposure arm from their first dose following randomization up until they switch to obinutuzumab.

The following will be summarized in the All Exposure safety analysis:

- AEs

- SAEs
- AESIs
- Deaths

All exposure safety analysis population: The All exposure safety analysis population is defined as patients who received any part of blinded infusion of obinutuzumab or placebo during blinded treatment or open label treatment period. Patients who were randomized to the study but who did not receive any part of blinded infusion of obinutuzumab or placebo will not be included in the all exposure safety analysis population. Patients will be grouped according to the treatment that patients actually received rather than the treatment assigned.

4.7 INTERIM ANALYSES

No efficacy interim analyses are planned.

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 for censored data. J Am Stat Assoc. 1977 Sep 1;72(359):557-65.

Higgins JPT, Li and 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.

Korbet SM, Schwartz MM and Evans J, et al. Collaborative Study Group. Severe lupus nephritis: racial differences in presentation and outcome. J Am Soc Nephrol. 2007;18:244-54.

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

Rubin DB. Multiple imputation for nonresponse in surveys John Wiley. New York. 1987.

Tesar V and Hruskova Z. Treatment of proliferative lupus nephritis: a slowly changing landscape. Nat Rev Nephrol. 2011;7:96-109.

White, Ian R., Patrick Royston, and Angela M. Wood. "Multiple imputation using chained equations: issues and guidance for practice." Statistics in medicine 30.4 (2011): 377-399.

Appendix 1 Changes to Protocol-Planned Analyses

The time points of the following exploratory endpoints have been updated following the schedule of assessment of the 24-hour UPCR as follows.

- Proportion of patients who achieve CRR at Weeks 24 and 50
- Proportion of patients who achieve proteinuric response at Weeks 24 and 50
- Proportion of patients who achieve ORR at Weeks 24 and 76

Week 36 has been replaced by Week 24 and Week 64 has been removed.

Appendix 2 Psychometric Validation of the FACIT-Fatigue Scale in LN Patients: Analysis Plan

BACKGROUND

Patients with lupus nephritis (LN) experience significant symptoms due to their disease that impact their daily lives. Of the identified core symptoms of LN, fatigue is one of the most prevalent and bothersome to patients (Williams-Hall et al. 2022; Raymond et al. 2021; Eudy et al. 2022; Gallop et al. 2012; Mathias et al. 2018). In studies of patients with active proliferative LN, more than 90% of patients reported experiencing fatigue at a baseline evaluation (Daleboudt et al. 2011; Grootscholten et al. 2007; Arends et al. 2014). Furthermore, in these patients fatigue was generally reported as their most bothersome symptom. The frequency of fatigue has been shown to be similar between patients with and without active LN (Bland et al. 2018). Due to the importance of fatigue in the LN patient experience and the clinical and patient-reported significance of the symptom, it was deemed important to assess the effect of obinutuzumab on fatigue symptoms in LN in the Phase III REGENCY trial. In this study, fatigue was assessed using the FACIT-Fatigue scale. The associated key secondary endpoint is mean change from Baseline at Week 76.

This analysis plan has been developed to provide psychometric validation for the use of the FACIT-fatigue scale in LN patients in order to facilitate USPI inclusion.

OBJECTIVE

- To establish the reliability, validity, and responsiveness of the FACIT-Fatigue scale in LN patients to facilitate potential label inclusion..

The FACIT-Fatigue is presented in Table 1. The FACIT-Fatigue is one of many different FACIT scales that are part of a collection of health-related quality of life (HRQOL) questionnaires targeted to the management of chronic illness and referred to as The FACIT Measurement System (Webster et al. 2003). The FACIT-Fatigue is a 13-item measure that assesses self-reported fatigue and its impact upon daily activities and function (Tennant 2015).

Table 1 FACIT-Fatigue Scale (Version 4)

Health Concept	FACIT Item	Question	Response Options
Fatigue	HI7	I feel fatigued	0 = Not at all 1 = A little bit 2 = Somewhat 3 = Quite a bit 4 = Very Much
	HI12	I feel weak all over	
	An1	I feel listless ("washed out")	
	An2	I feel tired	
Daily / Role Functioning	An3	I have trouble starting things because I am tired	
	An4	I have trouble finishing things because I am tired	
Physical Functioning	An5	I have energy	
Daily / Role Functioning	An7	I am able to do my usual activities	
	An8	I need to sleep during the day	
	An12	I am too tired to eat	
	An14	I need help doing my usual activities	
	An15	I am frustrated by being too tired to do the things I want to do	
Social Functioining	An16	I have to limit my social activity because I am tired	

FACIT-FATIGUE RESPONSE OPTIONS

All 13 items of the FACIT-Fatigue are measured on a Likert scale with 5 response options and range from 0 (Not at all) to 4 (Very Much).

