Division	:	Worldwide Development
Information Type	:	Reporting and Analysis Plan (RAP)

Title	:	Reporting and Analysis Plan for BEL115471, A Phase 3/4, Multi-Center, Randomized, Double-Blind, Placebo-Controlled, 52-Week Study to Evaluate the Efficacy and Safety of Belimumab (HGS1006) in Adult Subjects of Black Race with Systemic Lupus Erythematosus (SLE) **Open-label Phase**, End of Study
Compound Number	:	GSK1550188
Effective Date	:	14-FEB-2019

Description:

- This is a Phase 3/4, multi-center, international, randomized, double-blind, placebo-controlled, 52-week study to evaluate the efficacy and safety of belimumab in adult subjects of black race with active systemic lupus erythematosus (SLE). All subjects will receive stable background standard of care therapy throughout the study. Efficacy will be measured by the SLE Responder Index (SRI) at Week 52, defined by a composite SELENA SLEDAI score (Protocol Amendment 2 modified the primary endpoint to use the modified SLEDAI-2K scoring for proteinuria), Physician's Global Assessment (PGA) and BILAG A and B organ domain scores. In addition, oral corticosteroid use, flares and biomarkers (serum immunoglobulins anti-dsDNA, complement, and B cell) will be assessed. Safety will be assessed by adverse events, clinical laboratory evaluations, immunogenicity, and vital signs. In addition, subjects who successfully complete the double-blind phase may enter into an open-label, 6-month extension phase of the study and receive IV belimumab at a dose of 10 mg/kg. This Reporting and Analysis Plan (RAP) prospectively describes the efficacy and safety analyses that will be performed for the **open-label phase** of the study. Details regarding the analyses and summaries for the double-blind phase of the study were described in a separate RAP.

Author(s)

Author	Date
PPD Statistics Leader (Clinical Statistics)	14-FEB-2019

Copyright 2019 the GlaxoSmithKline group of companies. All rights reserved.
Unauthorized copying or use of this information is prohibited.

RAP Team Review Confirmations
(Method: E-mail)

Reviewer	Date
PPD Principal Programmer (Clinical Programming)	DD-MMM-YYYY
PPD Clinical Development Manager (Clinical Development)	DD-MMM-YYYY
PPD Senior Clinical Development Director (Clinical Development)	DD-MMM-YYYY
PPD SERM Senior Director (Safety Evaluation and Risk Management)	DD-MMM-YYYY
PPD Medical Writer (Clinical Development)	DD-MMM-YYYY
PPD Clinical Development Manager (Clinical Development)	DD-MMM-YYYY
PPD Clinical Data Manager (Clinical Data Management)	DD-MMM-YYYY

Clinical Statistics & Clinical Programming Line Approvals
(Method: Signature)

Approver	Date
PPD Statistics Director (Clinical Statistics)	DD-MMM-YYYY
PPD Director Clinical Programming (Clinical Programming)	DD-MMM-YYYY

Approved by: _____ _____
PPD Date
Director Statistics
GSK

Approved by: PPD _____
PPD 13-Feb-2019
Date
Director Clinical Programming
GSK

RAP Team Review Confirmations
(Method: E-mail)

Reviewer	Date
PPD Principal Programmer (Clinical Programming)	DD-MMM-YYYY
PPD Clinical Development Director (Clinical Development)	DD-MMM-YYYY
PPD Senior Clinical Development Director (Clinical Development)	DD-MMM-YYYY
PPD SERM Senior Director (Safety Evaluation and Risk Management)	DD-MMM-YYYY
PPD Medical Writer (Clinical Development)	DD-MMM-YYYY
PPD Clinical Development Manager (Clinical Development)	DD-MMM-YYYY
PPD Clinical Data Manager (Clinical Data Management)	DD-MMM-YYYY

Clinical Statistics & Clinical Programming Line Approvals
(Method: Signature)

Approver	Date
PPD Statistics Director (Clinical Statistics)	DD-MMM-YYYY
PPD Director Clinical Programming (Clinical Programming)	DD-MMM-YYYY

Approved by: PPD
PPD
Director Statistics
GSK

14-Feb-2019
Date

Approved by: PPD
Director Clinical Programming
GSK

Date

RAP Team Review Confirmations
(Method: E-mail)

Reviewer	Date
PPD Principal Programmer (Clinical Programming)	14-FEB-2019
PPD Clinical Development Manager (Clinical Development)	13-FEB-2019
PPD Senior Clinical Development Director (Clinical Development)	14-FEB-2019
PPD SERM Director (Safety Evaluation and Risk Management)	14-FEB-2019
PPD Medical Writer (Clinical Development)	13-FEB-2019
PPD Clinical Development Manager (Clinical Development)	13-FEB-2019
PPD Clinical Data Manager (Clinical Data Management)	14-FEB-2019

Clinical Statistics & Clinical Programming Line Approvals
(Method: Signature)

Approver	Date
PPD Director Statistics (Clinical Statistics)	14-FEB-2019
PPD Director Programming (Clinical Programming)	13-FEB-2019

Approved by: _____ _____
PPD Date
Director Statistics
GSK

Approved by: _____ _____
PPD Date
Director Programming
GSK

CONFIDENTIAL

BEL115471

TABLE OF CONTENTS

PAGE

1. INTRODUCTION

This reporting and analysis plan (RAP) documents the planned analyses for the open-label phase of the BEL115471 study.

This is a phase 3/4, multi-center, international, randomized, double-blind, placebo-controlled study to evaluate the efficacy and safety of 10mg/kg belimumab administered monthly, compared with placebo over a 52-week treatment period (i.e. double-blind phase) in subjects of black race with active systemic lupus erythematosus (SLE) (defined as Safety of Estrogen in Lupus National Assessment Systemic Lupus Erythematosus Disease Activity Index (SELENA SLEDAI) score ≥8). Subjects who complete the initial 52-week double-blind treatment period may choose to enter a 6-month open-label extension and receive IV belimumab at a dose of 10 mg/kg.

The Double-blind phase was reported separately and there is a double-blind phase RAP. This Open-label phase RAP is based upon the following study documents:

[1] Study Protocol Amendment 2 (February 9, 2017)

[2] Final Case Report Form (CRF) (August 2, 2017)

[3] Program Safety Analysis Plan (PSAP) Version 5 (December 13, 2017). Note: for reporting purposes, the most current version of the PSAP and associated MedDRA version at the time of database release (DBR) will be used.

[4] Double-blind Phase RAP (July 19, 2018)

[5] Double-blind Phase CSR (February 01, 2019)

2. SUMMARY OF KEY PROTOCOL INFORMATION

2.1. Changes to the Protocol Defined Statistical Analysis Plan

Changes from the originally planned statistical analysis specified in the protocol and/or DB phase RAP are outlined in Table 1.

Table 1 Changes to Protocol Defined Analysis Plan

Protocol	Reporting & Analysis Plan	
Statistical Analysis Plan	**Statistical Analysis Plan**	**Rationale for Changes**
[1] Day 0 as Baseline/Treatment Start date	Baseline/Treatment Start date will appear as Day 1 in the displays and will be referenced as Day 1 in this document. A table indicating the target open-label study day for each planned visit starting at Day 1 (instead of Day 0) is in Section 12.2.1.	The CDISC standard is to refer to the Baseline/Treatment Start date as Day 1;

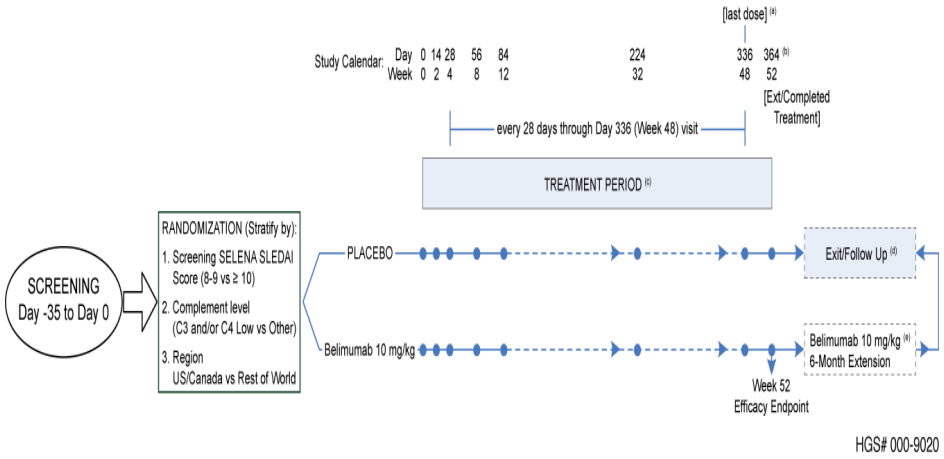
Protocol	Reporting & Analysis Plan	
Statistical Analysis Plan	**Statistical Analysis Plan**	**Rationale for Changes**
[2] Indicates the prednisone analyses will be conducted using dosing on all days between visits	Average daily prednisone dose is based on the 7-day average prednisone dose	Consistent with the DB RAP and reporting of prior Belimumab studies.

2.2. Open-label Study Objective(s) and Endpoint(s)

Objectives	Endpoints
Primary Objectives	**Primary Endpoints**
• To explore the efficacy of belimumab in adult SLE subjects of black race.	• Systemic lupus erythematosus responder index (SRI) response rate with the modified SLEDAI-2K scoring for proteinuria at OL Week 24 (or Week 28/Exit if Week 24 missing)
• To evaluate the safety and tolerability of belimumab in adult SLE subjects of black race.	• Safety will be evaluated by adverse events, changes in laboratory parameters, and immunogenicity.
Secondary Objectives	**Major Secondary Endpoints**
•	• SRI-SS Response rate with the SELENA SLEDAI for scoring of proteinuria at OL Week 24 (or Week 28/Exit if Week 24 missing)
•	• Time to first severe flare (as measured by the modified SLE Flare Index; with SELENA SLEDAI as the SLEDAI criterion of the SFI).
•	• Percent of subjects whose average prednisone dose has been reduced to ≤ 7.5 mg/day at Open-Label Phase Week 28/Exit, in subjects receiving greater than 7.5 mg/day at pre-belimumab baseline.

The Double-blind primary and major secondary analyses are described in the double-blind phase RAP. Analyses of the Open-label phase will be descriptive only.

2.3. Study Design

Overview of Study Design and Key Features	
	
Design Features	(a)The last dose of study agent is given on the Day 337 (Week 48) visit to subjects NOT participating in the 6-month open-label extension phase of the study. (b)Subjects participating in the 6-month open-label extension phase of the study are dosed on the Day 365 (Week 52) visit of the double-blind period after the completion of all Day 365(Week 52) assessments. This Day 365 (Week 52) represents the first dose (i.e., Day 1) of the 6-month open-label extension phase. For subjects not participating in the 6-month open-label extension phase of the study, the Day 365 (Week 52) visit serves as the Exit visit with follow-up visits occurring 8 weeks after the last dose of study agent. (c)The primary treatment period includes 48 weeks of study agent administration (Day 1 to the Day 337 visit) **and** a final visit for the primary endpoint assessment at Week 52 which is 4 weeks after the last dose of the study agent. (d) An Exit visit (1-4 weeks after the last dose of study agent) and a follow-up visit 8 weeks after the last dose of study agent will be performed for subjects withdrawing at any time during the study. The 8-week follow-up visit is not required in subjects entering the separate continuation protocol. (e) At the end of the 6-month extension period, subjects who wish to continue treatment may do so by being prescribed the IV commercially available product. If IV belimumab is not commercially available in a subject's country of participation, subjects may continue to receive belimumab administered intravenously every 4 weeks under a separate continuation protocol. NOTE: This schematic uses Day 0 nomenclature consistent with the protocol. To be consistent with CDISC reporting, add 1 to any Day ≥ 0. For the PLACEBO group, there should be an arrow at the end of the treatment period immediately before the Exit/Follow-up box which points to the Belimumab 10 mg/kg 6-month extension box.
Dosing	• During the open-label phase belimumab 10 mg/kg IV every 28 days for 6 months
Time & Events	• Refer to Appendix 1: Schedule of Activities

Overview of Study Design and Key Features	
Treatment Assignment	• During the open-label extension phase all subjects will receive belimumab 10 mg/kg IV every 28 days for 6 months
Interim Analysis	• No interim analysis is planned for this study.

2.4. Statistical Hypotheses / Statistical Analyses

There is no planned hypothesis testing for the Open-label phase data.

3. PLANNED ANALYSES

3.1. Interim Analyses

No interim analysis is planned for this study.

3.2. Final Analyses

There will be two database locks for this study corresponding to the primary analysis (end of double-blind) and the end of study (end of open-label). The database was locked for the primary analysis (SRI-S2K at Week 52) after data through the Week 52 visit (or Exit visit for those subjects who withdraw during double-blind treatment) for all subjects had been collected, verified and validated. This RAP is for the analysis following the second database lock which will occur after data for the 6-month open-label phase through the 8-week follow-up period have been collected, verified and validated. All subjects and study site personnel are unblinded for the OL phase treatment. All subjects and study site personnel (except the unblinded site pharmacist) remain blinded to the DB treatment until after the Final database lock. This RAP details the analyses for the open-label phase, all outputs required for the open-label phase analyses will be listed in the Mock Tables, Listings and Figures (TLFs) Shells document. A separate RAP was provided for the double-blind phase.

The start of the open-label phase is defined as follows:

- The first dose date of open-label medication.

The final planned open-label phase analyses will be performed after the completion of the following sequential steps:

1. All open-label participants have completed through week 28 of the open-label phase and 8-week follow-up visit if not entering the continuation protocol
2. All required database cleaning activities have been completed and final database release (DBR) and database freeze (DBF) has been declared by Data Management.

In addition to the open-label phase data, any data from the double-blind that was not available at the time of the data cut for the double-blind phase report will be reported.

This includes the following laboratory parameters which were not delivered at the time of the double-blind phase report.

Beta-2-glycoprotein IgA Category
Beta-2-glycoprotein IgG Category
Beta-2-glycoprotein IgM Category
Anti-Ribosomal P
Anti-Ro (SS-A)
Anti-Ro (SS-B)
Anti-RNP
In case of any other changes to DB phase data the DB phase data will be reported in listings.

4. ANALYSIS POPULATIONS

The Double-blind phase analysis sets are defined in the double-blind phase RAP Section 6. The definition of the Safety population from the double-blind phase is repeated here as it will be used for reports on the DB phase data.

Population	Definition / Criteria	Analyses Evaluated
Safety (referred to as ITT in protocol)	• The Safety population is defined as all subjects who are randomized and treated with at least one dose of study treatment in the double-blind phase. The safety population will be summarized according to the treatment that a subject was randomized to receive, regardless of the actual treatment received. The DB phase has completed and all subjects received their randomized treatment for the majority of their administrations. • All double-blind listings will be presented for the safety population unless otherwise specified.	• DB phase safety data
mITT	• During the study three sites (PPD , PPD and PPD) were investigated for potential GCP non-compliance; consequently, all three sites were terminated by the sponsor. A multi-functional team reviewed these concerns and agreed that data from these three sites should be excluded from all efficacy analyses. The team were confident that the subjects had been dosed and therefore agreed data from these sites should be retained in the safety analyses. • The Modified Intention-To-Treat (mITT) population is defined as the safety population excluding subjects who had any assessment at three sites (PPD , PPD or PPD). The database records subjects under final site and so subjects moved to PPD , PPD or PPD , but randomized at other sites are also excluded. Additionally, subjects randomized at one of the three sites but subsequently moved to another site will be excluded from the mITT. All double-blind efficacy analyses were performed on the mITT population.	• Double-blind phase Biomarker data

Population	Definition / Criteria	Analyses Evaluated
PK	• The pharmacokinetic (PK) population will comprise all subjects included in the mITT population for whom at least one post belimumab treatment PK sample was obtained and analyzed. Summaries using this population will be based on the actual treatment received if this differs from that to which the subject was randomized.	• Double-blind phase PK data
Intent-to-Treat Open-label	• All randomized participants who received at least one dose of open-label treatment (i.e. an OL study treatment at DB Week 52/OL Day 1 or a later OL visit). • All open-label listings will be presented for the ITT OL population unless otherwise specified.	• Open-label Safety and listings
Modified Intent-to-Treat Open-label	• The Modified Intent-To-Treat Open-label (mITT OL) population is defined as the Intent to Treat Open-label population excluding subjects who had any assessment at three sites (PPD , PPD or PPD). All open-label efficacy analyses will be performed on the mITT OL population unless otherwise specified.	• Open-label Efficacy
Completer Open-label	• The Completer OL population is defined as all subjects who complete all 28 weeks of the planned open-label treatment period. A summary for the primary efficacy endpoint will be produced for the Completers OL population. This population is a subset of the mITT OL population.	• Open-label Efficacy

Refer to Appendix 9: List of Data Displays which details the population used for each display.

4.1. Protocol Deviations

Important protocol deviations in the Open-label phase (including deviations related to study inclusion/exclusion criteria, conduct of the trial, patient management or patient assessment) will be summarized and listed. Important deviations up to and including the first day of OL treatment were reported in the DB phase CSR, Important PDs on or after the first day of OL treatment will be reported in this OL CSR.

Protocol deviations will be tracked by the study team throughout the conduct of the study in accordance with the Protocol Deviation Management Plan (PDMP): Dated: 10 July 2017 (Version 2.0).

- Data will be reviewed prior to freezing the database to ensure all important deviations are captured and categorized on the protocol deviations dataset.
- This dataset will be the basis for the summaries and listings of Open-label protocol deviations.

5. CONSIDERATIONS FOR DATA ANALYSES AND DATA HANDLING CONVENTIONS

5.1. Study Treatment & Sub-group Display Descriptors

Data Displays for Reporting					
Description	Order in TLF	Color	SAS Color	Line Style	Symbol
Placebo to Belimumab 10mg/kg	3	Purple	CX703070	Solid	Circle (filled)
Belimumab 10mg/kg to Belimumab 10mg/kg	4	Blue	CX0000FF	Dashed	Triangle (filled)

For the Open-label phase subjects will be reported in treatment groups based on their double-blind phase randomized treatment group. In the Open-label treatment group labels the first treatment is the double-blind randomized treatment and the second is the open-label phase treatment. The labels used for Open-label phase reporting will be Placebo to Belimumab 10mg/kg and Belimumab 10mg/kg to Belimumab 10mg/kg. For all tables, a total column for both open-label treatment groups combined will also be presented for subjects in the open-label phase.

5.2. Baseline Definitions

There are three baselines used in the Open-label reports:

Study Baseline

This is the same as the double-blind phase baseline (defined in Section 9.4.1 of the double-blind phase RAP) and is used for study population reports.

OL Baseline (pre-belimumab) (used for efficacy (except time to flare), laboratory values, concomitant medications and vital signs)

The main OL baseline is defined as the last available value prior to the initiation of treatment with belimumab. For subjects switching from placebo to belimumab 10 mg/kg IV in the open-label phase, baseline was defined as the last assessment at the end of the double-blind phase (i.e. Week 52) pre-OL treatment. For subjects that received belimumab 10 mg/kg IV during the double-blind phase and continued to receive belimumab 10 mg/kg IV during the open-label phase, baseline was defined as Day 1 of the double-blind phase (i.e. equivalent to Study baseline and defined as for baseline in section 9.4.1 of the double-blind phase RAP).

Open-label Phase Baseline (used for adverse events, and time to flare)

The open-label phase baseline value of a variable is defined as the last available value measured prior to dosing on or before the date of first open-label dose (Open-label Day

1). If the last available value occurs on Open-label Day 1 but the time of the assessment is not collected, then the assessment will be assumed to be prior to dosing.

For adverse events, these are considered present at open-label baseline if the start date is prior to open-label Day 1 and the end date is on or after open-label Day 1. Events with a start date on open-label Day 1 are considered as being open-label treatment-emergent. Flares that occur on or prior to the open-label treatment start date, are considered as double-blind flares rather than Open-label flares.

5.3. Multicenter Studies

This is a multi-center trial and subjects were centrally randomized at the start of the double-blind phase. Analyses will not be adjusted for center.

5.4. Examination of Covariates, Other Strata and Subgroups

5.4.1. Covariates and Other Strata

Refer to Section 8.2 in the Double-Blind Phase RAP for details of Other Strata and Covariates used in the DB phase.

5.4.2. Examination of Subgroups

Refer to Section 8.3 in the Double-Blind Phase RAP for definitions of subgroups used in the Double-blind phase. Subgroups used for the Open-label phase will be:

- OL phase completers
- Region (US/Canada and Rest of World)
- SLICC/ACR Damage Index Subgroup at Pre-Belimumab Baseline (No Damage and Damage Index >=1)
- Pre-Belimumab Baseline Proteinuria >0.5 g/24hr
- Pre-Belimumab Baseline Prednisone Dose (>7.5 mg/day and <=7.5 mg/day)
- Autoantibody positive at Pre-Belimumab Baseline
- Low complement at Pre-Belimumab Baseline

5.5. Multiple Comparisons and Multiplicity

There is no statistical testing of efficacy endpoints planned for the open-label phase.

5.6. Other Considerations for Data Analyses and Data Handling Conventions

Other considerations for data analyses and data handling conventions are outlined in the appendices:

Section	Component
12.2	Appendix 2: Assessment Windows
12.3	Appendix 3: Study Phases and Treatment Emergent Adverse Events

Section	Component
12.4	Appendix 4: Data Display Standards & Handling Conventions
12.5	Appendix 5: Derived and Transformed Data
12.6	Appendix 6: Reporting Standards for Missing Data
12.7	Appendix 7: Values of Potential Clinical Importance

6. STUDY POPULATION ANALYSES

6.1. Overview of Planned Study Population Analyses

The study population analyses will be based on the ITT Open-label and modified ITT OL population and use the study baseline, unless otherwise specified.

Study population analyses including analyses of subject's disposition, protocol deviations, demographic and baseline characteristics (using study baseline), prior and concomitant medications, and Open-label exposure will be based on GSK Core Data Standards. Details of the planned displays are presented in Appendix 9: List of Data Displays.

7. EFFICACY ANALYSES

7.1. Primary Efficacy Analyses

7.1.1. SRI-S2K Response

The primary endpoint is the systemic lupus erythematosus responder index (SRI) response rate with the modified SLEDAI-2K (S2K) scoring for proteinuria. This S2K rule scores proteinuria as 4 points anytime the value is >0.5 g/24hr. This endpoint will be referred to as the SRI-S2K for reporting and is defined as:

- >= 4-point reduction from OL baseline (pre-belimumab) in SELENA SLEDAI score using the SLEDAI-2K proteinuria scoring [SS-S2K 4pt],

AND

- No worsening (increase of <0.3 points from OL baseline) in Physician's Global Assessment (PGA) [PGA No Worsening],

AND

- No new British Isles Lupus Assessment Group of SLE Clinics (BILAG) A organ domain score or 2 new BILAG B organ domain scores compared with OL baseline at the time of assessment (i.e., at OL Week 24 visit) [BILAG No new 1A/2B].

The following text is instead of Section 9.3.3 in the DB phase RAP.

For the **SRI-S2K response** endpoint and its components (SS-S2K 4-point reduction from OL baseline, PGA no worsening, BILAG no new 1A/2B), observed data will be used. If OL baseline (pre-Belimumab) is missing for any of the components of the SRI-S2K, then

the SRI-S2K will be missing. Subjects who have a SS-S2K score <4 at OL baseline (pre-Belimumab) will be excluded from this analysis as they have no opportunity to meet the responder criteria.

For the SLE Disease Activity Scales these are performed at OL Week 24, or OL Week 28 if not done at OL Week 24. For the OL Week 24/Week 28 timepoint the OL Week 24 values are used for a subject in the first instance and Week 28 only used if Week 24 is missing.

If the partial data of the primary efficacy endpoint at OL baseline are missing (including individual items of any component of the primary endpoint), the LOCF method will be used for the missing item or component. This will be modified for laboratory items in the BILAG for which scoring is dependent on both the actual score and the change from the previous visit as described in Section 9.3.5 of the DB phase RAP.

7.1.2. Population of Interest

The Open-label efficacy summaries will be based on the modified Intent-To-Treat Open-label population and use the OL baseline (pre-Belimumab), unless otherwise specified.

7.1.3. Statistical Analyses / Methods

Details of the planned displays are provided in Appendix 9: List of Data Displays and will be based on GSK data standards and statistical principles.

The endpoints / variables defined in Section 7.1.1 will be summarized using descriptive statistics and listed.

7.1.3.1. Statistical Methodology Specification

Endpoint / Variables
• SRI-S2K Response
Model Specification
• No model will be applied
Model Checking & Diagnostics
• NA
Model Results Presentation
• The number and percentage responding at OL Week 24, OL Week 28 and OL Week 24/Week 28/Exit will be presented.
Subgroup Analyses
• The number and percentage responding by region will be presented.
Sensitivity and Supportive Analyses
• The analysis will be repeated on the subgroup of OL completers. • The 3 component responder endpoints (SS-S2K 4-point reduction from OL baseline (pre-

belimumab), PGA no worsening, BILAG no new 1A/2B) will also be summarized.

- The Disposition of SRI-S2K Response at the Open-Label Phase Week 24/Week 28 will be summarized by treatment group and overall.

7.2. Secondary Efficacy Analyses

7.2.1. Endpoint / Variables

- SRI-SS Response. The SRI will also be derived as described in Section 7.1.1 requiring a 4-point reduction using the SELENA SLEDAI proteinuria scoring as was originally recorded on the SELENA SLEDAI disease activity scale. Apart from proteinuria scoring the SRI-S2K response and the SRI-SS Response are equivalent.
- SRI5-S2K, SRI6-S2K, SRI7-S2K and SRI8-S2K. SRI5-S2K to SRI8-S2K are defined identically to the SRI-S2K except for using higher thresholds of improvement for SS-S2K reduction for a patient to be declared a responder (e.g., SS-S2K ≥5-point reduction for SRI5-S2K). Patients with baseline SS-S2K <5, <6, <7 or <8, will have SRI5-S2K - SRI8-S2K response missing respectively.
- EMA modified SRI-S2K response. This endpoint is defined in Section 9.4.20 of the DB phase RAP except using the OL baseline and for the timepoints OL Week 24, OL Week 28 and OL Week24/Week 28 rather than DB Week 52.
- Time to first severe SFI flare. See Section 12.5.3. Analysis excludes severe flares that were triggered only by an increase in SELENA SLEDAI score to >12. Data censored at last available assessment by Open-Label week 24/28 visit. For subjects who died, data are censored at death if no flares occurred before death. Time to first flare is defined as (event date - Open-Label treatment start date + 1). This endpoint will use the OL Phase baseline.
- Time to first SFI flare. See Section 12.5.3 and information for severe SFI flare above. This endpoint will use the OL Phase baseline.
- SLICC/ACR Damage Index, See Section 9.4.14 in DB phase RAP, except use the OL baseline.
- SLICC/ACR Damage Index Worsening. See Section 9.4.14 in DB phase RAP, except use the OL baseline.
- SELENA SLEDAI and SLEDAI-S2K. See Section 9.4.15 in the DB phase RAP, except use the OL baseline.
- SS-S2K >=4-point reduction. See Section 9.4.15.1 in the DB phase RAP, except use the OL baseline.
- SELENA SLEDAI and SLEDAI-S2K Organ system improvement. See Section 9.4.15 in the DB phase RAP. An improvement is defined as a decrease (compared to OL baseline (pre-Belimumab)) in the SELENA SLEDAI score within the same organ system at a post-Belimumab visit.
- BILAG no New 1A/2B Organ Domain scores compared with OL baseline (Pre-belimumab) at Open-Label Phase Week 24/Week 28. See Section 9.4.16.3 in the DB phase RAP.

- BILAG no 1A/2B Organ Domain scores. See Section 9.4.16.4 in the DB phase RAP, except use the OL baseline.
- BILAG improvement. See Section 9.4.16.1 in the DB phase RAP, except use the OL baseline.
- BILAG worsening. See Section 9.4.16.2 in the DB phase RAP, except use the OL baseline.
- PGA. See Section 9.4.17 of the DB phase RAP, except use the OL baseline.
- PGA No Worsening (Increase of <0.3 points from OL baseline (pre-belimumab)) See Section 9.4.17.1 of the DB phase RAP
- PGA improvement (increase of >=0.3 from OL baseline (pre-Belimumab)). See Section 9.4.17.2 of the DB phase RAP. Subjects with baseline PGA score < 0.3 will be treated as having no improvement.
- Time to first renal flare. See Section 9.4.28 in the DB phase RAP and Section 12.5.3. Time to first renal flare is defined as [date of first renal flare - Open-Label treatment start date + 1]. This endpoint will use the OL phase baseline.
- Proteinuria. See Section 9.4.12 in the DB phase RAP. For Proteinuria shifts from OL baseline (pre-Belimumab) Normal is <=0.5 g/24 hour and High is >0.5 g/24 hour.
- Doubling of serum creatinine (patients whose serum creatinine attains a level double that of the OL baseline (pre-Belimumab) value and is confirmed with a second measurement at least 3 weeks later).
- Change from OL baseline in prednisone (based on 7-day average at the visits). See Section 9.4.13 in the DB phase RAP, except use the OL baseline, and Section 12.5.3.
- Any increase prednisone (based on 7-day average at the visits). See Section 9.4.13 in the DB phase RAP, except use the OL baseline, and Section 12.5.3.
- Prednisone reduced to <=7.5mg/day. Daily prednisone dose reduced to ≤7.5 mg/day from >7.5 mg/day at OL baseline (pre-Belimumab) by visit (based on 7-day average at the visits). See Section 12.5.3.
- Prednisone increased to >7.5mg/day. Daily prednisone dose increased to >7.5 mg/day from ≤7.5 mg/day at OL baseline (pre-Belimumab) by visit (based on 7-day average at the visits). See Section 12.5.3.
- FACIT – Fatigue Scale score. See section 9.4.30 of the DB phase RAP, except use the OL baseline.

7.2.2. Population of Interest

The Open-label efficacy summaries will be based on the modified Intent-To-Treat Open-label population, unless otherwise specified.

7.2.3. Statistical Analyses / Methods

Details of the planned displays are provided in Appendix 9: List of Data Displays and will be based on GSK data standards and statistical principles.

Unless otherwise specified, endpoints / variables defined in Section 7.2.1 will be summarized using descriptive statistics and listed.

7.2.3.1. Statistical Methodology Specification

Endpoint / Variables
• SRI-SS Response • SRI5-S2K to SRI8-S2K • EMA modified SRI-S2K response • SLICC/ACR Damage Index Worsening (Change >0) Compared with OL Baseline (pre-Belimumab) • SS-S2K >= 4-point Reduction from OL Baseline (pre-Belimumab) • BILAG no New 1A/2B Organ Domain scores • BILAG no 1A/2B Organ Domain scores • PGA No Worsening (Increase of <0.3 points from OL Baseline (pre-Belimumab)) • PGA Improvement (Increase of >=0.3 points from OL Baseline (pre-Belimumab)) • Doubling of Serum Creatinine from OL Baseline (pre-Belimumab) • Any Increase in Prednisone Compared to OL Baseline (pre-Belimumab) • Prednisone Reduced to <=7.5 mg/day • Prednisone Increased to > 7.5 mg/day • Improvement in FACIT-Fatigue Scale Score exceeding the MCID (>=4)
Model Specification
• No model will be applied
Model Checking & Diagnostics
• NA
Model Results Presentation
• The number and percentage responding/worsening (xx (xx.x%)) will be presented. • The number of patients 'n' that the percentage is out of will also be presented if this is not equal to N= for the population in the column heading.
Subgroup Analyses
• The SLICC/ACR Damage Index Worsening analyses will be repeated by SLICC/ACR Damage Index Subgroup at OL Baseline (pre-Belimumab) (No Damage vs. Damage Index >=1). • The Doubling of Serum Creatinine from OL Baseline (pre-Belimumab) analyses will be run on the subgroup with OL Baseline (pre-Belimumab) Proteinuria >0.5 g/24hr. • The Prednisone Reduced to <=7.5 mg/day analysis will be run on the subgroup with OL Baseline (pre-Belimumab) Prednisone Dose >7.5 mg/day. • The Prednisone Increased to > 7.5 mg/day analysis will run on the subgroup with OL Baseline (pre-Belimumab) Prednisone Dose <=7.5 mg/day.

Endpoint / Variables
• Time to First Severe SFI Flare • Time to First SFI Flare • Time to First Renal Flare
Model Specification
• PROC LIFETEST will be used to estimate the median, 25th, and 75th percentile of OL days to first (severe) flare
Model Checking & Diagnostics
• NA
Model Results Presentation
• The table will display the number and percentage of subjects with a (severe SFI) flare over 24/28 weeks, the median, 25th, and 75th percentile of days to first (severe) flare. For subjects who experience a (severe) flare, the OL study day of the flare will be summarized and the table will display the median, 25th and 75th percentiles and minimum and maximum.
Sensitivity and Supportive Analyses
• The time to first renal flare analysis will be repeated on the subgroup with OL Phase Proteinuria > 0.5 g/24hr.

Endpoint / Variables
• SLICC/ACR Damage Index Change from OL Baseline (pre-Belimumab)
Model Specification
• The data will be summarized only, no model will be applied. • Only subjects with a OL Baseline (pre-Belimumab) and post Belimumab assessment are included in the analysis.
Model Checking & Diagnostics
• NA
Model Results Presentation
• The number, mean, SD, SE, median and minimum and maximum will be presented. • The OL Baseline (pre-Belimumab), OL Week 28 and Change from OL Baseline (pre-Belimumab) timepoints will be summarized.
Subgroup Analyses
• The analyses will be repeated by SLICC/ACR Damage Index Subgroup at OL Baseline (pre-Belimumab) (No Damage vs. Damage Index >=1).

Endpoint / Variables
• SELENA SLEDAI Organ System Improvement among Subjects with Organ System Involvement at OL Baseline (pre-Belimumab) • BILAG Improvement by Organ Domain among Subjects with an A or B Domain Score at OL

Baseline (pre-Belimumab). • BILAG Worsening by Organ Domain among Subjects with no A Domain Score at OL Baseline (pre-Belimumab).
Model Specification
• The data will be summarized only, no model will be applied.
Model Checking & Diagnostics
• NA
Model Results Presentation
• For SELENA SLEDAI the number with involvement at OL Baseline (pre-Belimumab) will be presented then the number with improvement at OL Week 24, OL Week 28 and OL Week 24/Week 28 and the percentage this is of subjects with involvement at OL Baseline (pre-Belimumab). • For BILAG Improvement the number with A/B at OL Baseline (pre-Belimumab) will be presented then the number with improvement at OL Week 24, OL Week 28 and OL Week 24/Week 28 and the percentage this is of subjects with A/B at OL Baseline (pre-Belimumab). • For BILAG Worsening the number with no A at OL Baseline (pre-Belimumab) will be presented then the number with worsening at OL Week 24, OL Week 28 and OL Week 24/Week 28 and the percentage this is of subjects with no A at OL Baseline (pre-Belimumab).
Subgroup Analyses
• For SELENA SLEDAI for each organ domain separately the Subjects with Organ System Involvement at OL Baseline (pre-Belimumab) are included. • For BILAG Improvement for each organ domain separately the Subjects with Organ Domain A/B at OL Baseline (pre-Belimumab) are included. • For BILAG Worsening for each organ domain separately the Subjects with no A in Organ Domain at OL Baseline (pre-Belimumab) are included.

Endpoint / Variables
• SELENA SLEDAI Percent Change from OL Baseline (pre-Belimumab) • SELENA SLEDAI Change from OL Baseline (pre-Belimumab) • PGA Percent Change from OL Baseline (pre-Belimumab) • PGA Change from OL Baseline (pre-Belimumab) • Proteinuria Percent Change from OL Baseline (pre-Belimumab) • Proteinuria Change from OL Baseline (pre-Belimumab) • Prednisone Change from OL Baseline (pre-Belimumab) • FACIT-Fatigue Scale Score Change from OL Baseline (pre-Belimumab)
Model Specification
• The data will be summarized only, no model will be applied. • For percent change from OL Baseline (pre-Belimumab) subjects with a OL Baseline (pre-Belimumab) score of zero are excluded from the analysis due to division by zero.

Model Checking & Diagnostics
• NA
Model Results Presentation
• The number, mean, SD, SE, median, 25th and 75th percentile and minimum and maximum will be presented. • The OL Baseline (pre-Belimumab) and change from OL Baseline (pre-Belimumab) timepoints will be summarized.
Subgroup Analyses
• The Proteinuria Percent Change from OL Baseline (pre-Belimumab) analyses will be run on the subgroup with OL Baseline (pre-Belimumab) Proteinuria > 0.5 g/24hr.

Endpoint / Variables
• Proteinuria Shifts from OL Baseline (pre-Belimumab)
Model Specification
• The data will be summarized only, no model will be applied.
Model Checking & Diagnostics
• NA
Model Results Presentation
• The number of subjects with OL Baseline (pre-Belimumab) values will be presented, then the number and percentage that are Normal or High at OL Baseline (pre-Belimumab). • For 'Any time post-baseline', the baseline categories Normal and High will be split into n, No Change and Normal to High or High to Normal. • For OL Week 24 and Week 28 the same breakdown will be presented as for 'Any time post baseline'.

8. SAFETY ANALYSES

The safety analyses will be based on the Intent to Treat Open-label population and utilize the open-label baseline, unless otherwise specified.

8.1. Adverse Events Analyses

Adverse events analyses including the analysis of adverse events (AEs), Serious (SAEs) and other significant AEs will be based on GSK Core Data Standards. The details of the planned displays are provided in Appendix 9: List of Data Displays.

8.2. Adverse Events of Special Interest Analyses

The primary source for rules governing identification, adjudication, and reporting of Adverse Events of Special Interest is the Program Safety Analysis Plan (PSAP). AESI are defined using preferred terms from the current version of MedDRA. The intent is to

update these definitions semi-annually using the newest MedDRA version. See Section 18.12 (Appendix 12) of the Double-blind phase RAP for further details. The details of the planned Open-label displays are provided in Appendix 9: List of Data Displays.

8.3. Clinical Laboratory Analyses

Laboratory evaluations including the analyses of Chemistry laboratory tests (Electrolytes, Other Chemistries, Immunoglobins), Hematology laboratory tests, Urinalysis, and liver function tests will be based on GSK Core Data Standards. The details of the planned displays are in Appendix 9: List of Data Displays.

8.4. Other Safety Analyses

The analyses of non-laboratory safety test results including vital signs will be based on GSK Core Data Standards, unless otherwise specified. The details of the planned displays are presented in Appendix 9: List of Data Displays.

9. PHARMACOKINETIC ANALYSES

9.1. Primary Pharmacokinetic Analyses

9.1.1. Endpoint / Variables

9.1.1.1. Drug Concentration Measures

PK data is not collected in the Open-label phase of this study and so no OL tables or listings are planned.

In the DB phase there were 2 cases of PK samples being assigned to the same date. This has since been corrected. Two subjects and 4 samples were affected. A subset of the DB PK reports, i.e. DB listing 5.01 and DB table 5.01, will be re-run using the updated PK data.

10. BIOMARKER ANALYSES

10.1. Biomarker Analyses – Open-label Phase

10.1.1. Endpoint / Variables

- Immunoglobin Levels
    - Immunoglobulin isotype IgG,
    - Immunoglobulin isotype IgA,
    - Immunoglobulin isotype IgM,
- Autoantibody levels
    - Autoantibody anti-dsDNA,
- Complement Levels
    - complement C3
    - complement C4
- B-cells

- CD19 (/uL)
- CD20 (/uL)
- Naive CD19+CD20+CD27- (/uL)
- Naive CD19+CD20+CD27- (%CD19) – only change from baseline table (not percent change)
- Activated CD19+CD20+CD69+ Normalized (COUNT/mL)
- Memory CD19+CD20+CD27+ (/uL)
- Memory CD19+CD20+CD27+ (%CD19) – only change from baseline table (not percent change)
- Plasmacytoid CD19+CD20+CD138+ Normalized (COUNT/mL)
- Plasma CD19+CD20-CD138+ Normalized (COUNT/mL)
- Short-lived Plasma CD19+CD20-CD27b+ Normalized
- SLE Subset CD19+CD38b+CD27b+Lymph Normalized (COUNT/mL)
- Transitional CD19+CD24b+CD38b+CD27- Normalized (COUNT/mL)

10.1.2. Population of Interest

The biomarker analyses will be based on the modified Intent-To-Treat Open-label population, unless otherwise specified.

10.1.3. Statistical Analyses / Methods

Details of the planned displays are provided in Appendix 9: List of Data Displays and will be based on GSK Data Standards and statistical principles.

Unless otherwise specified, endpoints / variables defined in Section 10.1.1 will be summarized using descriptive statistics and listed.

10.1.3.1. Statistical Methodology Specification

Endpoint / Variables
• Immunoglobulin Levels Percent Change from OL Baseline (pre-Belimumab) • Immunoglobulin Levels Change from OL Baseline (pre-Belimumab) • Autoantibody Levels Percent Change from OL Baseline (pre-Belimumab) • Autoantibody Levels Change from OL Baseline (pre-Belimumab) • Complement Levels Percent Change from OL Baseline (pre-Belimumab) • Complement Levels Change from OL Baseline (pre-Belimumab) • B-cells Percent Change from OL Baseline (pre-Belimumab) • B-cells Change from OL Baseline (pre-Belimumab)
Model Specification
• The data will be summarized only, no model will be applied.

Model Checking & Diagnostics
• NA
Model Results Presentation
• The number, mean, SD, SE, median, 25th and 75th percentiles, minimum and maximum will be presented. • The OL Baseline (pre-Belimumab) and Change from OL Baseline timepoints will be summarized.
Subgroup Analyses
• The Autoantibody Levels Percent Change from OL Baseline (pre-Belimumab) and Autoantibody Levels Change from OL Baseline will be repeated for the subgroup positive at OL Baseline. • The Complement Levels Percent Change from OL Baseline (pre-Belimumab) and Complement Levels Change from OL Baseline will be repeated for the subgroup with Low complement at OL Baseline.

Endpoint / Variables
• Immunoglobulin Levels Shifts from OL Baseline (pre-Belimumab) • Autoantibody Levels Shifts from OL Baseline (pre-Belimumab) • Complement Levels Shifts from OL Baseline (pre-Belimumab)
Model Specification
• The data will be summarized only, no model will be applied.
Model Checking & Diagnostics
• NA
Model Results Presentation
• The number of subjects with OL Baseline (pre-Belimumab) values will be presented, then the number and percentage that are Low or Normal/High at OL Baseline (pre-Belimumab). • For OL Week 24 and Week 28, the Baseline Categories Low and Normal/High will be split into n, Low to Normal/High, Low to Low, Normal/High to Normal/High, and Normal/High to Low (this is the breakdown using the levels for Immunoglobin and Complement). • The levels for the Autoantibodies are Positive or Negative, and for the shifts with no change the label “No change” will be used.

10.2. Biomarker Analyses – for the Double-blind Phase

These parameters were not available to be reported in the DB CSR so will be summarized for the DB period in this End of study RAP/Report.

10.2.1. Endpoint / Variables

The following subset of parameters excluded from the DB phase laboratory data delivery will be summarized:

- Autoantibody levels
    - beta-2-glycoprotein IgA,
    - beta-2-glycoprotein IgG,
    - beta-2-glycoprotein IgM,
    - Anti-RNP,
    - Anti-Ro (SS-A),
    - Anti-Ro (SS-B),
    - Anti-Ribosomal P

10.2.2. Population of Interest

These laboratory parameters were excluded from the DB phase lab data delivery, these will be reported for the DB phase using the mITT population.

10.2.3. Statistical Analyses / Methods

Details of the planned displays are provided in Appendix 9: List of Data Displays and will be based on GSK Data Standards and statistical principles.

Unless otherwise specified, endpoints / variables defined in Section 10.2.1 will be summarized using descriptive statistics and listed.

11. REFERENCES

GlaxoSmithKline Document Number 2018N366586_00. Double-blind Phase CSR, Effective Date February 1, 2019

12. APPENDICES

12.1. Appendix 1: Schedule of Activities

12.1.1. Protocol Defined Schedule of Events

Table 2 Study calendar 6-month open-label extension

Study Day	Day 0 visit (Week 52/Exit Visit On-Treatment [a])	Day 28 visit ± 7 days	Day 56 visit ± 7 days	Day 84 visit ± 7 days	Day 112 visit ± 7 days	Day 140 visit ± 7 days	Day 168 visit ± 7 days	Day 196/EXIT (4 weeks post [a])	8-week Follow-up (8 wks post last dose) ± 7 Days[f]
Study Week	Day 0	Wk 4	Wk 8	Wk 12	Wk 16	Wk 20	Wk 24	Wk 28	
IV administration of study agent[h]	X	X	X	X	X	X	X		
Clinical Assessments:									
Weight[d]	See Day 364/Exit visit in Week 52 Treatment Period	X	X	X	X	X	X		
Symptom-driven Physical Exam									
C-SSRS Since Last Visit									
Assess/Record Adverse Events		X	X	X	X	X	X	X	X
Record all current medications		X	X	X	X	X	X	X	X
SLE Disease Activity Scales i							X	X[e]	
SLICC/ACR Damage Index								X	
FACIT-Fatigue Scale[g]								X	
Laboratory Assessments:									
Labs: Hematology & Modified Chem 20 (non-fasting)	See Day 364/Exit visit in Week 52 Treatment Period						X	X[e]	X
PT/PTT							X		
Urinalysis							X	X[e]	X
Spot urine (protein to creatinine ratio)							X	X[e]	
Urine Pregnancy Test[c]		X	X	X	X	X	X	X	X
Pharmacokinetic Sampling									
Immunogenicity[f]							X	X[e]	X
B cells							X	X[e]	
Complement (C3/C4) and anti-dsDNA							X	X[e]	
Extractable nuclear antigens (ENAs)									
Anti-phospholipid antibodies (aCL, lupus anticoagulant, ±beta-2-glycoprotein-1)									
Serum Immunoglobulin IgG, IgA, IgM				X			X	X[e]	

a. Day 0 of the 6-month open-label extension is the Day 364/Week 52/Exit study visit from the On-Treatment phase of the study (see Table 6-1 in the protocol) and represents the first belimumab administration for those subjects continuing in the 6-month open-label extension.
b. All subjects, including subjects who have discontinued study agent prior to Day 196, will return for an Exit visit approximately 4 weeks after their last dose of study agent as well as a follow-up visit approximately 8 weeks after the last dose of study agent.
c. Urine pregnancy test, as required, with results available prior to dosing. Women of child-bearing potential must be reminded of the requirement to report any pregnancy that occurs through 16 weeks following the last dose of study agent (see Section 7.6 in the protocol).
d. The weight at the current visit will be used for calculating the dose. At sites where it is impractical to use the subject's current body weight to calculate the dose, it is permissible to use the subject's body weight from the previous visit. However, the subject's body weight will be measured prior to dosing and if the previous visit weight and current visit weight vary by more than 10%, then the weight measured at the current visit must be used.
e. Assessment to be performed only if assessment was not done at Day 168/Week 24.
f. The 8-week follow up visit is not required for subjects entering the separate continuation protocol.
g. Must be completed by the subject prior to any study-related discussion with the investigator or study coordinator. The FACIT-Fatigue Scale will only be completed by subjects for whom a survey exists in the subject's language.
h. Subjects should remain under clinical supervision for 3 hours after completion of the first 2 infusions. Should symptoms of acute hypersensitivity occur, an extended period of monitoring may be appropriate, based on clinical judgment.
i. SLE Disease Activity Scales: SELENA SLEDAI, SLE Flare Index, BILAG, and PGA. Refer to Section 6.8.1 in the protocol for guidelines for scoring proteinuria for SELENA SLEDAI evaluation

12.2. Appendix 2: Assessment Windows

12.2.1. Definitions of Visits for Analyses

All visits including the exit visit used the CRF visit and were not slotted based on OL study day. See Section 9.4.3 of the double-blind phase RAP for details of the windows for the double-blind phase visits. See Section 12.5.1 for the OL study day definition. For completeness, the table also includes the open-label follow up visit and the baseline visits.

Analysis Visit	Analysis Visit Number	Target OL Study Day[a]
Open-label extension visits:		
DB baseline	15	DB day 1 (needed for DB phase reports)
OL Baseline	155	DB day 1 (DB belimumab) or OL day 1 (DB placebo) the pre-belimumab baseline
Week 52/Exit DB	160	1 (OL Phase baseline, analysis visit name selected to be suitable for DB phase too)
OL Week 4	180	29
OL Week 8	190	57
OL Week 12	200	85
OL Week 16	210	113
OL Week 20	220	141
OL Week 24	230	169
OL Week 28/Exit	240	197 (The Exit visit is also mapped to this analysis visit and does not have a target day)
OL Week 24/Week 28/Exit	245	169 (OL Week 24 or OL Week 28/Exit if OL Week 24 missing)
OL Follow-up[b]	250	225 (or 8 weeks post last dose i.e. target day not relevant for early withdrawal)

a. OL Study Day with OL Treatment Start Date as OL Day 1.
b. The open-label extension follow-up occurs 4-weeks after early withdrawal or 8 weeks after last dose (e.g. Week 32) for subjects who complete the open-label extension phase.

Some subjects were not administered their first OL treatment until OL Week 4 or OL Week 8 so it is possible that the OL Day 1 visit and the OL Week 4 visit will have a negative OL study day.

12.3. Appendix 3: Study Phases, and Treatment Emergent Open-label Adverse Events

12.3.1. Study Phases

Study Phase	Definition
Pre-OL Treatment	Date < OL Study Treatment Start Date
On-OL Treatment	OL Study Treatment Start Date <= Date

Note that Adverse Events and Concomitant Medications on the first Open-label treatment start date will be included for both the primary (Double-blind) and final (Open-label) reporting.