FACIT-FATIGUE SCORING

To score the FACIT-Fatigue, reversals are first performed for the scores on two items (An5 and An7) then all individual items are summed to obtain a score. The sum of the item scores is multiplied by the number of items in the subscale, then divided by the number of items answered. This produces the final score which ranges from 0 to 52. Higher scores indicate less fatigue (Yellen et al.1997).

FACIT-FATIGUE RECALL PERIOD

The recall period of the FACIT-Fatigue is the past 7 days.

DATA

All data will stem from a single pivotal Phase III trial, REGENCY.

ANALYSES TO SUPPORT THE OBJECTIVE

RELIABILITY

Internal consistency - Internal consistency is an estimate of reliability based on the average correlation between items in a scale. This is represented by calculating Cronbach's alpha.

Calculate Cronbach's alpha using data from the 13 scale items in relation to the total FACIT-Fatigue scale score at Baseline. Cronbac's alpha ≥0.7 suggests reasonable internal consistency. In addition, estimates of lambda 2, 4, and omega will be calculated for comparison.

Test-retest reliability - Calculate two-way, mixed effects ICC between FACIT-Fatigue scores derived at Baseline and Week 24. Correlations ≥0.7 indicate reasonable test-retest reliability.

VALIDITY

Construct validity - assesses the degree to which a scale actually measures the construct it was designed to measure. Two estimates of construct validity will be used - convergent validity and discriminant validity.

Convergent validity - Tests if a scale is correlated with other scales that measure the same or related construct. Calculate Spearman's correlation coefficient and Weighted Kappa between:

- FACIT-Fatigue score and SF-36 Vitality, Physical Functioning, Bodily Pain, Role Physical, and General Health domain scores derived at Baseline
- FACIT-Fatigue score and SF-36 Physical Component Summary (PCS) score derived at Baseline

Discriminant validity - Tests if a scale is independent from other scales that measure constructs that differ from the one being considered. Calculate Spearman's correlation coefficient and Weighted Kappa between:

- FACIT-Fatigue score and SF-36 Mental Health, Role Emotional, and Social Functioning domain scores derived at Baseline
- FACIT-Fatigue score and SF-36 Mental Component Summary (MCS) score derived at Baseline

Known groups validity - Tests if a scale can differentiate between groups known to differ in terms of a disease category. Patients will be grouped according to the British Isles of Lupus Assessment Group (BILAG) General domain and BILAG Musculoskeletal domain scores (A, B, or C/D/E). ANOVA will be used to test the mean differences in FACIT-Fatigue scores between groups. Effect sizes (mean difference/SD) will be calculated for the differences between adjacent groups. A second analysis will be conducted using patient groups based on differing scores for Safety of Estrogens in Lupus Erythematosus National Assessment-Systemic Lupus Erythematosus Disease Activity Index (SELENA-SLEDAI), a measure of reduction in global disease activity . Two groups will be formed: a high SELENA-SLEDAI score group (indicating more active disease - SELENA-SLEDAI score ≥10); and a low SELENA-SLEDAI score group (SELENA-SLEDAI <6). Both analyses will be based on Week 76 data.

Content Validity - This form of validity is typically established using qualitative interviews with patients, and sometimes physicians, to determine if a scale's content maps well with the patient's perspective of their disease burden. There is reasonable evidence for content validity in SLE/LN that will be culled from the existing literature.

RESPONSIVENESS OF FACIT-FATIGUE TO CHANGES IN DISEASE ACTIVITY

In this section, the responsiveness of the FACIT-Fatigue score in relation to changes in disease status is assessed. For this analysis, the extent to which important longitudinal changes in Physician's Global Assessment (PGA), Subject's Global Assessment (SGA), and BILAG grade are associated with changes in the FACIT-Fatigue score will be examined. PGA and SGA score changes will be classified as better, worse, or unchanged, with a 30% increase or decrease used as the criterion for classifying meaningful change. BILAG grade change between consecutive assessments will be categorized as more active, stable, or less active. Moving from a BILAG score of D/E to C or vice versa will be classified as stable since those changes are not considered clinically meaningful. ANOVA will be used to compare mean FACIT-Fatigue score change between groups. Both analyses will be based on Week 76 data.