12.3.1.1. Study Phases for OL Concomitant Medication

Study Phase	Definition
Prior DB Phase	If medication end date is not missing and is before Double-blind Treatment Start Date
DB Concomitant Prior OL phase	If medication is not Prior DB Phase and medication end date is not missing and is before Open-label Treatment Start Date
OL Concomitant	Any medication that is not Prior DB phase nor DB Concomitant Prior OL phase

NOTES:

- Please refer to Appendix 6: Reporting Standards for Missing Data for handling of missing and partial dates for concomitant medication. Use the rules in this table if concomitant medication date is completely missing.

In the OL Concomitant medication analysis dataset all Prednisone medication data is required in order to calculate the OL (pre-belimumab) baseline 7-day average prednisone dose.

12.3.2. Treatment Emergent Open-label Flag for Adverse Events

Flag	Definition
Treatment Emergent Open-label	• An OL treatment-emergent AE is an adverse event that starts on or after the first open-label treatment dose. • AEs with missing start dates will be assumed to be treatment-emergent OL. • Open-label Treatment Start Date ≤ AE Start Date.

NOTES:

- Time of open-label treatment dosing and start time of AEs should be considered, if collected.

12.4. Appendix 4: Data Display Standards & Handling Conventions

12.4.1. Reporting Process

Software	
• The currently supported versions of SAS software 9.4 will be used.	
Reporting Area	
HARP Server	: US1SALX00259
HARP Compound	: GSK1550188 / bel115471
Analysis Datasets	
• Analysis datasets will be created according to CDISC standards (SDTM IG Version 3.2 & ADaM IG Version 1.0]. If the Study Data Standardization Plan (SDSP) exists for a study, ensure the CDISC versions are consistent.	
Generation of RTF Files	
• RTF files will be generated.	

12.4.2. Reporting Standards

General

• The current GSK Integrated Data Standards Library (IDSL) will be applied for reporting, unless otherwise stated (IDSL Standards Location: https://spope.gsk.com/sites/IDSLLibrary/SitePages/Home.aspx):
  • 4.03 to 4.23: General Principles
  • 5.01 to 5.08: Principles Related to Data Listings
  • 6.01 to 6.11: Principles Related to Summary Tables
  • 7.01 to 7.13: Principles Related to Graphics
• Do not include subject level listings in the main body of the GSK Clinical Study Report. All subject level listings should be located in the modular appendices as ICH or non-ICH listings

Formats
• GSK IDSL Statistical Principles (5.03 & 6.06.3) for decimal places (DP's) will be adopted for reporting of data based on the raw data collected, unless otherwise stated. • Numeric data will be reported at the precision collected on the eCRF. • The reported precision from non eCRF sources will follow the IDSL statistical principles but may be adjusted to a clinically interpretable number of DP's.
Planned and Actual Time

• Reporting for tables, figures and formal statistical analyses:
  o Planned time relative to dosing will be used in figures, summaries, statistical analyses and calculation of any derived parameters, unless otherwise stated.
  o The impact of any major deviation from the planned assessment times and/or scheduled visit days on the analyses and interpretation of the results will be assessed as appropriate.
• Reporting for Data Listings:
  o Planned and actual time relative to study drug dosing will be shown in listings (Refer to IDSL Statistical Principle 5.05.1).
  o Unscheduled or unplanned readings will be presented within the subject's listings.

Unscheduled Visits
• Unscheduled visits will not be included in summary tables and/or figures. • All unscheduled visits will be included in listings.

Descriptive Summary Statistics	
Continuous Data	Refer to IDSL Statistical Principle 6.06.1
Categorical Data	N, n, frequency, %
Graphical Displays	
• Refer to IDSL Statistical Principals 7.01 to 7.13.	

12.5. Appendix 5: Derived and Transformed Data

12.5.1. General

Multiple Measurements at One Analysis Time Point
• If there are multiple visits within a visit window, the visit closest to the target date will be used. If there are two visits equidistant from the target date, the first will be used. • Participants having both High and Low values for Normal Ranges at any post-baseline visit for safety parameters will be counted in both the High and Low categories of "Any visit post-baseline" row of related summary tables. This will also be applicable to relevant Potential Clinical Importance summary tables.
OL Study Day
Open-label study Day is the number of days from the open-label treatment start date to a study date of interest (e.g., adverse event start date) and is calculated as follows: • Calculated as the number of days from First OL Dose Date: o Ref Date = Missing → Study Day = Missing o Ref Date < First OL Dose Date → Study Day = Ref Date – First OL Dose Date o Ref Data ≥ First OL Dose Date → Study Day = Ref Date – (First OL Dose Date) + 1 Note: Study Day cannot be zero. If either date is missing, then Study Day is missing.
Windows for Assessment of Post-injection/infusion sensitivity reactions (PISR) and hypersensitivity reactions (HSR)
• See Section 9.4.35 of the DB phase RAP.

12.5.2. Study Population

Extent of Exposure
• Number of days of exposure to OL study drug will be calculated based on the formula: **Duration of OL Exposure in Days = Last infusion date in open-label – (First infusion in open-label) + 28** • Participants who did not report a OL treatment start date will be categorized as having zero days of OL exposure. • First and last OL infusion dates will be used, regardless of any missed doses.

12.5.3. Efficacy

SLE Flare Index (SFI) Scoring
• SFI is used to report the first mild /moderate or severe flare occurrence since the last assessment. In the OL phase there is a LOG question and an assessment at OL Week 24 or OL Week 28 (if Week 24 not done). • The change in SELENA SLEDAI criteria will be assessed via the SFI form. • The mild/moderate or severe classification will be re-derived using the subcategory scores. • Flares originally marked severe will be downgraded to "Not Severe" if the only reason marked is a change in SELENA SLEDAI score to >12. o In this case, if any of the mild/moderate reasons are checked, then the flare will be considered Mild/Moderate; otherwise the flare will not be counted.

Time to First SFI Flare, Censoring and Disposition Rules

- The rules described in this section apply to SLE Flares and SLE Severe Flares.

Up to Open-label 28 Weeks

Time to first (severe) SLE flare is defined as the number of days from first open-label exposure until the subject meets an event (event date – first open-label exposure date + 1). The disposition of subjects is defined as follows in Table 3. Only flares post first open-label treatment are included in the analysis, therefore flares (not subjects) occurring on the open-label treatment start date should be removed from the analysis set prior to determining the first open-label flare. For the Open-label phase, time to first flare over 28 weeks is calculated as:

Time to first OL flare (days) = Date of first OL flare – Open-label treatment start date +1.

Table 3 Subject Disposition Rules for SLE Flare (Open-label Phase)

Subject disposition	Event met	Event date
Subject has an OL flare		
Subject has an OL flare	Yes	Date of first OL flare
Subject does not have an OL flare		
Subject withdraws	No	Censored at last OL flare assessment date
Subject dies	No	Censored at date of death
Subject completes	No	Censored at last OL flare assessment date

Baseline Prednisone Dose

The OL reports will use the OL baseline from before the first dose of belimumab. For subjects on Belimumab in the DB phase the baseline average daily prednisone dose is the sum of all prednisone doses over 7 consecutive days up to, but not including DB Day 1, divided by 7. For subjects on Placebo in the DB phase the baseline average daily prednisone dose is the sum of all prednisone doses over 7 consecutive days up to, but not including OL Day 1, divided by 7.

Prednisone Change from OL Baseline (Observed)

While on treatment, the average daily prednisone dose at the visit is the sum of all prednisone doses over 7 consecutive days up to and including the day of interest, divided by 7, unless otherwise specified. Days on which a subject does not have prednisone use recorded will be considered as 0 mg for the day in the calculation of average daily prednisone dose.

The average daily prednisone equivalent dose will be calculated for OL Week 28/Exit. For subjects who withdraw from the study without an Exit scheduled visit, the average dose will be calculated at the date of withdrawal. For a subject who withdrew at OL baseline then the average dose will be the OL baseline value (i.e. zero change from OL baseline).

The time period considered is the 7 days up to and including the OL Week 28 / Exit visit. No imputation will be done for missing data.

Any Increase in Prednisone at Week 28/Exit Compared to OL Baseline (Observed)

Subjects are considered to have 'any increase' from OL baseline (pre-belimumab) when the change from baseline in average daily prednisone dose is >0.

The time period considered is the 7 days up to and including the OL Week 28 / Exit visit.

Prednisone Reduced to <=7.5 mg/day at Open-Label Phase Week 28/Exit (Observed)
A responder is defined as a subject who decreased their daily prednisone dose to <=7.5 mg/day from an OL baseline dose >7.5 mg/day. This endpoint is only defined for subjects who used prednisone >7.5 mg/day at OL baseline. The time period considered is the 7 days up to and including up to the OL Week 28 / Exit visit. If a subject misses their Week 28 or Exit visit, then the subject will be missing for that visit (Observed).

Prednisone Increased to > 7.5 mg/day at Open-Label Phase Week 28/Exit (Observed)
This is a summary of subjects who increased their daily prednisone dose to >7.5 mg/day from an OL baseline dose <=7.5 mg/day. This endpoint is only defined for subjects who used prednisone <=7.5 mg/day at OL baseline. The time period considered is the 7 days up to and including up to the OL Week 28 / Exit visit.

Time to First Renal Flare, Censoring and Disposition Rules
Time to first renal flare is defined as the number of days from first open-label exposure until the subject meets an event (event date – first open-label exposure date + 1). The disposition of subjects is defined as follows in Table 4. Open-label Renal flares are defined as reproducible results comparative to Open-label Phase Baseline values, so can only occur post-open-label Baseline. Time to first renal flare over 28 weeks is calculated as: *Time to first OL flare (days) = Date of first OL flare – open-label treatment start date +1* NB: Patients who withdraw or are lost to follow up without an OL renal flare assessment will be censored at zero days. Patients who withdraw or are lost to follow up with only one renal flare assessment (which therefore cannot be reproduced) will be censored at their last open-label flare assessment date. It is possible for a renal flare to occur at the Week 28/Exit visit as it could be reproduced at an unscheduled visit or the 8 week follow up visit.

Table 4 Subject Disposition Rules for Renal Flares (Open-label Phase)

Subject disposition	Event met	Event date
Subject has an OL flare		
Subject has an OL renal flare	Yes	Date of first renal flare which has been reproduced by a later assessment
Subject does not have an OL flare		
Subject withdraws	No	Censored at last OL flare assessment date
Subject dies	No	Censored at last OL flare assessment date
Subject completes	No	Censored at last OL flare assessment date

12.6. Appendix 6: Reporting Standards for Missing Data

12.6.1. Premature Withdrawals

Element	Reporting Detail
Premature Withdrawals General	• Subjects who drop out and do not complete the OL study will return for an OL Exit visit 4 weeks after their final dose of study agent. • During the 6-month open-label extension, the assessments for the Week 28 visit will be performed for the OL Exit visit. • All post OL treatment data from participants who were withdrawn from the OL study will be listed and all available data will be included in summary tables and figures, unless otherwise specified. • The SLE Disease Activity Scales are completed at OL Week 24, and OL Week 28/Exit should be completed only if Week 24 is missing.

12.6.2. Handling of Missing Data

Element	Reporting Detail
Last Observation Carried Forward (LOCF)	• The LOCF principle is applied whereby missing values will be replaced with the last previous non-missing value. • If a subject withdraws the LOCF value will be handled by using the result from the last visit post OL day 1 and prior to the date of withdrawal. • If any observed dates exist, the date for the LOCF record will be determined from the observed dates. If there are multiple observed dates, the latest of these dates (closest to the scheduled visit) will be taken as the derived date. If the record is completely missing then the date will be missing, otherwise missing dates will be imputed as the target open-label study day from the schedule of events.
Observed	• Observed data are the data collected or observed for the subject with no imputation for missing data. The observed case analysis includes all data collected including data post IP withdrawal.
Worst Observation Carried Forward (WOCF)	• The worst observation carried forward (WOCF) principle will be applied to the SLICC/ACR Damage Index data at an item level. As the SLICC/ACR Damage Index is meant to be a monotonic, non-decreasing value (except for renal domain), items that decrease over the course of the study will be replaced with the highest previous item score achieved among previous visits (including study baseline). Within the renal domain, damage scored on the first two items (estimated glomerular filtration rate < 50%, proteinuria > 3.5 gm/24hours) should only be carried forward to OL Week 24/28 if the third item (end stage renal disease) has not been scored at OL Week 24/28. The reason for this is that scores on the first two items can decrease but only if the third item is then scored.

12.6.2.1. Handling of Missing and Partial Dates

Element	Reporting Detail
General	• Partial dates will be displayed as captured in subject listing displays.
Adverse Events	• The eCRF does not allow for the possibility of partial AE dates. • Completely missing start or end dates will remain missing, with no imputation applied.

Element	Reporting Detail
	Consequently, time to onset and duration of such events will be missing. AEs with missing start dates will be considered as OL treatment emergent.
Concomitant Medications/	Partial dates for any concomitant medications recorded in the CRF will be imputed using the following convention: Medication Start Date (CMSTDT) • If CMENDT is missing OR CMENDT >= Belimumab TRTSDT (whether CMENDT is a complete or partial date) then CMSDT is imputed as the Belimumab TRTSDT. • If CMENDT is < Belimumab TRTSDT (whether CMENDT is a complete or partial date) then CMSTDT is imputed with JAN for missing month and 01 for missing day, whatever is applicable. Medication End Date (CMENDT) • If month and year are present, then set to the earlier of (last contact date or last day of that month of the CMENDT). • If only year present, then set to the earlier of (31DEC of the year or last contact date). Missing dates for any concomitant medications recorded in the CRF will be imputed using the following convention: Medication Start Date (CMSTDT) CMSTDT is imputed as Belimumab TRTSDT unless: • CMENDT is < Belimumab TRTSDT (whether CMENDT is a complete (DD/MM/YY) or partial date (some combination of CMENDT day, month or year imputed) OR • The month and/or month and year of the partial CMENDT date are before the month and/or year of Belimumab TRTSDT OR • "Taken prior to study?" is checked. Medication End Date (CMENDT) • End dates for concomitant medications will not be imputed, and the medication will be considered ongoing. Note that medications with partial or missing start and/or stop dates will be assumed to be OL concomitant unless there is evidence through comparison of partial dates to suggest otherwise, for example if the day is missing, then the month and year will be compared to the month and year of the first dose date of study treatment and if the month and year are the same or later, then the medication will be considered concomitant. See Section 9.3.1 and Section 9.3.2 of the DB phase RAP.

12.7. Appendix 7: Values of Potential Clinical Importance

12.7.1. Laboratory Values

The Adverse Event and Laboratory Toxicity Grading Tables for this study can be found in the Double-Blind Phase RAP in Section 18.11 (Appendix 11).

12.8. Appendix 8: Abbreviations & Trade Marks

12.8.1. Abbreviations

Abbreviation	Description
ACR	American College of Rheumatology
ADaM	Analysis Data Model
AE	Adverse Event
BILAG	British Isles Lupus Assessment Group of SLE Clinics
BILAG No 1A/2B	No new BILAG A organ domain score or 2 new BILAG B organ domain scores compared with baseline
BMI	Body mass index
CDISC	Clinical Data Interchange Standards Consortium
CRF	Case Report Form
CSR	Clinical Study Report
C-SSRS	Columbia Suicide Severity Rating Scale
DB	Double-blind
DBF	Database Freeze
DBR	Database Release
DP	Decimal Places
eCRF	Electronic Case Record Form
EMA	European Medicines Agency
FACIT	Functional Assessment of Chronic Illness Therapy
FDA	Food and Drug Administration
FDAAA	Food and Drug Administration Clinical Results Disclosure Requirements
GSK	GlaxoSmithKline
ICH	International Conference on Harmonization
IDSL	Integrated Data Standards Library
ITT	Intent-To-Treat
IV	Intravenous
IVRS	Interactive Voice Recognition System
LOCF	Last Observation Carried Forward
MedDRA	Medical Dictionary for Regulatory Activities
mITT OL	Modified Intention-To-Treat Open-label
OL	Open-label
OLE	Open-label extension
PDMP	Protocol Deviation Management Plan
PGA	Physician's Global Assessment
PK	Pharmacokinetic
PSAP	Program Safety Analysis Plan
PSRQ	Possible Suicidality Related Questionnaire
PT	Preferred Term
RAP	Reporting & Analysis Plan
SAC	Statistical Analysis Complete
SAE	Serious Adverse Event
SDL	Source Data Lock

Abbreviation	Description
SDSP	Study Data Standardization Plan
SDTM	Study Data Tabulation Model
SELENA	Safety of Estrogen in Lupus National Assessment
SFI	SLE Flare Index
SLE	Systemic Lupus Erythematosus
SLEDAI	Systemic Lupus Erythematosus Disease Activity Index
SLICC	Systemic Lupus International Collaborating Clinics
SOC	System Organ Class
SOP	Standard Operation Procedure
SRI	SLE Responder Index
SRI-S2K	SLE Responder Index (SRI) using the SELENA SLEDAI modified with S2K scoring for proteinuria
TFL	Tables, Figures & Listings
WBC	White Blood Cell
WOCF	Worst Observation Carried Forward

See also the Abbreviation section of the DB phase RAP.

12.8.2. Trademarks

Trademarks of the GlaxoSmithKline Group of Companies
NONE

Trademarks not owned by the GlaxoSmithKline Group of Companies
SAS

12.9. Appendix 9: List of Data Displays

12.9.1. Data Display Numbering

The following numbering will be applied for Open-label extension RAP generated displays:

Section	Tables	Figures
Study Population	11.1 to 11.n	11.1 to 11.n
Efficacy	12.1 to 12.n	12.1 to 12.n
Safety	13.1 to 13.n	13.1 to 13.n
Health Outcomes	14.1 to 14.n	14.1 to 14.n
Pharmacokinetic	15.1 to 15.n	15.1 to 15.n
Biomarker	16.1 to 16.n	16.1 to 16.n

12.9.2. Deliverables

Delivery [Priority] [1]	Description
SAC [X]	Final Statistical Analysis Complete

NOTES:

- Indicates priority (i.e. order) in which displays will be generated for the reporting effort

12.9.3. Study Population Tables

Study Population Tables					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
Subject Disposition					
11.1.	mITT OL	ES1/ ES1A / ES8 / ES8A	Subject Completion Status by End of Open-Label Phase Week 28	ICH E3, FDAAA, EudraCT	SAC [1]
11.2.	Intent-to-Treat Open-label	ES1/ ES1A / ES8 / ES8A	Subject Completion Status by End of Open-Label Phase Week 28	ICH E3, FDAAA, EudraCT	SAC [1]
Protocol Deviation					
11.3.	Intent-to-Treat Open-label	DV1	Subjects with Important Protocol Deviations (Open-Label Phase)	ICH E3	SAC [1]
Population Analyzed					
11.4.	Intent-to-Treat Open-label	SP1 / SP1A	Summary of Study Populations (Open-Label Phase)	IDSL	SAC [1]
Demographic and Baseline Characteristics					
11.5.	mITT OL	DM1 / DM3	Demographic and Baseline Characteristics (Open-Label Phase)	ICH E3, FDAAA, EudraCT	SAC [1]
11.6.	Intent-to-Treat Open-label	DM1 / DM3	Demographic and Baseline Characteristics (Open-Label Phase)	ICH E3, FDAAA, EudraCT	SAC [1]
11.7.	mITT OL		Baseline Disease Activity (Open-Label Phase)	Duration is defined as (Belimumab treatment start date - SLE diagnosis date + 1)/365.25. Baseline is pre-Belimumab.	SAC [1]

Study Population Tables					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
11.8.	Intent-to-Treat Open-label		Baseline Disease Activity (Open-Label Phase)	Duration is defined as (Belimumab treatment start date - SLE diagnosis date + 1)/365.25. Baseline is pre-Belimumab.	SAC [1]
11.9.	mITT OL		Summary of Race and Racial Combination (Open-Label Phase)		SAC [1]
11.10.	mITT OL		SELENA SLEDAI Organ and Item Involvement at Baseline (Open-Label Phase)	Baseline is Pre-belimumab baseline.	SAC [1]
11.11.	Intent-to-Treat Open-label		SELENA SLEDAI Organ and Item Involvement at Baseline (Open-Label Phase)	Baseline is Pre-belimumab baseline.	SAC [1]
11.12.	mITT OL		Immunoglobulin Levels at Baseline (Open-Label Phase)	Baseline is Pre-belimumab baseline.	SAC [1]
11.13.	Intent-to-Treat Open-label		Immunoglobulin Levels at Baseline (Open-Label Phase)	Baseline is Pre-belimumab baseline.	SAC [1]
11.14.	mITT OL		Complement Levels and Other Biomarkers at Baseline (Open-Label Phase)	Baseline is Pre-belimumab baseline.	SAC [1]
11.15.	Intent-to-Treat Open-label		Complement Levels and Other Biomarkers at Baseline (Open-Label Phase)	Baseline is Pre-belimumab baseline.	SAC [1]
11.16.	Intent-to-Treat Open-label		Allowable SLE Medication Usage at Baseline (Open-Label Phase)	Baseline is Pre-belimumab baseline.	SAC [1]
11.17.	Intent-to-Treat Open-label		Steroid, Anti-malarial and Immunosuppressant Use at Baseline (Open-Label Phase)	Baseline is Pre-belimumab baseline.	SAC [1]
Prior and Concomitant Medications					
11.18.	Intent-to-Treat Open-label	CM1	Concomitant Medications by ATC Level 1 and ATC Level 4 Term (Open-Label Phase)	ICH E3	SAC [1]
11.19.	Intent-to-Treat Open-label		Concomitant Medications by ATC Level 4 and Preferred Term (Open-Label Phase)		SAC [1]

Study Population Tables					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
Exposure and Treatment Compliance					
11.20.	Modified Intent-to-Treat Open-label	EX1 / EX5	Study Drug Exposure during Open-Label Phase Week 24	ICH E3 For ClinPharm, a listing often substitutes for a table.	SAC [1]
11.21.	Intent-to-Treat Open-label	EX1 / EX5	Study Drug Exposure during Open-Label Phase Week 24	ICH E3	SAC [1]
Biomarker parameters missed from DB phase delivery					
11.22.	mITT		Autoantibody Levels at Baseline Continued (Double-blind Phase)	Report • Beta-2-glycoprotein IgA (U/mL), • Beta-2-glycoprotein IgG (U/mL) • Beta-2-glycoprotein IgM (U/mL) as not delivered at time of DB report. (format as the DB phase report)	SAC [1]

12.9.4. Efficacy Tables

Efficacy: Tables					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
SRI-S2K Response					
12.1.	mITT OL		SRI-S2K Response at Open-Label Phase Week 24 /Week 28 (Observed)		SAC [1]
12.2.	mITT OL		SRI-S2K Response at Open-Label Phase Week 24 /Week 28 and the 3 Components (Observed)		SAC [1]
12.3.	mITT OL		Disposition of SRI-S2K Response at Open-Label Phase Week 24 /Week 28 (Observed)		SAC [1]
12.4.	Completer OL		SRI-S2K Response at Open-Label Phase Week 24 /Week 28 (Completer Sensitivity Analysis) (Observed)		SAC [1]
12.5.	mITT OL		SRI-SS Response at Open-Label Phase Week 24 /Week 28 (SELENA SLEDAI) [1] (Observed)		SAC [1]
12.6.	mITT OL		SRI5-S2K Response at Open-Label Phase Week 24 /Week 28 (Observed)		SAC [1]
12.7.	mITT OL		SRI6-S2K Response at Open-Label Phase Week 24 /Week 28 (Observed)		SAC [1]
12.8.	mITT OL		SRI7-S2K Response at Open-Label Phase Week 24 /Week 28 (Observed)		SAC [1]
12.9.	mITT OL		SRI8-S2K Response at Open-Label Phase Week 24 /Week 28 (Observed)		SAC [1]
12.10.	mITT OL		EMA Modified SRI-S2K Response [1] at Open-Label Week 24 /Week 28 (Observed)		SAC [1]
12.11.	mITT OL		SRI-S2K Response at Open-Label Phase Week 24 /Week 28 (Observed) by Region		SAC [1]

Efficacy: Tables					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
SFI Flare					
12.12.	mITT OL		Time to First Severe SFI Flare over Open-Label Phase [1] (DO=Censor)		SAC [1]
12.13.	mITT OL		Time to First SFI Flare over Open-Label Phase [1] (DO=Censor)		SAC [1]
SLICC/ACR Damage Index					
12.14.	mITT OL		SLICC/ACR Damage Index Change from OL Baseline at Open-Label Phase Week 28 (Observed Case, Corrected)	Pre-Belimumab Baseline	SAC [1]
12.15.	mITT OL		SLICC/ACR Damage Index Change from OL Baseline at Open-Label Phase Week 28 by SLICC/ACR Damage Index Subgroup at OL Baseline (No Damage vs. Damage Index >=1) (Observed Case, Corrected)	Pre-Belimumab Baseline	SAC [1]
12.16.	mITT OL		SLICC/ACR Damage Index Worsening (Change >0) Compared with OL Baseline at Open-Label Phase Week 28 (WOCF)	Pre-Belimumab Baseline	SAC [1]
12.17.	mITT OL		SLICC/ACR Damage Index Worsening (Change >0) Compared with OL Baseline at Open-Label Phase Week 28 (WOCF) by SLICC/ACR Damage Index Subgroup at OL Baseline (No Damage vs. Damage Index >=1)	Pre-Belimumab Baseline	SAC [1]
SS-S2K					
12.18.	mITT OL		SS-S2K >= 4-point Reduction from OL Baseline at Open-Label Phase Week 24/Week 28 (Observed)	Pre-Belimumab Baseline	SAC [1]
12.19.	mITT OL		SELENA SLEDAI-S2K Organ System Improvement by Organ System at Open-Label Phase Week 24/Week 28 (Observed) among Subjects with Organ System Involvement at OL Baseline [1]	Pre-Belimumab Baseline	SAC [1]
12.20.	mITT OL		SS-S2K Percent Change from OL Baseline at Open-Label Phase Week 24 and Week 28 (Observed)	Pre-Belimumab Baseline	SAC [1]
12.21.	mITT OL		SS-S2K Change from OL Baseline at Open-Label Phase Week 24 and Week 28 (Observed)	Pre-Belimumab Baseline	SAC [1]

Efficacy: Tables					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
SELENA SLEDAI					
12.22.	mITT OL		SELENA SLEDAI Percent Change from OL Baseline at Open-Label Phase Week 24 and Week 28 (Observed)	Pre-Belimumab Baseline	SAC [1]
12.23.	mITT OL		SELENA SLEDAI Change from OL Baseline at Open-Label Phase Week 24 and Week 28 (Observed)	Pre-Belimumab Baseline	SAC [1]
BILAG					
12.24.	mITT OL		BILAG no New 1A/2B Organ Domain scores compared with OL Baseline at Open-Label Phase Week 24/Week 28 (Observed)	Pre-Belimumab Baseline	SAC [1]
12.25.	mITT OL		BILAG no 1A/2B Organ Domain scores at Open-Label Phase Week 24/Week 28 (Observed)		SAC [1]
12.26.	mITT OL		BILAG Improvement by Organ Domain at Open-Label Phase Week 24/Week 28 (Observed) among Subjects with an A or B Domain Score at OL Baseline [1]	Pre-Belimumab Baseline	SAC [1]
12.27.	mITT OL		BILAG Worsening by Organ Domain at Open-Label Phase Week 24/Week 28 (Observed) among Subjects with no A Domain Score at OL Baseline [1]	Pre-Belimumab Baseline	SAC [1]
PGA					
12.28.	mITT OL		PGA No Worsening (Increase of <0.3 points from OL Baseline) at Open-Label Phase Week 24/Week 28 (Observed)	Pre-Belimumab Baseline	SAC [1]
12.29.	mITT OL		PGA Percent Change from OL Baseline at Open-Label Phase Week 24 and Week 28 (Observed)	Pre-Belimumab Baseline	SAC [1]
12.30.	mITT OL		PGA Change from OL Baseline at Open-Label Phase Week 24 and Week 28 (Observed)	Pre-Belimumab Baseline	SAC [1]
12.31.	mITT OL		PGA Improvement (Increase of >=0.3 points from OL Baseline) at Open-Label Phase Week 24 or Week 28 (Observed)	Subjects with baseline PGA score < 0.3 will be treated as having no improvement. Pre-Belimumab Baseline.	SAC [1]

Efficacy: Tables					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
Renal Flare					
12.32.	mITT OL		Time to First Renal Flare over Open-Label Phase 28 weeks [1] (DO=Censor)		SAC [1]
12.33.	mITT OL		Time to First Renal Flare over Open-Label Phase 28 weeks [1] (DO=Censor) among Subjects with OL Phase Baseline Proteinuria > 0.5 g/24hr		SAC [1]
Proteinuria					
12.34.	mITT OL		Proteinuria Percent Change from OL Baseline at Open-Label Phase Week 24 and Week 28 (Observed) among Subjects with OL Baseline Proteinuria > 0.5 g/24hr	Pre-Belimumab Baseline	SAC [1]
12.35.	mITT OL		Proteinuria Change from OL Baseline at Open-Label Phase Week 24 and Week 28 (Observed)	Pre-Belimumab Baseline	SAC [1]
12.36.	mITT OL		Proteinuria Shifts from OL Baseline at Open-Label Phase Week 24 and Week 28 (Observed)	Pre-Belimumab Baseline	SAC [1]
12.37.	mITT OL		Percent of Subjects with Doubling of Serum Creatinine from OL Baseline at Open-Label Phase Week 24 and Week 28 (Observed) among Subjects with OL Baseline Proteinuria >0.5 g/24hr	Pre-Belimumab Baseline	SAC [1]
Prednisone					
12.38.	mITT OL		Prednisone Change from OL Baseline at Open-Label Phase Week 28/Exit (Observed)	Pre-Belimumab Baseline	SAC [1]
12.39.	mITT OL		Any Increase in Prednisone Compared to OL Baseline at Open-Label Phase Week 28/Exit (Observed)	Pre-Belimumab Baseline	SAC [1]
12.40.	mITT OL		Prednisone Reduced to <=7.5 mg/day at Open-Label Phase Week 28/Exit (Observed) among Subjects with OL Baseline Prednisone Dose >7.5 mg/day	Pre-Belimumab Baseline	SAC [1]
12.41.	mITT OL		Prednisone Increased to > 7.5 mg/day at Open-Label Phase Week 28/Exit (Observed) among Subjects with OL Baseline Prednisone Dose <=7.5 mg/day	Pre-Belimumab Baseline	SAC [1]

12.9.5. Safety Tables

Safety: Tables					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
Adverse Events (AEs)					
13.1.	Intent-to-Treat Open-label	Similar to AE13	Adverse Events Summary (Open-Label Phase)	ICH E3	SAC [1]
13.2.	Intent-to-Treat Open-label		Adverse Events by SOC (Open-Label Phase)	ICH E3 .	SAC [1]
13.3.	Intent-to-Treat Open-label	As AE1 but with total column	Adverse Events by SOC and PT (Open-Label Phase)		SAC [1]
13.4.	Intent-to-Treat Open-label		Adverse Events by PT (Open-Label Phase)		SAC [1]
13.5.	Intent-to-Treat Open-label	AE15	Common [1] Non-Serious Adverse Events by SOC and PT (Number of Subjects and Occurrences) (Open-Label Phase)	FDAAA, EudraCT ‘	SAC [1]
13.6.	Intent-to-Treat Open-label		Adverse Events by SOC and Severity (Open-Label Phase)		SAC [1]
13.7.	Intent-to-Treat Open-label		Adverse Events by SOC and PT and Severity (Open-Label Phase)		SAC [1]
13.8.	Intent-to-Treat Open-label		Relationship between System Organ Class and Verbatim Text (Open-Label Phase)		SAC [1]

Safety: Tables					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
Serious and Other Significant Adverse Events					
13.9.	Intent-to-Treat Open-label		Serious Adverse Events by SOC (Open-Label Phase)	FDAAA, EudraCT	SAC [1]
13.10.	Intent-to-Treat Open-label	AE1	Serious Adverse Events by SOC and PT (Open-Label Phase)		SAC [1]
13.11.	Intent-to-Treat Open-label		Serious Adverse Events by PT (Open-Label Phase)		SAC [1]
13.12.	Intent-to-Treat Open-label	AE16	Serious Adverse Events by SOC and PT (Number of Subjects and Occurrences) (Open-Label Phase)	As DB Phase table 3.21	SAC [1]
13.13.	Intent-to-Treat Open-label		Study Agent Related Adverse Events by SOC (Open-Label Phase)		SAC [1]
13.14.	Intent-to-Treat Open-label	AE1	Study Agent Related Adverse Events by SOC and PT (Open-Label Phase)		SAC [1]
13.15.	Intent-to-Treat Open-label		Study Agent Related Adverse Events by PT (Open-Label Phase)		SAC [1]
13.16.	Intent-to-Treat Open-label		Adverse Events Resulting in Study Agent Discontinuation by SOC (Open-Label Phase)		SAC [1]
13.17.	Intent-to-Treat Open-label	AE1	Adverse Events Resulting in Study Agent Discontinuation by SOC and PT (Open-Label Phase)		SAC [1]
13.18.	Intent-to-Treat Open-label	AE3	Adverse Events Resulting in Study Agent Discontinuation by PT (Open-Label Phase)		SAC [1]
13.19.	Intent-to-Treat Open-label	AE1	Study Agent Related Serious Adverse Events by SOC and PT (Open-Label Phase)		SAC [1]

Safety: Tables					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
13.20.	Intent-to-Treat Open-label	AE1	Fatal Serious Adverse Events by SOC and PT (Open-Label Phase)		SAC [1]
13.21.	Intent-to-Treat Open-label		Adverse Events of Special Interest by Category (Open-Label Phase)		SAC [1]
13.22.	Intent-to-Treat Open-label		Malignant Neoplasm Adverse Events of Special Interest by Category and PT (Open-Label Phase)		SAC [1]
13.23.	Intent-to-Treat Open-label		Post-Infusion Systemic Reactions Adverse Events of Special Interest by Category and PT (Open-Label Phase)		SAC [1]
13.24.	Intent-to-Treat Open-label		Serious Post-Infusion Systemic Reactions Adverse Events of Special Interest by Category and PT (Open-Label Phase)		SAC [1]
13.25.	Intent-to-Treat Open-label		Infection Adverse Events of Special Interest by Category and PT (Open-Label Phase)		SAC [1]
13.26.	Intent-to-Treat Open-label		Infection Adverse Events of Special Interest Leading to Study Agent Discontinuation by Category and PT (Open-Label Phase)		SAC [1]
13.27.	Intent-to-Treat Open-label		Depression/Suicide/Self-injury Adverse Events of Special Interest by Category and PT (Open-Label Phase)		SAC [1]
13.28.	Intent-to-Treat Open-label		Deaths by Category and PT (Open-Label Phase)		SAC [1]
13.29.	Intent-to-Treat Open-label		Post-Infusion Systemic Reactions per Anaphylactic Reactions CMQ Broad Search by PT in First Six Infusions (Open-Label Phase)		SAC [1]
13.30.	Intent-to-Treat Open-label		Serious Post-Infusion Systemic Reactions per Anaphylactic Reactions CMQ Broad Search by PT in First Six Infusions (Open-Label Phase)		SAC [1]
13.31.	Intent-to-Treat Open-label		Serious Post-Infusion Systemic Reactions/Hypersensitivity per GSK Adjudication by PT in First Six Infusions (Open-Label Phase)		SAC [1]

Safety: Tables					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
Laboratory: Chemistry					
13.32.	Intent-to-Treat Open-label		Laboratory Results by Visit: Electrolytes (Open-Label Phase)		SAC [1]
13.33.	Intent-to-Treat Open-label		Laboratory Results by Visit: Other Chemistries (Open-Label Phase)		SAC [1]
13.34.	Intent-to-Treat Open-label		Laboratory Results by Visit: Immunoglobulins (Open-Label Phase)		SAC [1]
13.35.	Intent-to-Treat Open-label		Worst Laboratory Toxicity Grade: Electrolytes (Open-Label Phase)		SAC [1]
13.36.	Intent-to-Treat Open-label		Worst Laboratory Toxicity Grade: Other Chemistries (Open-Label Phase)		SAC [1]
13.37.	Intent-to-Treat Open-label		Worst Laboratory Toxicity Grade: Immunoglobulins (Open-Label Phase)		SAC [1]
13.38.	Intent-to-Treat Open-label		Laboratory Toxicity Grade Worsening of at Least 2 Grades from OL Baseline: Electrolytes (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
13.39.	Intent-to-Treat Open-label		Laboratory Toxicity Grade Worsening of at Least 2 Grades from OL Baseline: Other Chemistries (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
13.40.	Intent-to-Treat Open-label		Laboratory Toxicity Grade Worsening of at Least 2 Grades from OL Baseline: Immunoglobulin G (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
13.41.	Intent-to-Treat Open-label		Laboratory Reference Range Shifts from OL Baseline by Visit: Electrolytes (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]

Safety: Tables					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
13.42.	Intent-to-Treat Open-label		Laboratory Reference Range Shifts from OL Baseline by Visit: Other Chemistries (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
13.43.	Intent-to-Treat Open-label		Laboratory Reference Range Shifts from OL Baseline by Visit: Immunoglobulins (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
13.44.	Intent-to-Treat Open-label		Immunoglobulin Levels below the Lower Limit of Normal (LLN) by Visit (Open-Label Phase)		SAC [1]
13.45.	Intent-to-Treat Open-label		Immunoglobulin Levels below the Lower Limit of Normal (LLN) at each Visit among Subjects with Immunoglobulins >= LLN at OL Baseline (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
13.46.	Intent-to-Treat Open-label		Immunoglobulin Levels above the Lower Limit of Normal (LLN) by Visit (Open-Label Phase)		SAC [1]
13.47.	Intent-to-Treat Open-label		Immunoglobulin Levels above the Lower Limit of Normal (LLN) at each Visit among Subjects with Immunoglobulins < LLN at OL Baseline (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
Laboratory: Hematology					
13.48.	Intent-to-Treat Open-label		Laboratory Results by Visit: Hematology (Open-Label Phase)		SAC [1]
13.49.	Intent-to-Treat Open-label		Worst Laboratory Toxicity Grade: Hematology (Open-Label Phase)		SAC [1]
13.50.	Intent-to-Treat Open-label		Laboratory Toxicity Grade Worsening of at Least 2 Grades from OL Baseline: Hematology (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]

Safety: Tables					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
13.51.	Intent-to-Treat Open-label		Laboratory Reference Range Shifts from OL Baseline by Visit: Hematology (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
Laboratory: Urinalysis					
13.52.	Intent-to-Treat Open-label		Worst Laboratory Toxicity Grade: Urinalysis (Open-Label Phase)		SAC [1]
13.53.	Intent-to-Treat Open-label		Laboratory Toxicity Grade Worsening of at Least 2 Grades from OL Baseline: Urinalysis (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
13.54.	Intent-to-Treat Open-label		Laboratory Reference Range Shifts from OL Baseline by Visit: Urinalysis (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
Laboratory: Hepatobiliary (Liver)					
13.55.	Intent-to-Treat Open-label		Laboratory Results by Visit: Liver Function (Open-Label Phase)		SAC [1]
13.56.	Intent-to-Treat Open-label		Worst Laboratory Toxicity Grade: Liver Function (Open-Label Phase)		SAC [1]
13.57.	Intent-to-Treat Open-label		Laboratory Toxicity Grade Worsening of at Least 2 Grades from OL Baseline: Liver Function (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
13.58.	Intent-to-Treat Open-label		Laboratory Reference Range Shifts from OL Baseline by Visit: Liver Function (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]

Safety: Tables					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
Immunogenicity					
13.59.	Intent-to-Treat Open-label		Immunogenic Response by Visit (Open-Label Phase)		SAC [1]
Vital Signs					
13.60.	Intent-to-Treat Open-label	VS1	Vital Signs by Visit (Observed) (Open-Label Phase)	ICH E3	SAC [1]
13.61.	Intent-to-Treat Open-label		Vital Signs Change from OL Baseline by Visit (Observed) (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]

12.9.6. Health Outcomes Tables

Efficacy: Tables					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
FACIT- Fatigue Scale					
14.1.	mITT OL		FACIT-Fatigue Scale Score Change from OL Baseline (Observed) at Open-Label Phase Week 28	Pre-Belimumab Baseline	SAC [1]
14.2.	mITT OL		Improvement in FACIT-Fatigue Scale Score exceeding the MCID (>=4) at Open-Label Phase Week 28		SAC [1]

12.9.7. Pharmacokinetic Tables

Efficacy: Tables					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
Belimumab Concentration					
15.1	PK		Belimumab Concentrations (ug/mL) (Observed) (Double-Blind Phase)	As DB table 5.01	SAC [1]

12.9.8. Biomarker Tables

Biomarker: Tables					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
Immunoglobin Levels					
16.1.	Modified Intent-to-Treat Open-label		Immunoglobulin Levels Percent Change from OL Baseline by Visit (Observed) (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
16.2.	Modified Intent-to-Treat Open-label		Immunoglobulin Levels Change from OL Baseline by Visit (Observed) (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
16.3.	Modified Intent-to-Treat Open-label		Immunoglobulin Levels Shifts from OL Baseline by Visit (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
Autoantibody levels					
16.4.	Modified Intent-to-Treat Open-label		Autoantibody Levels Percent Change from OL Baseline by Visit (Observed) (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]

Biomarker: Tables					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
16.5.	Modified Intent-to-Treat Open-label		Autoantibody Levels Percent Change from OL Baseline by Visit among Subjects Positive at OL Baseline (Observed) (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
16.6.	Modified Intent-to-Treat Open-label		Autoantibody Levels Change from OL Baseline by Visit (Observed) (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
16.7.	Modified Intent-to-Treat Open-label		Autoantibody Levels Change from OL Baseline by Visit among Subjects Positive at OL Baseline (Observed) (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
16.8.	Modified Intent-to-Treat Open-label		Autoantibody Levels Shifts from OL Baseline by Visit (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
Complement Levels					
16.9.	Modified Intent-to-Treat Open-label		Complement Levels Percent Change from OL Baseline by Visit (Observed) (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
16.10.	Modified Intent-to-Treat Open-label		Complement Levels Percent Change from OL Baseline by Visit among Subjects with Low Complement at OL Baseline (Observed) (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
16.11.	Modified Intent-to-Treat Open-label		Complement Levels Change from OL Baseline by Visit (Observed) (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
16.12.	Modified Intent-to-Treat Open-label		Complement Levels Change from OL Baseline by Visit among Subjects with Low Complement at OL Baseline (Observed) (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]

Biomarker: Tables					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
16.13.	Modified Intent-to-Treat Open-label		Complement Levels Shifts from OL Baseline by Visit (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
16.14.	Modified Intent-to-Treat Open-label		B-cells Percent Change from OL Baseline by Visit (Observed) (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]
16.15.	Modified Intent-to-Treat Open-label		B-cells Change from OL Baseline by Visit (Open-Label Phase)	Pre-Belimumab Baseline	SAC [1]

Biomarker: Tables					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
Autoantibody levels Double-blind Phase					
16.16.	mITT		Autoantibody Levels Percent Change from Baseline by Visit continued (Observed) (Double-Blind Phase)	Report parameters not delivered in time for DB phase report: a. Beta-2 Glycoprotein 1 IgA Antibody, b. Beta-2 Glycoprotein 1 IgG Antibody, c. Beta-2 Glycoprotein 1 IgM Antibody, d. Anti-Ribosomal P Antibody (U/mL), e. Sjogrens SS-A Antibody (INDEX), f. Sjogrens SS-B Antibody (INDEX), g. Anti-RNP, NB: should use Wilcoxon rank sum. (format as the DB phase report)	SAC [1]
16.17.	mITT		Autoantibody Levels Percent Change from Baseline by Visit (Observed) among Subjects Positive at Baseline continued (Double-Blind Phase)	Parameters as above. NB: should use Wilcoxon rank sum. (format as the DB phase report)	SAC [1]
16.18.	mITT		Autoantibody Levels Change from Baseline by Visit continued (Observed) (Double-Blind Phase)	Parameters as above NB: should use Wilcoxon rank sum. (format as the DB phase report)	SAC [1]

Biomarker: Tables					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
16.19.	mITT		Autoantibody Levels Change from Baseline by Visit (Observed) among Subjects Positive at Baseline continued (Double-Blind Phase)	Parameters as above NB: should use Wilcoxon rank sum. (format as the DB phase report)	SAC [1]
16.20.	mITT		Autoantibody Levels Shifts from Baseline by Visit continued (Double-Blind Phase)	Parameters as above. (format as the DB phase report)	SAC [1]

12.9.9. ICH Listings

ICH: Listings					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
Subject Disposition					
11.01	Intent-to-Treat Open-label	ES2 / ES3	Reasons for Study Withdrawal (Open-Label Phase)	ICH E3	SAC [1]
11.02	Intent-to-Treat Open-label	TA1 / CP_RD1x	Listing of Planned and Actual Treatments (Open-Label Phase)	IDSL (some subjects received Placebo rather than Belimumab in the OL phase these should be identified here)	SAC [1]
Protocol Deviations					
11.03	Intent-to-Treat Open-label	DV2	Important Protocol Deviations (Open-Label Phase)	ICH E3	SAC [1]
Populations Analyzed					
11.04	Intent-to-Treat Open-label	SP3/SP3a	Listing of Participants Excluded from Any Population (Open-Label Phase)	ICH E3 (list subjects excluded from mITT OL)	SAC [1]
Demographic and Baseline Characteristics					
11.05	Intent-to-Treat Open-label	DM2 / DM4	Demographic and Baseline Characteristics (Open-Label Phase)	ICH E3	SAC [1]
Prior and Concomitant Medications					
11.06	Intent-to-Treat Open-label	CP_CM3 / CP_CM4	Concomitant Medications (Open-Label Phase)	IDSL	SAC [1]

ICH: Listings					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
11.07	Intent-to-Treat Open-label		Concomitant Procedures/Surgeries (Open-Label Phase)		SAC [1]
Exposure and Treatment Compliance					
11.08	Intent-to-Treat Open-label	EX3 / EX4	Treatment Exposure (Open-Label Phase)	ICH E3	SAC [1]
11.09	Intent-to-Treat Open-label		Study Agent Administration (Open-Label Phase)		SAC [1]
Double-blind phase Prior and Concomitant Medications					
11.10	Safety	Listing 1.08 in DB Phase	Concomitant Medications (Double-Blind Phase)	IDSL (Listing of DB Phase concomitant medications, to capture any changes)	SAC [1]
11.11	Safety	Listing 3.08 in DB phase	Concomitant Procedures/Surgeries (Double-Blind Phase)	IDSL (Listing of DB Phase Concomitant Procedures/Surgeries to capture any changes)	SAC [1]
Adverse Events					
13.01	Intent-to-Treat Open-label	AE8 / AE8CP / AE9 / AE9CP	All Adverse Events (Open-Label Phase)	ICH E3	SAC [1]
Serious and Other Significant Adverse Events					
13.02	Intent-to-Treat Open-label		Serious Adverse Events (Open-Label Phase)		SAC [1]
13.03	Intent-to-Treat Open-label		Study Agent Related Adverse Events (Open-Label Phase)		SAC [1]

ICH: Listings					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
13.04	Intent-to-Treat Open-label	AE8 / AE8CPa / AE9 / AE9CPa	Deaths (Open-Label Phase)		SAC [1]
13.05	Intent-to-Treat Open-label	AE8 / AECP8 / AE9 / AE9CP	Adverse Events Resulting in Study Agent Discontinuation (Open-Label Phase)	ICH E3	SAC [1]
13.06	Intent-to-Treat Open-label		Adverse Events of Special Interest (Open-Label Phase)		SAC [1]
Hepatobiliary (Liver)					
13.07	Intent-to-Treat Open-label		Laboratory Results: Liver Function (Open-Label Phase)		SAC [1]
13.08	Intent-to-Treat Open-label		Grade 3 or Grade 4 Laboratory Toxicity Results: Liver Function (Open-Label Phase)		SAC [1]
All Laboratory					
13.09	Intent-to-Treat Open-label		Laboratory Results: Hematology (Open-Label Phase)		SAC [1]
13.10	Intent-to-Treat Open-label		Laboratory Results: Electrolytes (Open-Label Phase)		SAC [1]
13.11	Intent-to-Treat Open-label		Laboratory Results: Other Chemistries (Open-Label Phase)		SAC [1]
13.12	Intent-to-Treat Open-label		Laboratory Results: Urinalysis (Open-Label Phase)		SAC [1]
13.13	Intent-to-Treat Open-label		Laboratory Results: Immunoglobulins (Open-Label Phase)		SAC [1]
13.14	Intent-to-Treat Open-label		Grade 3 or Grade 4 Laboratory Toxicity Results: Hematology (Open-Label Phase)		SAC [1]

ICH: Listings					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
13.15	Intent-to-Treat Open-label		Grade 3 or Grade 4 Laboratory Toxicity Results: Electrolytes (Open-Label Phase)		SAC [1]
13.16	Intent-to-Treat Open-label		Grade 3 or Grade 4 Laboratory Toxicity Results: Other Chemistries (Open-Label Phase)		SAC [1]
13.17	Intent-to-Treat Open-label		Grade 3 or Grade 4 Laboratory Toxicity Results: Urinalysis (Open-Label Phase)		SAC [1]
13.18	Intent-to-Treat Open-label		Grade 3 or Grade 4 Laboratory Toxicity Results: Immunoglobulins (Open-Label Phase)		SAC [1]
Vital Signs					
13.19	Intent-to-Treat Open-label	VS4 / VS5	Vital Signs (Open-Label Phase)	IDSL	SAC [1]
Double-blind Phase Adverse Events					
13.20	Safety	As DB Listing 3.01	All Adverse Events (Double-Blind Phase)	ICH E3 (Listing of DB Phase Adverse Events to capture any changes. If no changes this is not required)	SAC [1]

12.9.10. Non-ICH Listings

Non-ICH: Listings					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
Efficacy					
12.01	Intent-to-Treat Open-label		SRI-S2K Results (Open-Label Phase)		SAC [1]
12.02	Intent-to-Treat Open-label		SELENA-SLEDAI and SS-S2K Results (Open-Label Phase)		SAC [1]
12.03	Intent-to-Treat Open-label		PGA Results (Open-Label Phase)		SAC [1]
12.04	Intent-to-Treat Open-label		BILAG Results (Open-Label Phase)		SAC [1]
12.05	Intent-to-Treat Open-label		Daily Prednisone Dose (Open-Label Phase)		SAC [1]
12.06	Intent-to-Treat Open-label		SFI Flares (Open-Label Phase)		SAC [1]
12.07	Intent-to-Treat Open-label		SLICC/ACR Damage Index Results (Open-Label Phase)		SAC [1]
Health Outcomes					
14.01	Intent-to-Treat Open-label		FACIT-Fatigue Scale Questionnaire (Open-Label Phase)		SAC [1]
Pharmacokinetic					
15.01	Randomized		Serum Belimumab PK Concentration-Time Data	As DB listing 5.01 (Possibly restrict listing to only subjects with updated data since DB phase)	SAC [1]

Non-ICH: Listings					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
Biomarker					
16.01	Intent-to-Treat Open-label		Immunogenicity Results (Open-Label Phase)		SAC [1]
16.02	Intent-to-Treat Open-label		Biomarker Results (Open-Label Phase)	As well as expected OL Biomarker parameters include any unscheduled OL biomarker results e.g. Antinuclear Antibody, B-Lymphocyte Stimulator, Cardiolipin IgA Antibody, Cardiolipin IgG Antibody, Cardiolipin IgM Antibody etc	SAC [1]
16.03	Intent-to-Treat Open-label		B Cell Results (Open-Label Phase)		SAC [1]

Non-ICH: Listings					
No.	**Population**	**IDSL / Example Shell**	**Title**	**Programming Notes**	**Deliverable [Priority]**
16.04	Safety		Biomarker Results continued (Double-blind phase)	Report parameters not delivered in time for DB phase report: • Beta-2 Glycoprotein 1 IgA Antibody, • Beta-2 Glycoprotein 1 IgG Antibody, • Beta-2 Glycoprotein 1 IgM Antibody, • Anti-Ribosomal P Antibody (U/mL), • Sjogrens SS-A Antibody (INDEX), • Sjogrens SS-B Antibody (INDEX), • Anti-RNP.	SAC [1]

12.10. Appendix 10: Example Mock Shells for Data Displays

These will be defined in a separate document.