DERIVATION OF MINIMUM CLINICALLY IMPORTANT DIFFERENCE (MCID) - A RESPONDER THRESHOLD FOR ASSESSING WITHIN-PATIENT IMPROVEMENT

In order to derive an MCID for the FACIT-Fatigue scale, the SF-36 Vitality Domain as an anchor will be employed. The Vitality domain has the advantage of assessing a similar construct to the FACIT-Fatigue scale and is built from fairly well-delineated categorical response choices ("all of the time"=1; 'most of the time"=2; "a good bit of the time"=3; "some of the time"=4; "a little of the time"=5; and "none of the time"=6). For this analysis, an SF-36 Vitality domain responder will be defined as a patient who satisfies the following:

How much of the time during the past 7 days…

VT1 - did you feel full of pep? (must score a 3 or less);

VT2 - did you have a lot of energy? (must score 3 or less);

VT3 - did you feel worn out? (must score 4 or higher);

VT4 - did you feel tired? (must score 4 or higher)

The mean improvement in FACIT-Fatigue score (rounded up to the nearest integer) for the subpopulation of SF-36 VT responders between Baseline and Week 76 will serve as the estimated MCID for the FACIT-Fatigue scale. The proportion of subjects from each treatment group who have improved by the MCID value or greater between Baseline and Week 76 will be calculated. A chi-square test will be conducted to evaluate the differences.

REFERENCES

Arends S, Berden JH, Grootscholten C, et al. Induction therapy with short-term high-dose intravenous cyclophosphamide followed by mycophenolate mofetil in proliferative lupus nephritis. Neth J Med 2014; 72: 481–490.

Bland A, Hunziker S, Barraclough M, et al. Motor and cognitive fatigue in SLE is associated with mood and health related quality of life (HRQoL) in patients with SLE: results from the patient reported outcomes in lupus (PRO-Lupus) study. Rheumatol 2018;57:iii93–94.

Daleboudt GMN, Berger SP, Broadbent E, et al. Health related quality of life in patients with systemic lupus erythematosus and proliferative lupus nephritis. Psychol Health Med 2011; 16:393–404.

Eudy A M, Rogers J L, Corneli A, et al. Intermittent and Persistent Type 2 lupus: patient perspectives on two distinct patterns of Type 2 SLE symptoms. Lupus science & medicine 2022;9(1), e000705.

Gallop K, Nixon A, Swinburn P, et al. Development of a conceptual model of health-related quality of life for systemic lupus erythematosus from the patient's perspective. Lupus 2012;21(9):934–943.

Grootscholten C, Snoek FJ, Bijl M, et al. Health-related quality of life and treatment burden in patients with proliferative lupus nephritis treated with cyclophosphamide or azathioprine/methylprednisolone in a randomized controlled trial. J Rheumatol 2007;34:1699–1707.

Mathias S D, Berry P, De Vries J, et al. Patient experience in systemic lupus erythematosus: development of novel patient-reported symptom and patient-reported impact measures. Journal of patient-reported outcomes 2018;2(1):11.

Raymond K, Park J, Joshi A V, et al. Patient Experience With Fatigue and Qualitative Interview-Based Evidence of Content Validation of The FACIT-Fatigue in Systemic Lupus Erythematosus. Rheumatology and therapy 2021;8(1):541–554.

Tennant K. Assessment of fatigue in older adults: the FACIT Fatigue Scale (version 4). Supportive Care in Cancer 2015;23(5):1355–64.

Webster K, Cella D, Yost K. The functional assessment of chronic illness therapy (FACIT) measurement system: Properties, applications, and interpretation. Health and Quality of Life Outcomes 2003;1(79):1–7.

Williams-Hall R, Berry P, Williamson N, et al. Generation of evidence supporting the content validity of SF-36, FACIT-F, and LupusQoL, and novel patient-reported symptom items for use in patients with systemic lupus erythematosus (SLE) and SLE with lupus nephritis (LN). Lupus science & medicine 2022;9(1):e000712.

Yellen S B, Cella D F, Webster K., et al. Measuring fatigue and other anemia-related symptoms with the Functional Assessment of Cancer Therapy (FACT) measurement system. Journal of pain and symptom management 1997;13(2):63–74.

Signature Page for Statistical Analysis Plan - CA41705 - Published
System identifier: RIM-CLIN-546298

Approval Task	[redacted] Company Signatory 19-Jul-2024 15:33:20 GMT+0000