Division: World Wide Development

Retention Category: GRS019

Information Type: Reporting and Analysis Plan Addendum

Title:	Addendum to the Reporting and Analysis Plan (RAP) for BEL115471, A Phase 3/4, Multi-Center, Randomized, Double-Blind, Placebo-Controlled, 52-Week Study to Evaluate the Efficacy and Safety of Belimumab (HGS1006) in Adult Subjects of Black Race with Systemic Lupus Erythematosus (SLE)

Compound Number: GSK1550188

Effective Date: 08-AUG-2018

Description: This is a Phase 3/4, multi-center, international, randomized, double-blind, placebo-controlled, 52-week study to evaluate the efficacy and safety of belimumab in adult subjects of black race with active systemic lupus erythematosus (SLE). All subjects will receive stable background standard of care therapy throughout the study. Efficacy will be measured by the SLE Responder Index (SRI) at Week 52, defined by a composite SELENA SLEDAI score (Protocol Amendment 2 modified the primary endpoint to the SRI-S2K), Physician's Global Assessment (PGA) and BILAG A and B organ domain scores. In addition, oral corticosteroid use, flares and biomarkers (immunoglobulins, complement, anti-dsDNA and ANA autoantibodies, and B cell subsets) will be assessed. Safety will be assessed by adverse events, clinical laboratory evaluations, immunogenicity, and vital signs. In addition, subjects who successfully complete the double-blind phase may enter into an open-label, 6-month extension phase of the study.

Subject: Systemic Lupus Erythematosus, SLE, belimumab, GSK1550188, Benlysta, black race, efficacy, safety, placebo, SRI, SRI-S2K, SELENA SLEDAI, BILAG, PGA, Lymphostat-B, superiority, logistic regression

Copyright 2018 the GlaxoSmithKline group of companies. All rights reserved.
Unauthorised copying or use of this information is prohibited.

Addendum Author's Name, Title and Functional Area:

PPD
Statistics Leader
Clinical Statistics
Addendum Approved by:

Date

PPD
Principal Programmer
Clinical Programming

Date

PPD
Director Clinical Programming
Clinical Programming

Date

PPD
Clinical Development Director,
Clinical Development

Date

PPD MD
Senior Clinical Development Director,
Clinical Development

Date

PPD
Statistics Director
Clinical Statistics

Date

PPD
SERM Senior Director
Safety Evaluation and Risk Management

Date

PPD
Clinical Development Manager
Clinical Development

PPD **Date**

Clinical Data Manager
Clinical Data Management

TABLE OF CONTENTS

PAGE

ABBREVIATIONS

ADaM	Analysis Data Model
ANCOVA	Analysis of Covariance
BILAG	British Isles Lupus Assessment Group of SLE Clinics
CI	Confidence Interval
CRF	Case Report Form
DBR	Database Release
DO=NR	Dropout = Non-Responder
DO/TF=NR	Dropout/Treatment Failure = Non-Responder
FDA	Food and Drug Administration
ITT	Intention-to-treat
LOCF	Last Observation Carried Forward
LS	Least squares
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified Intention to Treat
PGA	Physician's Global Assessment
PSAP	Program Safety Analysis Plan
RAP	Reporting and Analysis Plan
SAC	Statistical Analysis Complete
SD	Standard deviation
SDTM	Study Data Tabulation Model
SE	Standard Error
SELENA	Safety of Estrogen in Lupus National Assessment
SFI	SLE Flare Index
SLE	Systemic Lupus Erythematosus
SLEDAI	Systemic Lupus Erythematosus Disease Activity Index
SRI	SLE Responder Index
SRI-S2K	SLE Responder Index (SRI) using the SELENA SLEDAI modified with S2K scoring for proteinuria
SS	SELENA SLEDAI
SS-S2K	SELENA SLEDAI (SS) with the modified S2K scoring for proteinuria

Trademark Information

Trademarks of the GlaxoSmithKline group of companies
NONE

Trademarks not owned by the GlaxoSmithKline group of companies
PASS
SAS

1. INTRODUCTION

After finalization of the reporting and analysis plan (RAP), but before database freeze and unblinding of the database to the blinded study team, a response was received from the FDA requesting additional analyses and data displays for BEL115471. This addendum to the RAP documents these additions to the RAP for the BEL115471 study. These additional analyses may not be included as part of the main Statistical Analysis Complete (SAC) package, but will be available soon after. In addition to the FDA requests this RAP Addendum documents some clarifications about BILAG renal scoring and the last visit in the double-blind phase.

The RAP and this addendum were based upon the following study documents:

- Study Protocol Amendment 2 (February 9, 2017)
- Final Case Report Form (CRF) (August 2, 2017)
- Program Safety Analysis Plan (PSAP) Version 5 (December 13, 2017). Note: for reporting purposes, the most current version of the PSAP and associated MedDRA version at the time of database release (DBR) will be used.

2. ADDITIONS TO THE RAP

2.1. Alternate Imputation Approach for Primary Endpoint continuous components

The supportive analyses of the components of SRI-S2K were defined in the RAP to use LOCF imputation. The following endpoints will have their observed data summarized by visit and Week 52 analyses repeated:

- SS-S2K Percent Change from baseline
- SELENA SLEDAI Percent Change from baseline
- SS-S2K Change from baseline
- SELENA SLEDAI Change from baseline
- PGA Percent Change from baseline
- PGA change from baseline

The observed data will be summarized by visit. At Week 52 belimumab and placebo will be compared using an ANCOVA model with treatment group, baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world) as covariates.

Further exploration of the impact of missing data for the above endpoints may be performed post-SAC.

2.1.1. SELENA SLEDAI and SS-S2K

If a subject misses an entire visit, the missing data on SELENA SLEDAI will not be imputed and the score will remain missing. If a subject is a treatment failure, and has post treatment failure efficacy assessments, then the reported post treatment failure SELENA SLEDAI data will be used i.e., the pre-treatment failure data will not be carried forward. The SS-S2K will be managed in the same manner.

If the data for one or more items of the 24 SELENA SLEDAI questions are missing, then the last available answer(s) to the corresponding question(s) from the most recent visit where the corresponding item(s) are non-missing will be assigned to the missing item(s) to obtain a total score. The SS-S2K will be managed in the same manner.

2.1.2. PGA

For the Observed Case analysis if a subject does not have a value recorded for the visit being evaluated, then the PGA value will remain missing. If a subject is a treatment failure and has post treatment failure PGA assessment then this post treatment failure PGA data will be included in the analysis. If the baseline PGA score was zero (or missing) then the percent change from baseline will be missing.

2.2. Evaluation of Treatment Effect on SS-S2K and key secondary endpoints, ignoring treatment failures

As per FDA request the Observed Case Analyses described in Section 11.4.1 “Observed Case Analyses” of the RAP will be performed regardless of the criteria of more than 5% of subjects contributing off treatment efficacy data.

2.3. Point Estimates and 95% CI for Difference in Proportions for SS-S2K Endpoint

The primary analysis will be repeated displaying the point estimates and respective 95% CIs for the proportions and the difference in proportions for the primary endpoint SS-S2K. In addition, the 95% CIs will be included for the key secondary endpoint SRI response. For the proportions of responders in each treatment group the 95% CI for the proportion will be produced using the Wald method (simple asymptotic) without continuity correction. For the difference in proportions the 3 binary covariates will be consolidated to give 8 strata levels and the Cochran-Mantel-Haenszel strata-adjusted proportion difference and stratified Wald 95% confidence intervals will be produced (Kim 2013).

2.4. BILAG Scoring

In Section 9.4.16 of the RAP for the BILAG Renal scoring replace “For items 72 and 73, use item b recorded as urine protein-creatinine ratio (mg/mmol).” with “For item 72 use

item b recorded as urine protein-creatinine ratio (mg/mmol) and for 73, use item a or b recorded as yes/no."

Question 72

As a change from RAP Section 9.3.5 if a subject completed Question 72a "24-hour urinary protein(g)" this value will not be used to calculate the BILAG. This is consistent with Section 9.4.16 of the RAP which refers to 72B but not 72A.

Question 73

Several subjects had 73A completed but not 73B for these subjects replace 73B with 73A in Section 9.4.16 BILAG system Renal function of the RAP.

2.5. Date of last Double-Blind Treatment Phase/Period

This is to supplement Section 9.4.31 and Section 9.4.32 in the RAP. In the ADaM datasets when all parts of that endpoint are missing for a visit then the analysis date will be imputed. Some imputed visit dates (Week 52/Exit in particular) may fall after the date of first open label dose. In this circumstance the imputed visit will be assigned to the double-blind period and/or phase based on the visit number regardless of if the imputed date falls after the first open label treatment.

3. REFERENCE

Kim Y, Won S "Adjusted proportion difference and confidence interval in stratified randomized trials" PharamSUG 2013 Paper SP04

Division: World Wide Development

Retention Category: GRS019

Information Type: Reporting and Analysis Plan

Title:	Reporting and Analysis Plan (RAP) for BEL115471, A Phase 3/4, Multi-Center, Randomized, Double-Blind, Placebo-Controlled, 52-Week Study to Evaluate the Efficacy and Safety of Belimumab (HGS1006) in Adult Subjects of Black Race with Systemic Lupus Erythematosus (SLE)

Compound Number: GSK1550188

Effective Date: 19-Jul-2018

Description: This is a Phase 3/4, multi-center, international, randomized, double-blind, placebo-controlled, 52-week study to evaluate the efficacy and safety of belimumab in adult subjects of black race with active systemic lupus erythematosus (SLE). All subjects will receive stable background standard of care therapy throughout the study. Efficacy will be measured by the SLE Responder Index (SRI) at Week 52, defined by a composite SELENA SLEDAI score (Protocol Amendment 2 modified the primary endpoint to the SRI-S2K), Physician's Global Assessment (PGA) and BILAG A and B organ domain scores. In addition, oral corticosteroid use, flares and biomarkers (immunoglobulins, complement, anti-dsDNA and ANA autoantibodies, and B cell subsets) will be assessed. Safety will be assessed by adverse events, clinical laboratory evaluations, immunogenicity, and vital signs. In addition, subjects who successfully complete the double-blind phase may enter into an open-label, 6-month extension phase of the study. This Reporting and Analysis Plan (RAP) prospectively describes the efficacy and safety analyses that will be performed for the double-blind phase of the study. Details regarding the analyses and summaries for the open-label phase of the study will be reported in a separate RAP.

Subject: Systemic Lupus Erythematosus, SLE, belimumab, GSK1550188, Benlysta, black race, efficacy, safety, placebo, SRI, SRI-S2K, SELENA SLEDAI, BILAG, PGA, Lymphostat-B, superiority, logistic regression

Copyright 2018 the GlaxoSmithKline group of companies. All rights reserved.
Unauthorised copying or use of this information is prohibited.

Author's Name, Title and Functional Area:

PPD
Statistics Leader
Clinical Statistics

Date

Approved by:

PPD
Principal Programmer
Clinical Programming

Date

PPD
Director Clinical Programming
Clinical Programming

Date

PPD
Clinical Development Director,
Clinical Development

Date

PPD MD
Senior Clinical Development Director,
Clinical Development

Date

PPD
Statistics Director
Clinical Statistics

Date

PPD
SERM Senior Director
Safety Evaluation and Risk Management

Date

PPD **Date**

Clinical Development Manager
Clinical Development

PPD **Date**

Clinical Data Manager
Clinical Data Management

TABLE OF CONTENTS

PAGE

ABBREVIATIONS

ACR	American College of Rheumatology
AE	Adverse Event
AESI	AEs of special interest
ANA	Anti-nuclear antibody
ANCOVA	Analysis of Covariance
AUC	Area under the curve
aCL	Anti-cardiolipin
Anti-dsDNA	Anti-double-stranded DNA
ATC	Anatomical Therapeutic Chemical
BILAG	British Isles Lupus Assessment Group of SLE Clinics
BILAG No 1A/2B	No new BILAG A organ domain score or 2 new BILAG B organ domain scores compared with baseline
BMI	Body mass index
CI	Confidence Interval
CMQ	Customized MedDRA Query
CRF	Case Report Form
C-SSRS	Columbia Suicide Severity Rating Scale
DO/TF=NR	Dropout/Treatment Failure = Non-Responder
ENA	Extractable nuclear antigen
FACIT	Functional Assessment of Chronic Illness Therapy
hpf	High-power field
IDSL	Integrated Data Standards Library
IgA	Immunoglobulin A
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IM	Intramuscular
ITT	Intention-to-treat
IV	Intravenous
LLN	Lower Limit of Normal
LOCF	Last Observation Carried Forward
LOQ	Limit of Quantification
LS	Least squares
MedDRA	Medical Dictionary for Regulatory Activities
MCID	Minimally Clinically Important Difference
mITT	Modified Intention to Treat
NMSC	Non-melanoma skin cancer
NTM	Non-tuberculous mycobacterium
PD	Pharmacodynamic
PGA	Physician's Global Assessment
PK	Pharmacokinetic
PP	Per Protocol
PSAP	Program Safety Analysis Plan
PSRQ	Possible Suicidality Related Questionnaire
PT	Preferred Term
RAP	Reporting and Analysis Plan
RBC	Red blood cell

SAE	Serious Adverse Event
SC	Subcutaneous
SD	Standard deviation
SELENA	Safety of Estrogen in Lupus National Assessment
SFI	SLE Flare Index
SLE	Systemic Lupus Erythematosus
SLEDAI	Systemic Lupus Erythematosus Disease Activity Index
SLICC	Systemic Lupus International Collaborating Clinics
SOC	System Organ Class
SRI	SLE Responder Index
SRI-S2K	SLE Responder Index (SRI) using the SELENA SLEDAI modified with S2K scoring for proteinuria
SRT	Safety Review Team
SS	SELENA SLEDAI
SS-S2K	SELENA SLEDAI (SS) with the modified S2K scoring for proteinuria
TB	Tuberculosis
TLFs	Tables, Listings and Figures
ULN	Upper Limit of Normal
WOCF	Worst Observation Carried Forward

Trademark Information

Trademarks of the GlaxoSmithKline group of companies
NONE

Trademarks not owned by the GlaxoSmithKline group of companies
PASS
SAS

1. INTRODUCTION

This reporting and analysis plan (RAP) documents the planned analyses for the BEL115471 study.

This is a phase 3/4, multi-center, international, randomized, double-blind, placebo-controlled study to evaluate the efficacy and safety of 10mg/kg belimumab administered monthly, compared with placebo over a 52-week treatment period in subjects of black race with active systemic lupus erythematosus (SLE) (defined as Safety of Estrogen in Lupus National Assessment Systemic Lupus Erythematosus Disease Activity Index (SELENA SLEDAI) score ≥8).

This RAP is based upon the following study documents:

- Study Protocol Amendment 2 (February 9, 2017)
- Final Case Report Form (CRF) (August 2, 2017)
- Program Safety Analysis Plan (PSAP) Version 5 (December 13, 2017). Note: for reporting purposes, the most current version of the PSAP and associated MedDRA version at the time of database release (DBR) will be used.

2. STUDY OBJECTIVES AND ENDPOINTS

2.1. Study Objectives

- To evaluate the efficacy of belimumab in adult subjects of black race with SLE.
- To evaluate the safety and tolerability of belimumab in adult subjects of black race with SLE.

2.2. Study Endpoints

2.2.1. Primary Endpoint

The primary efficacy endpoint is the Systemic lupus erythematosus Responder Index (SRI) response rate with the modified SLEDAI-2K (S2K) scoring for proteinuria at Week 52. This S2K rule scores proteinuria as 4 points anytime the value is >0.5 g/24hr. This endpoint will be referred to as the SRI-S2K for reporting and is defined as:

- ≥4-point reduction from baseline in SELENA SLEDAI score with the modified SLEDAI-2K scoring for proteinuria (SS-S2K),

AND

- No worsening (increase of <0.30 points from baseline) in Physician's Global Assessment (PGA),

AND

- No new British Isles Lupus Assessment Group of SLE Clinics (BILAG) A organ domain score or 2 new BILAG B organ domain scores compared with baseline at the time of assessment (i.e., at Week 52).

2.2.2. Secondary Endpoints

Major secondary efficacy endpoints:

1. SRI response rate with the SELENA SLEDAI (SS) scoring of proteinuria at Week 52.
2. Time to first severe flare (as measured by the modified SLE Flare Index (SFI); performed with SS as the SLEDAI criterion of the SFI).
3. Percent of subjects whose average prednisone dose has been reduced by ≥25% from baseline to ≤7.5 mg/day during Weeks 40 through 52, in subjects receiving greater than 7.5 mg/day at baseline.

2.2.3. Other Endpoints

Components of SRI and other SRI related Endpoints Supporting Primary Efficacy Endpoint:

- SRI-S2K by visit.
- SRI (without proteinuria correction) by visit.
- Percent of subjects with a ≥4-point reduction from baseline in SS-S2K at Week 52 and by visit (referred to as SS-S2K 4-point reduction).
- Percent of subjects with no new BILAG A organ domain score or 2 new BILAG B organ domain scores compared with baseline at Week 52 and by visit (referred to as BILAG No 1A/2B).
- Percent of subjects without PGA worsening (increase of <0.30 points from baseline) at Week 52 and by visit (referred to as PGA No Worsening).
- Percent of subjects with durable SRI-S2K from Week 44-52.
- Time to the 1st SRI-S2K response that is maintained through Week 52.
- Duration of longest SRI-S2K response among subjects with at least 1 SRI-S2K response.
- SRI5-S2K to SRI8-S2K at Week 52 and by visit.

The SRI5-S2K to SRI8-S2K are defined identically to the SRI-S2K except for using higher thresholds of improvement for SS-S2K reduction for a patient to be declared a responder (e.g., SS-S2K ≥5 point reduction for SRI5-S2K)

Disease activity:

- Percent change and change in PGA by visit.
- Percent change and change in SELENA SLEDAI (modified to use SLEDAI-2K scoring for proteinuria) score by visit (referred to as SS-S2K).

Organ Specific:

- Percent of subjects with SELENA SLEDAI organ improvement by visit. The renal domain will use the SS-S2K proteinuria scoring.
- Percent of subjects with SELENA SLEDAI organ worsening by visit. The renal domain will use the SS-S2K proteinuria scoring.
- Percent of subjects with BILAG organ improvement by visit.
- Percent of subjects with BILAG organ worsening by visit.
- Proteinuria change and percent change by visit among patients with proteinuria at baseline.
- Time to first renal flare over 52 weeks (see Section 9.4.27).
- Percent of patients developing at least one renal flare over 52 weeks.

- Percent of subjects with doubling of serum creatinine (patients whose serum creatinine attains a level double that of the baseline value and is confirmed with a second measurement at least 3 weeks later).

SFI Flare (using SELENA SLEDAI scoring for Proteinuria):

- Time to first SFI flare over 52 weeks.
- Time to first SFI flare after Week 24.
- Time to first severe SFI flare after Week 24.
- Rate of SFI flare per 100 subject years.
- Rate of severe SFI flare per 100 subject years.

Steroids (based on 7-day average at the visits):

- Percent of subjects with daily prednisone dose reduced to ≤7.5 mg/day from >7.5 mg/day at baseline by visit
- Percent of subjects with daily prednisone dose increased to >7.5 mg/day from ≤7.5 mg/day at baseline by visit.
- Percent of subjects with any increase in prednisone use by visit
- Percent of subjects with a 50% decrease in prednisone dose by visit (in patients receiving steroids at baseline).
- Percent of subjects with a 50% increase in prednisone dose by a minimum ≥5 mg/day by visit (in all patients).
- Changes in prednisone dose (mg/day) by visit.

Systemic Lupus International Collaborating Clinics (SLICC)/ American College of Rheumatology (ACR) Damage Index:

- Change from baseline in SLICC Damage Index at Week 52.
- Percent of patients with any SLICC Damage Index worsening (change >0) at Week 52 compared with baseline.

Patient Reported Outcome:

- Change in Functional Assessment of Chronic Illness Therapy (FACIT)-Fatigue score at Week 52 and by visit.
- Percent of patients with improvement in FACIT-Fatigue score exceeding the Minimally Clinically Important Difference (MCID) (≥4) by visit.

Biomarkers:

Percent change and change from baseline in

- absolute B cell subsets (CD19+, CD20+, CD20+/27+ memory, CD20+/27– naïve, CD20+/69+ activated, CD20+/138+ plasmacytoid, CD27+ BRIGHT/CD20– Short-lived plasma cells, CD19+/27BRIGHT/38BRIGHT SLE subset, CD19+/CD24HIGH/CD38HIGH transitional B cell, and CD20-/138+ plasma cells) by visit. (Only in regions/countries where assessment of B cells is feasible.)

Change from baseline in

- CD20+/27– naïve and CD20+/27+ memory B cell subsets expressed as a percentage of CD19.
- Autoantibodies* (anti-dsDNA, ANA, aCL (IgA, IgG, IgM), lupus anticoagulant, beta-2-glycoprotein, anti-sm, anti-RNP, anti-SS-A, anti-SS-B, anti-ribosomal P]).
- Complement (C3, C4) levels.

Percent of patients with normalized serological activity at Week 52 and over time

- IgG, IgM and IgA (high to normal/low).
- Autoantibodies* [anti-dsDNA, ANA, aCL (IgA, IgG, IgM), lupus anticoagulant, beta-2-glycoprotein-1, present to absent.
- Complement (C3, C4, and C3 AND C4) levels low to normal/high.

* Anti-dsDNA will be collected monthly throughout the double-blind phase of the study for SELENA SLEDAI scoring. ANA will be collected at screening and baseline only. Other autoantibodies will be collected from all patients at baseline, then at regular intervals in patients positive at baseline.

2.3. Statistical Hypotheses

The goal of this study is to demonstrate superiority of belimumab over placebo when comparing the SRI-S2K response rate at Week 52 in subjects of black race with SLE. A logistic regression analysis controlling for baseline SS-S2K score (≤9 vs. ≥10), baseline complement level (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs rest of the world) using a two-sided test at the α=0.05 significance level will be used to test the following null hypothesis. Low C3 is defined as less than the lower limit of normal (LLN) which is <90 mg/dL and low C4 is defined as less than the LLN which is <10 mg/dL.

Null Hypothesis

There is no difference between belimumab 10mg/kg and placebo in terms of the SRI - S2K response rate at Week 52 of the double blinded period.

Alternative Hypothesis

Belimumab 10mg/kg differs from placebo in terms of the SRI-S2K response rate at Week 52 of the double blinded period.

3. STUDY DESIGN

This is a Phase 3/4, multi-center, international, randomized, double-blind, placebo-controlled, 52-week study to evaluate the efficacy and safety of belimumab in adult subjects of black race with active SLE. Approximately 501 SLE subjects will be randomized, with a target of about 334 subjects receiving belimumab and 167 subjects receiving placebo.

In addition to receiving stable standard therapy, subjects will be randomized in a 2:1 ratio to treatment with either 10mg/kg belimumab or placebo. At randomization, subjects will be stratified by their screening SELENA SLEDAI score (8-9 vs. ≥10), complement level (C3 and/or C4 low vs. other), and region (US/Canada vs. rest of the world). Subjects will be dosed with study agent on Days 1, 15, 29, and then every 28 days through 48 weeks, with a final evaluation at Week 52 (4 weeks after the last dose). Study agent will be administered intravenously (IV) over 1 hour. Subjects will remain under clinical supervision for 3 hours after completion of the first 2 infusions during the 52-week double-blind phase and 6-month open-label extension.

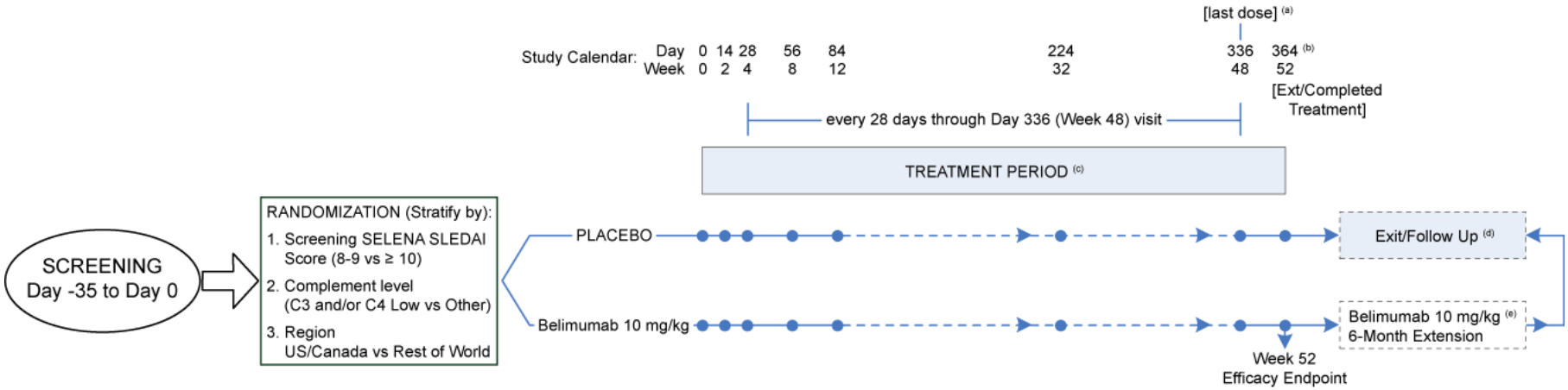
All subjects will continue the stable standard therapy they were receiving during the screening period. Subjects who complete dosing up to 48 weeks will return for a Day 365/Week 52 (final evaluation) visit. Subjects who successfully complete the initial 52-week double blind phase may enter into a 6-month open-label extension phase of this study. The Day 365/Week 52 visit will serve as the Day 1 visit for subjects entering the 6-month open-label extension phase. During the open-label extension phase, all subjects will receive belimumab 10 mg/kg IV every 28 days for 6 months. The first dose of the 6-month open-label extension will be given on the Day 365 (Week 52) visit of the double-blind treatment phase following the completion of all Day 365 (Week 52) assessments. Subjects participating in the extension phase will continue to be monitored for safety and more latitude will be permitted for background medication changes as outlined in Protocol Section 5.5. At the end of the 6-month open-label extension period, subjects who wish to continue treatment may do so by being prescribed commercially available product. If belimumab is not commercially available in a subject's country of participation, the subject may continue to receive belimumab under a separate continuation protocol. Subjects will be required to sign a new informed consent to participate in the separate continuation protocol. Subjects who complete the 52-week double blind phase, but do not enter the extension phase will be required to return for an additional follow up visit 8 weeks after the last dose of study agent. For subjects withdrawing at any other time during the study, an Exit visit (1-4 weeks after the last dose of study agent) and a follow-up visit (approximately 8 weeks after the last dose of study agent) will be performed. The 8-week follow-up visit is not required in subjects entering the separate continuation protocol. In the event that a subject discontinues study agent at any time during the study or withdraws consent, an attempt will be made to ascertain survival status at approximately 52 weeks after the first dose of study agent.

Serum samples for anti-belimumab measurements will be obtained from all randomized subjects before administration of study agent on Day 1, and at any time on Days 57 (Week 8), 169 (Week 24), 365/Exit (Week 52), and at the 8-week follow-up visit. Prior to amendment 2, for any subject who had a positive antibody response at the 8-week follow-

up visit (or last study visit at which immunogenicity was assessed if 8-week follow-up immunogenicity sample is not available), an attempt will be made to obtain a serum sample for anti-belimumab antibodies at least 6 months after the last dose of study agent or after completion and/or unblinding of the study, whichever is later. Amendment 2 removed the follow-up immunogenicity assessment for subjects with a positive anti-belimumab antibody test at the 8-week follow-up visit.

A schematic of the study is provided in Figure 1.

Figure 1 Study schematic



(a)The last dose of study agent is given on the Day 337 (Week 48) visit to subjects NOT participating in the 6-month open-label extension phase of the study.

(b)Subjects participating in the 6-month open-label extension phase of the study are dosed on the Day 365 (Week 52) visit of the double-blind period after the completion of all Day 365(Week 52) assessments. This Day 365 (Week 52) represents the first dose (i.e., Day 1) of the 6-month open-label extension phase. For subjects not participating in the 6-month open-label extension phase of the study, the Day 365 (Week 52) visit serves as the Exit visit with follow-up visits occurring 8 weeks after the last dose of study agent.

(c)The primary treatment period includes 48 weeks of study agent administration (Day 1 to the Day 337 visit) **and** a final visit for the primary endpoint assessment at Week 52 which is 4 weeks after the last dose of the study agent.

(d) An Exit visit (1-4 weeks after the last dose of study agent) and a follow-up visit 8 weeks after the last dose of study agent will be performed for subjects withdrawing at any time during the study. The 8 week follow-up visit is not required in subjects entering the separate continuation protocol.

(e) At the end of the 6 month extension period, subjects who wish to continue treatment may do so by being prescribed the IV commercially available product. If IV belimumab is not commercially available in a subject's country of participation, subjects may continue to receive belimumab administered intravenously every 4 weeks under a separate continuation protocol.

NOTE: This schematic uses Day 0 nomenclature consistent with the protocol. To be consistent with CDISC reporting, add 1 to any Day ≥ 0. For the PLACEBO group, there should be an arrow at the end of the treatment period immediately before the Exit/Follow-up box which points to the Belimumab 10 mg/kg 6-month extension box.

4. PLANNED ANALYSES

4.1. Interim Analyses

No interim analysis is planned for this study.

4.2. Final Analysis

There will be two database locks for this study corresponding to the primary analysis (end of double-blind) and the end of study (end of open-label). The database will be locked for the primary analysis (SRI-S2K at Week 52) after data through the Week 52 visit (or Exit visit for those subjects who withdraw during double-blind treatment) for all subjects have been collected, verified and validated. The second database lock will occur after data for the 6-month open-label phase through the 8-week follow-up period have been collected, verified and validated. All subjects and study site personnel (except the unblinded site pharmacist) will remain blinded until the second database lock. This RAP details the analyses for the double-blind phase, all outputs required for the double-blind analyses will be listed in the Mock Tables, Listings and Figures (TLFs) Shells document. A separate RAP will be created to document the analyses for the open-label phase.

The end of the double-blind treatment phase is defined as follows:

For subjects continuing in the open-label phase:

- The first dose date of open-label medication.

For subjects not continuing in the open-label phase the latest of these dates will be used:

- 8-week follow-up visit date
- Week 52/Exit visit date
- Early termination date

4.3. Changes from Protocol Specified Analyses

- Average daily prednisone dose for each study visit is defined in the protocol as the average dose between visits (Protocol Section 8.5.4). To align with BEL112341 and BEL114055, average steroid use at each visit will be defined as the sum of all prednisone doses over 7 consecutive days up to and including the day of interest, divided by 7, unless otherwise specified.
- Section 8.5.2.1 of the protocol specified that a multivariate analysis will also be performed to evaluate how response to belimumab relative to placebo varied across different categories within relevant baseline demographic or disease characteristics. As the relationship between belimumab response and baseline demographic/disease characteristics is now well understood this analysis will not be performed.

- Section 8.5.3 of the protocol listed the second major secondary endpoint as Time to first severe flare (as measured by the modified SLE Flare Index; with SLEDAI-2K and SELENA SLEDAI as the SLEDAI criterion of the SFI). This is clarified to make it clear that Time to first severe flare, as assessed by the SELENA SLEDAI will be considered the second major secondary endpoint in the testing hierarchy. Time to first severe flare, using the SLEDAI-2K criteria is an equivalent endpoint (not different) and so will not be assessed.
- Due to study conduct concerns at three sites a modified ITT population has been defined. For clarity, the protocol defined ITT population has been renamed the Safety population; see Section 6 for further details.
- The protocol defined sub-group analysis of SRI by Baseline average daily prednisone dose [or equivalent] (≤7.5 mg/day vs. >7.5 mg/day) will not be performed. In previous studies this analysis has not been clinically important as this baseline steroid dose categorization has not been relevant when assessing the risk/benefit of belimumab.
- The protocol definition of renal flares has been modified to remove the hematuria criterion in keeping with the current practice in assessment of renal disease.
- Population Pharmacokinetics as described in protocol Section 8.7.1 are no longer planned. The key patient characteristic which may affect belimumab exposure that is expected to differ between this study and the prior global Phase 3 studies, is body size. The impact of body size on belimumab exposure will be addressed by stratifying the summaries of observed serum concentrations by body weight quartiles and BMI categories.

5. SAMPLE SIZE CONSIDERATIONS

5.1. Sample Size Assumptions

The original protocol required a sample size of 816 subjects with a target of at least 544 subjects in the arm receiving belimumab and 272 subjects in the arm receiving placebo.

This sample size provides at least 90% power at a 5% level of significance to detect a minimum of a 12% absolute improvement in the response rate for the 10 mg/kg belimumab group (assumed rate = 56%) relative to the placebo group (assumed rate = 44%) at Week 52. The selection of the 12% absolute improvement is evidence-based and is based on the observed response rate in the Phase 3 studies [BLISS 76 (HGS1006-C1056) and BLISS 52 (HGS1006-C1057)]. The sample size calculation uses the most conservative estimate for the standard deviation (SD) in the population (i.e., population SD = 50%).

5.2. Sample Size Re-estimation

Following completion of BEL112341 and BEL113750 and the final pooled data from these two studies for the SRI endpoint, the sample size was re-estimated. The protocol was modified (Amendment 2) to amend the sample size to 501 and the primary endpoint to the SRI-S2K.

Approximately 501 subjects will be randomized and treated in the study, with a target of at least 334 subjects in the arm receiving belimumab and 167 subjects in the arm receiving placebo. This sample size provides at least 90% power at a 5% level of significance to detect a minimum of a 15.55% absolute improvement in the SRI response rate with the modified SLEDAI-2K scoring for proteinuria for the 10 mg/kg belimumab group relative to the placebo group (assumed rate =43.95%) at Week 52. This sample size is also sufficient to provide approximately 80% power at a 5% level of significance to detect a minimum of a 13.4% absolute improvement in the SRI response rate with the SELENA SLEDAI scoring for proteinuria for the 10 mg/kg belimumab group relative to the placebo group (assumed rate = 44.8%).

The selection of the assumed treatment differences is based on the observed SRI data from studies BEL112341 (HGS1006-C1115/BLISS SC) and BEL113750 (Northeast Asia), which are two efficacy studies that concluded in 2015 and 2016, respectively, and have nearly identical eligibility criteria to study BEL115471 (HGS1006-C1112, EMBRACE) including requiring a screening SS score ≥8. SRI results, calculated using the SS results and calculated using a modification to the proteinuria scoring based on SLEDAI-2K rules, are shown in Table 1.

Table 1 SRI Results from Studies BEL112341 and BEL113750

	SRI			SRI-S2K		
Dataset	Placebo	Belimumab	Δ	Placebo	Belimumab	Δ
BEL112341*	48.4%	61.4%	12.98%	46.6%	61.7%	15.14%
BEL113750**	40.1%	54.3%	14.17%	40.2%	56.3%	16.16%
Pooled BEL112341 and BEL113750	44.8% (222/496)	58.2% (582/1000)	13.44%	43.95% (218/496)	59.50% (595/1000)	15.55%

* Inclusion criterion SS≥8, 200 mg SC dose used.

**Inclusion criterion SS≥8, 10 mg/kg IV dose used.

All sample size calculations were performed using Power Analysis and Sample Size software (PASS 2012).

6. ANALYSIS POPULATIONS

Screened

The Screened population is defined as all subjects who were screened for the trial, irrespective of whether they were randomized or not.

This study was initiated prior to the requirement to fully document screening failure information. The definitive source of screening events and reasons for screening failure will be provided by Data Management for summary in the CSR, Statistics and Programming will not create any summaries using the screened population.

Randomized

The Randomized population is defined as all subjects who are randomized. Summaries using the Randomized population will group subjects according to the treatment that a subject was randomized to receive, regardless of the actual treatment received.

Safety (referred to as ITT in protocol)

The Safety population is defined as all subjects who are randomized and treated with at least one dose of study treatment. The safety population will be summarized according to the treatment that a subject was randomized to receive, regardless of the actual treatment received. All safety summaries will be performed on the safety population unless otherwise specified.

mITT

During the study three sites PPD and PPD were investigated for potential GCP non-compliance; as a consequence all three sites were terminated by the sponsor. A multi-functional team reviewed these concerns and agreed that data from these three sites should be excluded from all efficacy analyses. The team were confident that the subjects had been dosed and therefore agreed data from these sites should be retained in the safety analyses.

The Modified Intention-To-Treat (mITT) population is defined as the safety population excluding subjects who had any assessment at three sites PPD or PPD The database records subjects under final site and so subjects moved to PPD or PPD but randomized at other sites are excluded. Additionally, subjects randomized at one of the three sites but subsequently moved to another site will be excluded from the mITT. All efficacy analyses will be performed on the mITT population unless otherwise specified.

Per Protocol

Prior to breaking the blind, data for all subjects in the mITT Population will be reviewed to identify protocol violations which could affect the primary endpoint. Subjects with violations with the potential to impact the efficacy analyses will be excluded from the PP Population; see Section 18.8 for definitions.

A sensitivity analysis for the primary efficacy endpoint will be performed on the Per Protocol population if more than 15% of mITT subjects had a violation that could affect the primary efficacy endpoint.

Completers

The Completers population is defined as all subjects who complete all 52 weeks of the planned double-blind treatment period. A sensitivity analysis for the primary efficacy endpoint will be performed on the Completers population. This population is a subset of the mITT population.

As-Treated

The As-Treated population is defined as all subjects who receive at least one dose of study treatment. Summaries using the As-Treated population will group subjects according to the actual treatment administered to the subject. If a subject receives an incorrect treatment, the As-Treated analysis will be performed according to the treatment that the subject receives most of the time (>50% of the time).

A sensitivity analysis using the As-treated population will be performed on the primary efficacy endpoint if more than 15% of subjects received the incorrect treatment. This population is a subset of the mITT population.

PK

The pharmacokinetic (PK) population will comprise all subjects included in the As-Treated population for whom at least one post belimumab treatment PK sample was obtained and analyzed. Summaries using this population will be based on the actual treatment received if this differs from that to which the subject was randomized.

6.1. Analysis Datasets

Drop Out/Treatment Failure = Non-Responder (DO/TF=NR)

The DO/TF=NR dataset (as described in Section 9.1 and Section 9.2) will be used for the primary response endpoint and each of the three components of the primary response, as well as the majority of other responder endpoints. The basic premise of the DO/TF=NR is that a subject who drops out prior to the Week 52 visit (and does not have a visit with assessments in a ± 28 day window around the Week 52 visit) and/or uses a prohibited medication or a non-allowed dose of a restricted medication resulting in a treatment failure will thus be counted as a non-responder in the analysis. Note, this analysis dataset was referred to in the protocol as the DO=F or Drop Out=Failure dataset which did not make it clear that the intent was that treatment failures were also to be imputed as non-responders. DO/TF=Non-Responder was the intention and is consistent with the analyses of previous belimumab studies.

Last Observation Carried Forward (LOCF)

The LOCF principle is applied whereby missing values will be replaced with the last previous non-missing value. If the first assessment of the Treatment period is missing then the missing data will be imputed with the baseline value.

If a subject withdraws or takes a protocol-prohibited medication or a dose of allowable (but protocol restricted) medication that results in treatment failure designation (see Section 9.1) prior to the study visit being evaluated, the LOCF value will be handled by using the result from the last visit on or prior to the date of withdrawal or date of treatment failure, whichever is first.

If any observed dates exist, the date for the LOCF record will be determined from the observed dates. If there are multiple observed dates, the latest of these dates (closest to the scheduled visit) will be taken as the derived date. If the record is completely missing then the date will be missing, otherwise missing dates will be imputed as the (treatment start date + planned study day [e.g. Week 8 visit planned study day as day 1+ 8*7=57 day]).

Observed

Observed data are the data collected or observed for the subject with no imputation for missing data. The observed case analysis includes all data collected including data post IP withdrawal and post treatment failure.

Worst Observation Carried Forward (WOCF)

The worst observation carried forward (WOCF) principle will be applied to the SLICC/ACR Damage Index data at an item level. As the SLICC/ACR Damage Index is meant to be a monotonic, non-decreasing value (except for renal domain), items that decrease over the course of the study will be replaced with the highest previous item score achieved among previous visits (including baseline). Within the renal domain, damage scored on the first two items (estimated glomerular filtration rate < 50%, proteinuria > 3.5 gm/24hours) should only be carried forward to Week 52 if the third item (end stage renal disease) has not been scored at Week 52. The reason for this is that scores on the first two items can decrease but only if the third item is then scored

7. TREATMENT COMPARISONS

The primary comparison of interest is the comparison between belimumab and placebo for the SRI-S2K response at Week 52 in the mITT population at the 0.05 significance level imputing a non-responder for dropouts and treatment failures as described in Section 6.1.

Other comparisons of interest include components of the primary endpoint, SRI at Week 52, Time to first severe flare, steroid reduction.

For further details on the hierarchy of comparisons see Section 8.4.

7.1. Data Display Treatment and Other Sub-Group Descriptors

The following treatment descriptors (Table 2) will be used on all data tabulations for the double-blind phase:

Table 2 Treatment Descriptors, Colors, Line Style and Symbols for Reporting in Double-Blind Phase

Treatment Descriptor	Color	SAS Color	Line Style	Symbol
Belimumab 10mg/kg	Blue	CX0000FF	Dashed	Triangle (filled)
Placebo	Black	CX000000	Solid	Circle (open)

For tables presenting baseline demographics and characteristics described in Section 10, a total column for both treatment groups combined will also be presented for subjects in the double-blind phase.

8. GENERAL CONSIDERATIONS FOR DATA ANALYSES

- All data summaries and analyses will be performed using the latest available version of SAS software (as available at GSK). R may be used to quality check statistical analyses. The latest available version of R will be used and the version will be documented in the CSR.
- Data displays will follow the shells outlined in Attachment 1 and Attachment 2 which will follow the Benlysta program standards and, as far as possible, follow the GSK Integrated Data Standards Library (IDSL).

Unless otherwise stated, the following will apply:

- Continuous variables will be summarized with the statistics mean, median, standard deviation (SD), 25th and 75th percentiles, minimum and maximum.
- Categorical variables will be summarized with frequency counts and percentages, or proportions where specified. A 'Missing' category will be added to frequency counts if there is at least one missing record.
- Percentages will be calculated using the number of non-missing observations as the denominator. If the unit of measurement is a subject, the percentage will be based on the total N for the population that are still in the study at the respective visit
- The minimum and maximum will be reported to the same number of decimal places as the raw data recorded in the database.
- The mean, median, 25th percentile and 75th percentile will be reported to one more decimal place than the raw data recorded in the database.
- The SD will be reported to two more decimal places than the raw data recorded in the database.
- A maximum of four decimal places will be used for any statistics displayed.
- The same decimal point rules described above apply to scores calculated in the derived datasets.
- Percentages will be presented to one decimal place.
- A count of zero will have no corresponding percentage.
- For statistical analyses, all tests will be two-sided and p-values will be presented to a maximum of 4 decimal places.
- When the data are summarized by visit, only scheduled visits will be presented. Unscheduled visits will be included in patient listings
- Listings will be sorted by treatment group, site ID, unique subject ID, and visit (where appropriate).
- Figures displaying means or medians also will include standard error bars or interquartile ranges, respectively.

- Distributions will be reviewed and if there is significant evidence of skewness, medians will be used as the summary measure instead of means; in this case the corresponding figures will display medians.

8.1. Multicenter Studies

This is a multi-center trial and subjects will be centrally randomized. Analyses will not be adjusted for center.

8.2. Other Strata and Covariates

- Randomization will include stratification by subjects' screening SELENA SLEDAI score (8- 9 vs. ≥10), complement level (C3 and/or C4 low vs. other), and region (US/Canada vs. rest of the world). Note, no subjects were enrolled in Canada but cross program region subgroup labels have been retained for consistency.
- The screening value is used for stratification for SELENA SLEDAI and complement level as these factors require the results of lab tests to be available to the sites prior to randomization (baseline).
- Analyses will control for baseline values of SELENA SLEDAI and complement as the baseline values represent the status of the subject at the time study treatment is first administered.
- To be consistent with the primary endpoint, the SS-S2K at baseline will be used to control for baseline disease activity in the statistical models (except for sensitivity analyses using SS, in these cases baseline SS will be used).

8.3. Examination of Subgroups

The comparison of the primary efficacy endpoint between belimumab and placebo will be performed by the following subgroups:

- Baseline SS-S2K score (≤9 vs. ≥10)
- Baseline anti-dsDNA (≥30 IU/mL vs. <30 IU/mL)
- Baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low, where low is defined as less than the LLN which is <90 mg/dL for C3 and <10 mg/dL for C4)
- Baseline complement & anti-dsDNA (At least one C3/C4 low and anti-dsDNA >=30 IU/mL vs. Not (At least one C3/C4 low and anti-dsDNA >=30 IU/mL)
- Region (US/Canada vs. Rest of World)
- Age Group (<65 years, ≥65 years)
- Baseline BLyS Protein (≥75th Percentile vs. <75th Percentile).

Gender sub-group analysis will not be performed due to the low number of male subjects.

Subgroup analyses will be adjusted for the relevant covariates (however some covariates may have to be dropped due to the convergence issues for small subgroups). The p-

values for the treatment by subgroup interaction effect will be presented and evaluated for significance at the α=0.10 significance level.

8.4. Multiple Comparisons and Multiplicity

For the analysis of the primary and the major secondary efficacy endpoints, a step-down sequential testing procedure will be used to control the overall type 1 error rate. With this procedure, the primary and three major secondary endpoints will be evaluated for statistical significance based on a pre-specified sequence for interpretation: (1)SRI-S2K response rate at Week 52 (2) SRI response rate with the SELENA SLEDAI for scoring of proteinuria at week 52, (3) time to first severe SFI flare using the SELENA SLEDAI without the proteinuria adjustment, and (4) percent of subjects with average prednisone dose that has been reduced by ≥25% from baseline to ≤7.5 mg/day during Weeks 40 through 52. Specifically, endpoints will be tested in the sequence above (2-sided alpha=0.05) provided that statistical significance is achieved by all prior tests. If at any point in the sequence statistical significance is not met, then subsequent endpoints in the sequence cannot be deemed statistically significant, although nominal p-values will be reported and considered descriptive.

Analyses of other efficacy endpoints other than the major secondary efficacy endpoints will not be subject to any multiple comparison procedure. All statistical tests will be 2-sided and performed at an overall significance level of 0.05.

9. DATA HANDLING CONVENTIONS

This section describes data handling conventions for the efficacy data during the double-blind treatment period including handling of premature withdrawal and missing data. Missing data for analyses at time points other than Week 52 will be managed in a similar manner.

9.1. Treatment Failures

A treatment failure is defined as any subject who:

- receive a protocol-prohibited medication or a dose of allowable (but protocol restricted) medication that results in treatment failure designation prior to Week 52.

The treatment failure rules are detailed in the protocol (Section 5.5) with the programming rules further clarified in Appendix 3.

9.2. Dropouts

For the purposes of the analysis imputation, a dropout is defined as any subject who:

- withdraws from the study prior to Week 52 and has no visit within ± 28 days of the target Week 52 visit date (excluding follow up visits).

This rule is applied consistently across all efficacy endpoints. The assessment of whether or not a subject had a visit within the ± 28 day of Week 52 is performed at the domain level [e.g., a subject who had visit within 28 days for SLEDAI but not SLICC would be assessed on their observed data for SLEDAI but would be considered a dropout in SLICC].

9.3. Missing Data Rules

9.3.1. Missing Dates

Element	Reporting Detail
Concomitant Medications	Medication Start Date (CMSTDT) CMSTDT is imputed as TRTSDT unless: • CMENDT is < TRTSDT (whether CMENDT is a complete (DD/MM/YY) or partial date (some combination of CMENDT day, month or year imputed) OR • The month and/or month and year of the partial CMENDT date are before the month and/or year of TRTSDT OR • "Taken prior to study?" is checked. Medication End Date (CMENDT) • End dates for concomitant medications will not be imputed, and the medication will be considered ongoing.

Element	Reporting Detail
Adverse Events	• The eCRF does not allow for the possibility of partial AE dates. • Completely missing start or end dates will remain missing, with no imputation applied. Consequently, time to onset and duration of such events will be missing. AEs with missing start dates will be considered as treatment emergent.

9.3.2. Partial Dates

Element	Reporting Detail
General	Partial dates will be displayed as captured in subject listings.
Concomitant Medications	<u>Medication Start Date (CMSTDT)</u> • If CMENDT is missing OR CMENDT >= TRTSDT (whether CMENDT is a complete or partial date) then CMSDT is imputed as the TRTSDT. • If CMENDT is < TRTSDT (whether CMENDT is a complete or partial date) then CMSTDT is imputed with JAN for missing month and 01 for missing day, whatever is applicable. <u>Medication End Date (CMENDT)</u> • If month and year are present, then set to the earlier of (last contact date or last day of that month of the CMENDT). • If only year present, then set to the earlier of (31DEC of the year or last contact date).
SLE Disease Duration	• For records where month and day are missing for start date, impute with 01 for day and January for month to assume that the duration was the longest possible duration. • For records where the day only is missing for start date, impute with 01 for day to assume that the duration was the longest possible duration.

9.3.3. SRI-S2K

For the SRI-S2K endpoint and its components (SS-S2K 4-point reduction from baseline, PGA no worsening, BILAG no new 1A/2B), any subject who is classified as a drop out or a treatment failure will be considered a non-responder for the primary efficacy analysis and the supportive analyses of the primary efficacy endpoint. This imputation method is referred to as "Drop Out/Treatment Failure = Non-Responder" (DO/TF=NR). If baseline is missing for any of the components of the SRI-S2K, then the SRI-S2K will be missing. Subjects who have a SS-S2K score <4 at baseline will be excluded from this analysis as they have no opportunity to meet the responder criteria.

Any subject not otherwise classified as a drop out or treatment failure who misses the Week 52 visit will be handled as follows:

- If the subject does not have a visit within ± 28 days of Day 365 (Week 52) Visit, the subject will be considered a dropout for the Week 52 analysis.

- If a subject has 1 visit within ± 28 days of the target Week 52 visit, the data from that visit will be used for the Week 52 primary efficacy analysis.
- If a subject has more than 1 visit within ± 28 days of the target Week 52 visit, the data from the visit closest to the target Week 52 visit will be used for the Week 52 primary efficacy analysis.
- If a subject has 2 visits with equal distance within ± 28 days of the target Week 52 visit, the data from the visit prior to the Week 52 visit will be used for the Week 52 primary efficacy analysis. If a subject has a visit within the required window, but partial data of the primary efficacy endpoint are missing (including individual items of any component of the primary endpoint), the LOCF method will be used for the missing item or component. This will be modified for laboratory items in the BILAG for which scoring is dependent on both the actual score and the change from the previous visit as described in Section 9.3.5.

9.3.4. SELENA SLEDAI and SS-S2K

The LOCF method will be employed for subjects with missing data on SELENA SLEDAI (SS) score at the visit being evaluated. The LOCF method is consistent with how missing data in previous trials was handled. Inclusion of this methodology will enable comparison to previously reported studies. See Section 6.1 for further details Specifically, if a subject misses a scheduled visit or if partial data are missing from a subject's visit, the missing data will be handled by using the last observation (or item) carried forward method. For example, if the data on one or more items of the 24 SELENA SLEDAI questions are missing, the last available answer(s) to the corresponding question(s) from the most recent visit where the corresponding item(s) are non-missing will be assigned to the missing item(s) to obtain a total score. If a subject misses an entire visit, the missing data on SELENA SLEDAI will be handled by using the last score from the previous visit. However, if a subject withdraws or is a treatment failure prior to the study visit being evaluated, the SELENA SLEDAI data will be handled by using the score from the last visit on or prior to the date of withdrawal or date of treatment failure, whichever is first. The SS-S2K will be managed in the same manner.

9.3.5. BILAG

The LOCF method will be employed for subjects with missing data on BILAG score at the visit being evaluated, see Section 6.1 for further details. Specifically, if a subject misses a scheduled visit or if partial data are missing from a subject's visit, the missing data will be handled by using the last observation (or item) carried forward method. For example, if the data on one or more items of the 86 BILAG questions are missing, the last available answer(s) to the corresponding question(s) from the previous visit will be assigned to the missing item(s) to obtain a score for each organ system domain.

For the following laboratory items, both the actual value from the last visit and the change observed at that visit will be carried forward:

- 24-hour urinary protein (g) (Question 72a),

- urine protein-creatinine ratio (mg/mmol) (Question 72b),
- serum creatinine (mg/dl) (Question 75), and
- creatinine clearance (ml/min) (Question 76).

If a subject misses an entire visit, the missing data on BILAG will be handled by using the last score from the previous visit. However, if a subject withdraws or is a treatment failure prior to the study visit being evaluated, the BILAG data will be handled by using the score from the last visit on or prior to the date of withdrawal or the date of treatment failure, whichever is first.

9.3.6. PGA

The LOCF method will be employed for subjects with missing data on PGA score at the visit being evaluated, see Section 6.1 for further details. Specifically, if a subject misses the visit being evaluated; the missing data will be handled by using the last observation available. However, if a subject withdraws or is a treatment failure prior to the visit being evaluated, the PGA data will be handled by using the score from the last visit on or prior to the date of withdrawal or date of treatment failure, whichever is first.

9.3.7. Proteinuria

All proteinuria endpoints will use Observed data.

9.3.8. FACIT Fatigue

The LOCF method will be employed for subjects with a missing total score data on the FACIT Fatigue scale. Specifically, if a subject misses a scheduled visit, missing data will be handled by using the last observation carried forward method.

If a subject misses an entire visit, the total score from the previous visit will be carried forward. However, if a subject withdraws or is a treatment failure prior to the study visit being evaluated, the FACIT Fatigue data will be handled by using the score from the last visit on or prior to the date of withdrawal or treatment failure date whichever is first.

Details for calculating FACIT for items that are not scored at any visit are described in Section 9.4.29.

9.3.9. FACIT Fatigue Scale Score Improvement Exceeding the MCID (>=4 points)

The DO/TF=NR method will be employed for subjects with missing data on the endpoint for FACIT Fatigue Scale score improvement exceeding the MCID. For visits occurring on or after study dropout or treatment failure, the endpoint will be set to ‘No’.

9.4. Derived and Transformed Data

9.4.1. Baseline

The baseline value of a variable is defined as the last available value measured prior to dosing on or before the date of first dose (Day 1), unless otherwise specified. If the last available value occurs on Day 1 but the time of the assessment is not collected, then the assessment will be assumed to be prior to dosing. For concomitant medications and adverse events, these are considered present at baseline if the start date is prior to Day 1 and the end date is on or after Day 1. Medications or events with a start date on Day 1 are considered as being on-treatment and treatment-emergent, respectively; see Appendix 3 Steroids or Anti-Malarials section for slightly different baseline definition comparing Steroids and non-Steroid medications.

Baseline flares are flares that occur on or prior to the treatment start date.

The protocol specifies Day 0 as Baseline/Treatment Start date, but the CDISC standard is to refer to the Baseline/Treatment Start date as Day 1; therefore Baseline/Treatment Start date will appear as Day 1 in the displays and will be referenced as Day 1 in this document. A table indicating the target study day for each planned visit starting at Day 1 (instead of Day 0) is in Section 9.4.3.

9.4.2. Study Day

Study Day is the number of days from the treatment start date to a study date of interest (e.g., adverse event start date) and is calculated as follows:

If condition is…	Then Study Day is…
study date< treatment start date	study date – treatment start date
study date is ≥ treatment start date	study date – treatment start date +1

Note: Study Day cannot be zero. If either date is missing then Study Day is missing.

9.4.3. Analysis Visit and Analysis Visit Number

The data are analyzed according to the planned visit assignment.

Exit/withdrawal visits (listed below) must be slotted to the appropriate planned visit according to the study phase. This will only be done for subjects who withdrew prior to the end of the respective study phase. Unscheduled laboratory visits will also be remapped into the appropriate planned visit based on the original visit number (e.g. if unscheduled visit number =80.01, set analysis visit number= 80). The following hierarchy rules may apply for all the analysis flag in the dataset:

1. If Scheduled visit exists, take as the first hierarchy;
2. If no scheduled visit exists, take the unscheduled visit closest to the target study day.
3. If the Exit visit was slotted into the same visit as the visit prior to withdrawal, select the visit closest to the target study day. For instance, if subject already had scheduled

Week 20 assessment, and the Exit visit assessment was mapping as another Week 20 assessment, select the visit closest to the target study day.

Phase	Visit	Visit Number
Double-Blind	Week 52/EXIT	160
Open-Label Extension	OL Week 28/OL EXIT	240

The Analysis Visit is assigned based on the interval in which the Study Day for the exit/withdrawal visit falls according to intervals (inclusive) provided below. For completeness, the table also includes visits which are not slotted; these visits will have 'na' for 'not available' listed for the Interval Start and End Day.

Analysis Visit	Analysis Visit Number [a]	Target Study Day[b]	Interval Start Day	Interval End Day
Screening	10	-35	na	na
Double-Blind visits:				
Week 0	20	1	-35	1
Week 2	30	15	2	21
Week 4	40	29	22	42
Week 8	50	57	43	70
Week 12	60	85	71	98
Week 16	70	113	99	126
Week 20	80	141	127	154
Week 24	90	169	155	182
Week 28	100	197	183	210
Week 32	110	225	211	238
Week 36	120	253	239	266
Week 40	130	281	267	294
Week 44	140	309	295	322
Week 48	150	337	323	350
Week 52	160	365	351	378
Survival[c]	163	365	na	na
Follow-up[d]	170	393	na	na
Open-label extension visits:				
EXT Week 4	180	393	379	406
EXT Week 8	190	421	407	434
EXT Week 12	200	449	435	462
EXT Week 16	210	477	463	490
EXT Week 20	220	505	491	518
EXT Week 24	230	533	519	546
EXT Week 28	240	561	547	574
EXT Follow-up[e]	250	589	na	na

Analysis Visit	Analysis Visit Number [a]	Target Study Day[b]	Interval Start Day	Interval End Day

a. If there are multiple visits within a visit window, the visit closest to the target date will be used. If there are two visits equidistant from the target date, the first will be used.
b. Study Day with Baseline/Treatment Start Date as Day 1. Baseline record will be derived in each analysis dataset, analysis visit number will be 15.
c. Survival visit for subjects who withdraw from the double-blind phase but consent to the Week 52 survival assessment.
d. The double-blind follow-up occurs 4-weeks after early withdrawal or 8 weeks after last dose (e.g. Week 56) for subjects who complete and do not enter the open-label extension phase.
e. The open-label extension follow-up occurs 4-weeks after early withdrawal or 8 weeks after last dose (e.g. Week 32) for subjects who complete the open-label extension phase.

9.4.4. Change from Baseline

Change from baseline will be calculated as:

Visit value – baseline value.

If either value is missing the change from baseline will be missing.

9.4.5. Percent Change from Baseline

Percent change from baseline will be calculated as

$$\frac{\textit{Visit Value - Baseline Value}}{\textit{Baseline Value}} \; X \; 100.$$

Subjects with a baseline value of zero will not have a value calculated due to division by zero. If the baseline value is zero or missing then the percent change will be set to missing.

9.4.6. SLE Disease Duration

SLE Disease Duration (years) is defined as

$$\frac{\textit{Treatment Start date - SLE diagnosis date + 1}}{\textit{365.25}}$$

If either date is missing then SLE Disease Duration will be missing.

9.4.7. Analysis Age

Analysis age is derived in years relative to the treatment start date for randomized subjects, and is calculated in SAS as follows:

INTCK ('YEAR', Date of birth, **Date of first treatment**, 'C'),

in the case that a subject was randomized but not treated then age is defined as

INTCK ('YEAR', Date of birth, **Date of randomization**, 'C').

9.4.8. Body Mass Index (BMI)

Baseline body mass index (BMI) will be calculated from baseline weight and height measurements as:

$$BMI\ (kg/m^2) = \frac{Weight\ (kg)}{[Height\ (m)]^2}$$

Since height is collected in centimetres (cm), it must to be converted to meters (m) by dividing by 100 before using it in the formula above. If weight or height is missing, then BMI will be missing.

9.4.9. Race Hierarchical Rule

If multiple race categories are checked on the CRF, the subject will be assigned to a unique race group based on which of the races checked appears first in the list below:

- Black/African American/African Heritage
- Native Hawaiian or Other Pacific Islander
- American Indian or Alaska Native
- Asian
- White/Caucasian

For example, if Black and Asian are both checked, then the subject will be assigned as Black since it appears highest in the list. Race assigned based on this hierarchical rule will be applied to all analyses related to race. In the baseline demographic characteristics table and the race and racial combination details table, subjects with multiple race categories checked will be reported in the race per the hierarchical rule as well as in the multiracial category.

9.4.10. ACR Criteria at Baseline

- Subjects were required to meet at least 4 of the 11 criteria to be eligible for the study.
- The criteria consist of the following: malar "Butterfly" rash, discoid rash, photosensitivity, oral ulcers, arthritis, serositis, renal disorder, neurologic disorder, hematologic disorder, immunologic disorder, and anti-nuclear antibody (ANA). See Appendix 4 for an example of the form.
- Sub-criteria exist for serositis, renal disorder, neurologic disorder, hematologic disorder, and immunologic disorder.
- The number of ACR criteria met will be summed for a total score with a maximum possible score of 11.
- Sub-criteria are not summed as part of the total score, but if at least one sub-criterion is met for the overall criterion category, then the category will count as one for the total score. For example, the neurologic criterion includes sub-criteria of seizures

and psychosis. Neurologic will only count as one toward the total ACR criteria even when both seizures and psychosis are applicable.

9.4.11. B Cell Unit Conversions and Normalization for Rare B Cell Subsets

- The Benlysta program standard is to report common B cells (CD19, CD20, naïve, and memory) in counts per μL (/μL).
- To convert values reported as count per GI/L (= 10^9/L) to μL (= cells/mm^3) simply multiply the value by 10^3 or 1000

  Example: (0.25 GI/L) x (1000) = 250 /μL

- Rare B cell subsets reported in GI/L (=10^9/L) will be normalized and converted to cells/ml using the following formula:

  Normalized count/mL = [(rare cell event count) / (CD19+ event count)] * (CD19+ count per mm3 or μL) * 1000.

- Note that the CD19+ concentration from the TBNK panel should be converted to count/μL prior to the normalization and conversion of the Rare B-cell subset.

For additional detail, please see Section 8.6.3 of the PSAP located in IMMS at the following path: /Study File/GSK1550188/_Project/Meta Analysis/PSAP.

The list of B cell item mappings and display labels are listed in Appendix 13.

Endpoints to be normalized and values to be substituted into the formula are given in Appendix 13.

Table 3 Rare B cell subsets requiring normalization

Rare B cell subset	Rare cell event count	CD19+ Event Count [1]	CD19+ Counts [2]
Activated	CD20+CD69+ (EVENTS)	CD19+ (EVENTS) [a]	CD19+ (/μL)
Plasmacytoid	CD20+ CD138+ (EVENTS)	CD19+ (EVENTS)[a]	CD19+ (/μL)
Plasma	CD20- CD138+ (EVENTS)	CD19+ (EVENTS) [a]	CD19+ (/μL)
Short-lived plasma	CD27+b CD20- (EVENTS)	CD19+ (EVENTS) [a]	CD19+ (/μL)
SLE subset	CD27+CD38+CD19+ (EVENTS)	CD19+ (EVENTS) [a]	CD19+ (/μL)
Transitional	CD19+ CD24b+ CD38b+ CD27- (EVENTS)	CD19+ (EVENTS) [b]	CD19+ (/μL)

[1] From corresponding panel – [a] Plasma panel, [b] Transitional panel.
[2] Source data require conversion from GI/L to /μL.

The required parameters in the source data can be identified in Table 4.

Table 4 SDTM data required for Normalization of Rare B Cell Subsets (SDTM.LB)

Rare B cell subset	Lab Test Code (LBTESTCD)	Lab Test (LBTEST)	B cell Panel (LBMETHOD) [1]	Units of Measurement (LBORRESU)
CD19 Concentration from TBNK Panel				
CD19[2]	CD19	CD19	FLWTBNK	10^9/L [2]
CD19 Event Counts from Plasma/Transitional Panels				
CD19 Event (Plasma B cell panel)	CD19E	CD19 Number of Events	FLWPLSM	EVENTS
CD19 Event (Transitional B cell panel)	CD19E	CD19 Number of Events	FLWTRANS	EVENTS
Rare B Cell Events				
Activated	CDX141E	CD20+CD69+ Number of Events	FLWPLSM	EVENTS
Plasmacytoid	CDX145E	CD20+CD138+ Number of Events	FLWPLSM	EVENTS
Plasma	CDX143E	CD20-CD138+ Number of Events	FLWPLSM	EVENTS
Short-lived plasma	CDX154E	CD27+bCD20- Number of Events	FLWPLSM	EVENTS
SLE subset	CDX156E	CD27+CD38+CD19+ Number of Events	FLWPLSM	EVENTS
Transitional	CDX199E	CD19+ CD24b+ CD38b+ CD27- Number of Events	FLWTRANS	EVENTS

[1] FLWPLSM=FLOW CYTOMETRY - PLASMA B CELLS PANEL; FLWTRANS= FLOW CYTOMETRY - TRANSITIONAL B CELLS PANEL, and LWTBNK=FLOW CYTOMETRY - TRUCOUNT TBNK (BD MULTITEST) PANEL.

[2] Source data require conversion from GI/L (=10^9/L) to /μL.

9.4.12. Proteinuria

For analysis, urine protein in g/24-hour will be approximated by the urine protein:creatine ratio (uPCR) in mg/mg which is stored in the variable for the original unit in the SI dataset. The SI dataset standard unit is reported in mg/mmol and is used for scoring the BILAG only.

9.4.13. Average Daily Prednisone Equivalent

- For steroid use analyses, all steroid dosages are converted to a prednisone equivalent in milligrams; therefore, analyses refer to average daily prednisone dose instead of average daily steroid dose. See Appendix 2 for instructions on converting steroids to prednisone equivalents.

- This document will be updated throughout the course of the study to include all corticosteroids taken by subjects and will be finalized prior to the first database lock and unblinding.
- The average daily prednisone dose takes account of all steroids taken intravenously (IV), intramuscularly (IM), SC, intradermally, and orally for both SLE and non-SLE reasons.
- Please note: the prednisone visit slotting is based on the earliest date of (exposure dose date at the visit, primary efficacy assessment date [BILAG, PGA, SLEDAI]), if these are missing then use target visit date.

9.4.13.1. Baseline Prednisone Dose

At baseline, the average daily prednisone dose is the sum of all prednisone doses over 7 consecutive days up to, but not including Day 1, divided by 7. This baseline is used for comparison to post baseline average dose between visits and average dose at the visit.

9.4.13.2. Average Daily Prednisone Dose Between Visits

The average daily prednisone dose between visits will be calculated for each scheduled post-baseline visit by summing all prednisone doses since the previous visit (previous visit date +1) up to and including the current visit and then dividing by the number of days in this period (Date of current visit – Date of previous scheduled visit). Days on which a subject does not have prednisone use recorded will be considered as 0 mg for the day in the calculation of average daily prednisone dose.

For subjects who withdraw from the study or are deemed treatment failures prior to a scheduled visit, the average dose will be calculated as the average of all prednisone doses since the previous visit up to and including the date of withdrawal or date of treatment failure, whichever is earlier. For a subject who withdrew/treatment failed at baseline then the average dose will be the baseline value.

9.4.13.3. Average Daily Prednisone Dose at the Visit

While on treatment, the average daily prednisone dose at the visit is the sum of all prednisone doses over 7 consecutive days up to and including the day of interest, divided by 7, unless otherwise specified. Days on which a subject does not have prednisone use recorded will be considered as 0 mg for the day in the calculation of average daily prednisone dose.

The average daily prednisone equivalent dose will be calculated for each scheduled visit. For subjects who withdraw from the study or are treatment failures prior to a scheduled visit, the average dose will be calculated at the date of withdrawal or date of treatment failure, whichever is earlier. For a subject who withdrew/treatment failed at baseline then the average dose will be the baseline value.

Note: for the treatment failure adjudication, an intermediate dataset containing the rolling 7-day average daily prednisone for each study day will be used to capture treatment failures occurring between visits.

9.4.13.4. Average Daily Prednisone Dose during Weeks 40 to 52

DO/TF

While on treatment, the average daily prednisone dose during Weeks 40 to 52 is the average of all prednisone doses starting the day after the Week 40 visit date up to but not including the Week 52 visit date.

For subjects who withdraw or are treatment failures on or before the Week 40 visit date, the average daily prednisone dose during Weeks 40 to 52 will be missing. For subjects who withdraw or are treatment failures after the Week 40 visit date, the average daily dose will be calculated as the average of all prednisone doses from the day after the Week 40 visit date up to and including the date of withdrawal or date of treatment failure, whichever is earlier.

Observed

A sensitivity analysis may be performed using all observed data. The average daily prednisone dose during Weeks 40 to 52 is the average of all prednisone doses starting the day following the Week 40 visit date up to but not including the Week 52 visit date. All data in the interval is used irrespective of drop out (IP withdrawal) or treatment failure.

9.4.13.5. Prednisone Increase by ≥50% from Baseline by a Minimum ≥5 mg/day

Subjects are considered to have a prednisone increase by ≥50% and a minimum of ≥5mg/day if the percent change from baseline in daily prednisone dose is ≥50% AND the change from baseline in daily prednisone dose is ≥5mg/day. For subjects with no prednisone use at baseline, a subject meets the definition if the change from baseline in prednisone use is ≥ 5mg/day; percent change from baseline cannot be calculated when there is no baseline prednisone use due to division by zero.

9.4.13.6. Prednisone Reduction by ≥50% from Baseline

Subjects are considered to have a prednisone reduction if the prednisone percent change from baseline is ≤ -50%. To have a reduction subjects must have a baseline average daily prednisone dose >0.

9.4.13.7. Any Increase from Baseline

Subjects are considered to have 'any increase' from baseline when the change from baseline in average daily prednisone dose is >0.

9.4.14. SLICC/ACR Scoring

- The SLICC/ACR Damage Index (Appendix 5) increases over time. Once a subject meets the criteria for positive scoring of an item, that item should always be marked as present, even if the subject subsequently recovered.
- In the event the SLICC/ACR Damage Index is scored inconsistently (a decrease relative to previous items has occurred) and the data are unable to be queried and/or

corrected, a worst observation carried forward (WOCF) approach will be used at the item level for the SLICC/ACR damage Index questions. These WOCF values will then be used to calculate the total score which will be the value summarized and displayed for reporting.

- Worsening is defined as an increase from baseline in SLICC/ACR Damage Index score [(post-baseline visit score – baseline score)] >0.
- No worsening is defined as no increase from baseline in SLICC/ACR Damage Index score [(post-baseline visit score – baseline score) = 0].

9.4.15. SELENA SLEDAI and SLEDAI-S2K scoring

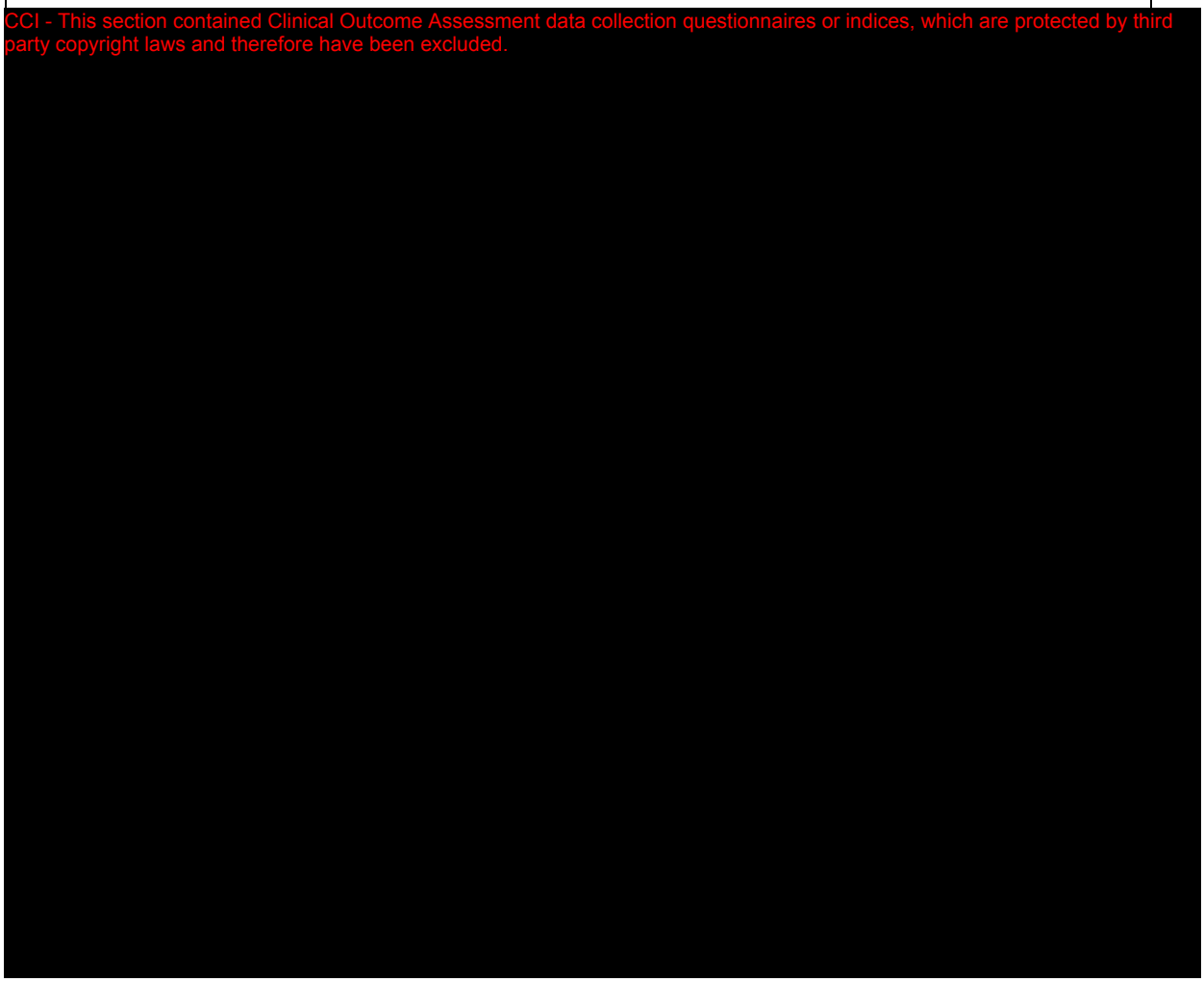
SELENA SLEDAI and SLEDAI-S2K Total and Organ System Domain Scores

SELENA SLEDAI assessments consist of 24 individual items in which signs and symptoms, laboratory tests, and physician's assessment for each of 9 organ systems are given a weighted score and summed if present (marked 'Yes') at the time of the visit or in the preceding 10 days. The maximum theoretical score is 105 (all 24 descriptors present simultaneously) with 0 indicating inactive disease (marked 'No'), but in practice few subjects have scores >45 [Buyon, 2005, Petri, 2005].

Organ system domain scores are the sum of the weights of items within the organ domain as defined below.

SELENA SLEDAI

CCI - This section contained Clinical Outcome Assessment data collection questionnaires or indices, which are protected by third party copyright laws and therefore have been excluded.



CCI - This section contained Clinical Outcome Assessment data collection questionnaires or indices, which are protected by third party copyright laws and therefore have been excluded.

[redacted]

to the Vascular organ system and the Constitutional organ system will be eliminated and its one component, fever, will be moved to the hematologic organ system.

SS-S2K- differs from SS [Gladman, 2002] in that it also reflects persistent disease activity in certain descriptors (listed below) whereas the SS only reflects new or recurrent disease activity.

Descriptors	Original SELENA SLEDAI	SELENA SLEDAI-S2K
Alopecia, Mucous Membrane Lesions, Rash	Items are only scored if they are new or recurrent.	Items are scored if they are present, regardless of history.
Proteinuria	Scored if it is a new onset or a recent increase of more than 0.5g/24hr.*	Scored if it was new, recurrent or persistent (i.e., any result >0.5g/24hr*).

*Urine protein:creatinine ratio of 0.5 mg/mg is considered equivalent to 0.5 g/24 hours for urine protein.

In this study, proteinuria is evaluated using spot-urine protein [represented as protein-creatinine ratio or PC Ratio (see Section 9.4.12) in the laboratory dataset] in lieu of the 24-hour urine collection.
For reporting, SS-S2K will be approximated by adjusting only the proteinuria descriptor of SELENA SLEDAI. If the laboratory PC ratio value exceeds 0.5 mg/mg the proteinuria item of the SLEDAI is given a positive score (4-point weight) creating an approximation to the SS-S2K assessment.

If the PC ratio value is missing then the value from the previous visit will be used. Specifically, the assessment for proteinuria indicated on the CRF will be used to calculate the SS score. For the SS-S2K score, the lab data will be assessed and if the PC ratio is >0.5 mg/mg for the visit then the proteinuria item will be scored as 4 points, otherwise it will receive a score of 0. If the PC ratio is missing, the proteinuria item from the previous SS-S2K will be carried forward.

In the eCRF, laboratory items on the SLEDAI may also be ticked 'unknown' to indicate the lab test was not available. The laboratory items are: urinary casts, hematuria, proteinuria, pyuria, low complement, increased DNA binding, thrombocytopenia, and leukopenia. In such instances where an item is ticked unknown, LOCF will be used

Missing items on the SS and SS-S2K will be managed per Section 9.3.4.

9.4.15.1. SLEDAI 4-point reduction

The SS 4-point reduction and SS-S2K 4-point reduction are defined below.

Endpoint	Criterion	Result
SS 4-point reduction	SS post-baseline – SS baseline ≤ - 4 (negative 4)	Y
	SS post-baseline – SS baseline > - 4 (negative 4)	N
	SS post-baseline = missing	See Section 9.3
	SS baseline = missing	Missing
SS-S2K 4-point reduction	SS-S2K post-baseline – SS-S2K baseline ≤ - 4 (negative 4)	Y
	SS-S2K post-baseline – SS-S2K baseline > - 4 (negative 4)	N
	SS-S2K post-baseline = missing	See Section 9.3
	SS-S2K baseline = missing	Missing

9.4.15.2. SLEDAI Organ System Improvement and Worsening

For both the SS and SS-S2K, each organ system (as defined in Section 9.4.15) will be evaluated for improvement and worsening as defined below:

Endpoint	Criterion	Result
Improvement	Organ score post-baseline – baseline organ score<0	Y
	Organ score post-baseline – baseline organ score ≥0	N
	Organ score post-baseline = missing	See Section 9.3
	Organ score baseline = missing	Missing
Worsening	Organ score post-baseline – baseline Organ score ≤ 0	N
	Organ score post-baseline – baseline Organ score > 0	Y
	Organ score post-baseline = missing	See Section 9.3
	Organ score baseline = missing	Missing

9.4.16. BILAG Scoring

BILAG
BILAG Score

• The British Isles Lupus Assessment Group (BILAG) score is an assessment of current lupus disease activity, as well as an indicator of historical disease activity in subjects with SLE. The Classic BILAG index was used in this study.
• Eight systems are given scores ranging from A to E where:
  • A = Active disease sufficient to require disease-modifying treatment (prednisolone >20mg or immunosuppressants)
  • B = Mild reversible problems requiring only symptomatic therapy (anti-malarials, NSAIDs, or prednisolone <20mg/day)
  • C = Stable, mild disease
  • D = Previous disease but currently inactive
  • E = Never active; no history
• If a subject meets the requirements for more than one letter score (A-E, with A being the highest), then the highest score met will be assigned for the organ system.
• Scoring of the BILAG is based on three publications including Hay, 1993; Isenberg, 2000, and a doctoral thesis written by Yee, 2008.
• The item numbers referred to below are CRF item numbers.

BILAG

BILAG Index Assessment (Scheduled Visits)

CCI - This section contained Clinical Outcome Assessment data collection questionnaires or indices, which are protected by third party copyright laws and therefore have been excluded.

BILAG System	Computational References Used	Source / Derivation / Comments *Variables named by BILAG item number (e.g. BILAG01 is item 1).*
General (Items 1-5)	Modified HGS BILAG Scoring using Hay	**First Assessment:** = 'A' if Pyrexia (BILAG01)>0 AND 2 of the other scores (BILAG02-BILAG05)>0 = 'B' if Pyrexia (BILAG01)>0 OR 2 of the other scores (BILAG02-BILAG05)>0 = 'C' if any of BILAG02-BILAG05 are >0 = 'E' if no involvement **Subsequent Assessments:** = 'A' if Pyrexia (BILAG01)>1 AND 2 of the other scores (BILAG02-BILAG05)>1 = 'B' if Pyrexia (BILAG01)>1 OR 2 of the other scores (BILAG02-BILAG05)>1 = 'C' if any of BILAG01-BILAG05 are >0 = 'D' if any previous value was in (A,B,C,D) = 'E' if at least one non-missing item score and no previous assessments were above E.
Mucocutaneous (Items 6-23)	Modified HGS BILAG Scoring using Hay for first assessments and Yee for subsequent assessments	**First Assessment:** = 'A' if any of BILAG06, BILAG08, BILAG13, or BILAG14 are >0 = 'B' if any one of BILAG16, BILAG07, BILAG12, BILAG09, BILAG10, BILAG17, or BILAG18 are >0 = 'C' if any one of BILAG19, BILAG20, BILAG21, BILAG22, BILAG23, BILAG11, or BILAG15 are >0 for 0-4 items or Yes for Yes/No items = 'E' if no involvement **Subsequent Assessments:** = 'A' if any of BILAG06, BILAG08, BILAG13, or BILAG14 are >1 = 'B' if any one of BILAG16, BILAG07, BILAG12, BILAG09, BILAG10, BILAG17, or BILAG18 are >1 = 'C' if (any one of BILAG19, BILAG20, BILAG21, BILAG22, BILAG23, BILAG11, or BILAG15 are >0 for 0-4 items or Yes for Yes/No items) or (any of BILAG06, BILAG08, BILAG13, BILAG14, BILAG16, BILAG07, BILAG12, BILAG09, BILAG10, BILAG17, or BILAG18 are =1) = 'D' if any previous value was in (A,B,C,D) = 'E' if at least one non-missing item score & no previous assessments were above E.
Neurological (Items 24-38)	Modified HGS BILAG Scoring using Yee	**All Assessments:** = 'A' if any of BILAG24, BILAG25, BILAG26, BILAG27, BILAG28, BILAG29, BILAG30, BILAG31, BILAG33, BILAG34 are in (3,4) = 'B' if (any of BILAG35, BILAG36, BILAG37, or BILAG32 are in (3, 4)) OR ((if any of BILAG24, BILAG25, BILAG26 are in (1,2)) = 'C' if BILAG38>0 OR (if any of BILAG27, BILAG28, BILAG29, BILAG30, BILAG31, BILAG33, BILAG34, BILAG35, BILAG36, BILAG37, or BILAG32 are in (1, 2)) = 'D' if any previous value was in (A,B,C,D) = 'E' if at least one non-missing item score & no previous assessments were above E.

BILAG System	Computational References Used	Source / Derivation / Comments *Variables named by BILAG item number (e.g. BILAG01 is item 1).*
Musculoskeletal (Items 39-47)	Modified HGS BILAG Scoring using Hay for first assessments and Yee for subsequent assessments	**First Assessment:** = 'A' if at least one of BILAG39 or BILAG40 is >0 = 'B' if at least one of BILAG41 or BILAG42 is >0 = 'C' if at least one of BILAG44, BILAG45, BILAG46, BILAG47, or BILAG43 is >0 for 0-4 items or Yes for Yes/No items = 'E' if no involvement **Subsequent Assessments:** = 'A' if at least one of BILAG39 or BILAG40 is >1 = 'B' if at least one of BILAG41 or BILAG42 is >1 = 'C' if (at least one of BILAG44, BILAG45, BILAG46, BILAG47, or BILAG43 is >0 for 0-4 items or Yes for Yes/No items) OR (one of BILAG39, BILAG40, BILAG41, BILAG42 is =1) = 'D' if any previous value was in (A, B, C, D) = 'E' if at least one non-missing item score and no previous assessments were above E.
Cardiovascular and Respiratory (Items 48-59)	Modified HGS BILAG Scoring using Hay for first assessments and Yee for subsequent assessments	**First Assessment:** = 'A' if 4 of the following are >0 for 0-4 items or Yes for Yes/No items: BILAG48, BILAG49, BILAG51, BILAG54, BILAG55, BILAG56, BILAG57, BILAG58, BILAG59 OR BILAG50>0 AND 2 of BILAG48, BILAG49, BILAG51, BILAG54, BILAG55, BILAG56, BILAG57, BILAG58, BILAG59 are >0 for 0-4 items or Yes for Yes/No items OR BILAG52>0 AND 2 of BILAG48, BILAG49, BILAG51, BILAG54, BILAG55, BILAG56, BILAG57, BILAG58, BILAG59 are >0 for 0-4 items or Yes for Yes/No items = 'B' if 2 of BILAG48, BILAG49, BILAG51, BILAG54, BILAG55, BILAG56, BILAG57, BILAG58, BILAG59, BILAG50, BILAG52 Are >0 for 0-4 items or Yes for Yes/No items = 'C' if any of BILAG53, BILAG50, BILAG52, BILAG48, BILAG49, BILAG51, BILAG54, BILAG55, BILAG56, BILAG57, BILAG58, or BILAG59 are >0 for 0-4 items or Yes for Yes/No items = 'E' if no involvement **Subsequent Assessments:** ='A' if 4 of the following are >1 for 0-4 items or Yes for Yes/No items: BILAG48, BILAG49, BILAG51, BILAG54, BILAG55, BILAG56, BILAG57, BILAG58, BILAG59 OR BILAG50>1 AND 2 of BILAG48, BILAG49, BILAG51, BILAG54, BILAG55, BILAG56, BILAG57, BILAG58, BILAG59 are >1 for 0-4 items or Yes for Yes/No items OR BILAG52>1 AND 2 of BILAG48, BILAG49, BILAG51, BILAG54, BILAG55, BILAG56, BILAG57, BILAG58, BILAG59 are >1 for 0-4 items or Yes for Yes/No items = 'B' if 2 of BILAG48, BILAG49, BILAG51, BILAG54, BILAG55, BILAG56, BILAG57, BILAG58, BILAG59, BILAG50, BILAG52

BILAG System	Computational References Used	Source / Derivation / Comments *Variables named by BILAG item number (e.g. BILAG01 is item 1).*
		Are >1 for 0-4 items or Yes for Yes/No items = 'C' if any of BILAG53, BILAG50, BILAG52, BILAG48, BILAG49, BILAG51, BILAG54, BILAG55, BILAG56, BILAG57, BILAG58, or BILAG59 are >0 for 0-4 items or Yes for Yes/No items = 'D' if any previous value was in (A,B,C,D) = 'E' if at least one non-missing item score and no previous assessments were above E.
Vasculitis (Items 60-67)	Modified HGS BILAG Scoring using Hay for first assessment and Yee for subsequent assessments	**First Assessment:** ='A' if at least one of BILAG60, BILAG61, or BILAG62 is >0 ='B' if at least one of BILAG66, BILAG65, or BILAG67 is >0 for 0-4 items or Yes for Yes/No item ='C' if at least one of BILAG63 or BILAG64 is >0 ='E' if no involvement **Subsequent Assessments:** ='A' if at least one of BILAG60, BILAG61, or BILAG62 is >1 ='B' if at least one of BILAG66, BILAG65, or BILAG67 is >1 for 0-4 items or Yes for Yes/No item ='C' if (at least one of BILAG63 or BILAG64 is >0) OR (at least one of BILAG60, BILAG61, BILAG62, BILAG65, BILAG66, or BILAG67 = 1 for 0-4 items or Yes for Yes/No item) ='D' if any previous value was in (A,B,C,D) ='E' if at least one non-missing item score and no previous assessments were above E.
Renal (Items 68-78) *All Assessments: For items 72 and 73, use item b recorded as urine protein-creatinine ratio (mg/mmol). Data requested per item a (24-hour urinary protein recorded in g/24 hours). was not collected per protocol*	Modified HGS BILAG Scoring using Isenberg and memo to file regarding items 72 and 73. Item 74 added to category D but no reference. Note: Typo in Isenberg (2000) paper where A3c creatinine clearance >50 ml/min and last time was >50/ml, in Yee this is <50 ml/min and last time >50 ml/min	**All Assessments:** For items 72 and 73, use item b recorded as urine protein-creatinine ratio (mg/mmol). Data requested per item a (24-hour urinary protein recorded in g/24 hours). was not collected per protocol **Define Proteinuria as:** **(**BILAG71 has a change from previous BILAG71 record >=2) OR (previous BILAG72B<0.2*100 mg/mmol and current BILAG72B>1*100 mg/mmol) OR (BILAG72B>1*100 mg/mmol and percentage change from previous BILAG72B value >=100%) or BILAG73B=1/yes and corresponding SLE relationship box has been checked for the criterion met **Define Deteriorating Renal Function as follows:** (BILAG75*88.4>130 mg/dl and percentage change from previous BILAG75>130%) **OR** (BILAG76<67% of previous BILAG76 value) **OR** (BILAG76<50 ml/min and previous BILAG76 value was >50 ml/min) and corresponding SLE relationship box has been checked for the criterion met ='A' if (two or more of the following occur: Proteinuria, BILAG70=1/yes, Deteriorating renal function, BILAG77=1/yes, BILAG78=1/yes) AND at least one of the two or more is Proteinuria, BILAG77, or BILAG78, and the corresponding SLE relationship has been checked for the criterion met. ='B' if (one or more of the following occur: Proteinuria, BILAG70=1/yes, Deteriorating renal function, BILAG77=1/yes, BILAG78=1/yes)

BILAG System	Computational References Used	Source / Derivation / Comments *Variables named by BILAG item number (e.g. BILAG01 is item 1).*
		OR (BILAG71>=2 and has risen from previous assessment by 1 or more) OR (BILAG72B has risen by >1*100 mg/mmol from previous BILAG72B record and increased by >50% but <100%) or (BILAG75*88.4>130 mg/dl and percentage change from previous record was >115%) and corresponding SLE relationship box has been checked for the criterion met. ='C' if (BILAG72B>0.25*100 mg/mmol) or (BILAG71>=1) OR (BILAG68>140 mm/Hg and change from previous BILAG68 >=30 mm/Hg and BILAG69>90 mm/Hg and change from previous BILAG69>=15 mm/Hg) ='D' if any previous value was in ('A', 'B', 'C', 'D') or BILAG74=1/yes and SLE related is checked ='E' if no involvement
Hematology (Items 79-86)	Modified HGS BILAG Scoring (all the same scoring algorithm), plus additional criterion for D from "See email from MChevrier dated 01Sep05 re neutrophils."	**All Assessments** (Note: 10^9/L=GI/L)**:** ='A' if BILAG80<1.0x10^9/L or BILAG83<25 x10^9/L or BILAG79<8 g/dL and the corresponding SLE related box is checked ='B' if BILAG80<2.5x10^9/L or BILAG83<100 x10^9/L or BILAG79<11 g/dL or (BILAG84=YES and BILAG85=YES) and the corresponding SLE related box is checked (there is no checkbox for BILAG85) ='C' BILAG80<4.0x10^9/L or BILAG83<150 x10^9/L or BILAG82<1.5x10^9/L or (BILAG85=YES and BILAG84^=YES) or BILAG86=YES and the corresponding SLE related box is checked (no checkbox for BILAG85 or BILAG86) ='D' if any previous value was in (A,B,C,D) or (.z<BILAG81 <0.5 x10^9/L and SLE related checkbox is checked) ='E' if at least one non-missing item score and no previous assessments were above E. NB: COMB' S Test Only done with evidence of hemolysis. So if #84 is yes, is evidence of hemolysis #84 no, #85 yes query

9.4.16.1. BILAG Improvement

BILAG improvement is limited to improvement from an A or B score, therefore within each organ domain, only subjects with an A or B score at baseline will be included. Subjects who have an A at baseline and change to a B, C, or D will be considered to have improvement. Similarly, subjects with a B at baseline who change to a C or D will be considered to have improvement.

9.4.16.2. BILAG Worsening

BILAG worsening is limited to worsening to an A or B score. Subjects who have a B at baseline and change to an A will be considered to have worsening. Similarly, subjects who have a C, D, or E at baseline and change to an A or B will be considered to have worsening. Subjects with an A at baseline are excluded since they cannot worsen.

9.4.16.3. BILAG No New 1A/2B

This responder endpoint (responder, non-responder) is a component of the SRI endpoint. A new A score is defined as any organ domain with an A score at a post-baseline visit that was not an A score at baseline. Similarly, a new B score is any organ domain with a B score post-baseline that was not an A or B at baseline (i.e., constitutes a worsening in the domain). To be a responder for this endpoint a subject cannot have any new A scores and no more than 1 new B score (i.e., not 2 new B scores).

9.4.16.4. BILAG No 1A/2B

This responder endpoint (responder, non-responder) is a component of the EMA SRI-S2K endpoint. To be a responder for this endpoint a subject cannot have any A scores and no more than 1 B score. This differs from the BILAG No New 1A/2B in that any subject with an A or more than one B score, even if they existed at baseline and are not 'new', will be considered a non-responder.

9.4.17. PGA Scoring

The PGA is collected on a 10cm visual analogue scale (VAS).

The standard scoring range for the PGA is 0 to 3, therefore the score will be rescaled for standard reporting by multiplying the collected score on the centimeter scale by 3/10.This rescaling will be within the SDTM datasets.

9.4.17.1. PGA No Worsening

PGA No Worsening is defined as (Post-baseline PGA – Baseline PGA) <0.3 using the re-scaled score (0-3 scale).

9.4.17.2. PGA Improvement

PGA Improvement is defined as (Post-baseline PGA – Baseline PGA) ≤ -0.3 using the re-scaled score (0-3 scale).

9.4.18. SRI-S2K

The primary endpoint is the systemic lupus erythematosus responder index (SRI) response rate with the modified SLEDAI-2K (S2K) scoring for proteinuria at Week 52. This S2K rule scores proteinuria as 4 points anytime the value is >0.5 g/24hr. This endpoint will be referred to as the SRI-S2K for reporting and is defined as:

- ≥4-point reduction from baseline in SELENA SLEDAI score using the SLEDAI-2K proteinuria scoring [SS-S2K 4pt],

AND

- No worsening (increase of <0.3 points from baseline) in Physician's Global Assessment (PGA) [PGA No Worsening],

AND

- No new British Isles Lupus Assessment Group of SLE Clinics (BILAG) A organ domain score or 2 new BILAG B organ domain scores compared with baseline at the time of assessment (i.e., at Week 52 visit) [BILAG No new 1A/2B].

The percentage of subjects achieving an SRI-S2K response at Week 52 will be presented for belimumab and placebo.

9.4.19. SRI using SELENA SLEDAI proteinuria scoring

The SRI will also be derived as described in Section 9.4.18 requiring a 4-point reduction using the SELENA SLEDAI proteinuria scoring as was originally recorded on the SELENA SLEDAI disease activity scale.

9.4.20. European Medicines Agency (EMA) modified SRI response at week 52 (EMA SRI-S2K)

The EMA modified SRI response rate with the modified SLEDAI-2K (S2K) scoring for proteinuria (EMA SRI-S2K) at Week 52 is defined as:

- ≥4-point reduction from baseline in SELENA SLEDAI - S2K score,

AND

- No worsening (increase of <0.30 points from baseline) in PGA,

AND

- No BILAG A organ domain scores or 2 BILAG B organ domain scores at the time of assessment (i.e., at Week 52).

This endpoint differs from the SRI-S2K in the BILAG component. The SRI-S2K requires no new BILAG 1A/2B scores whereas the EMA SRI-S2K requires no BILAG A and no more than one BILAG B score (No BILAG 1A/2B).

9.4.21. Date of SRI-S2K Response

SRI-S2K response is derived from three assessments (SELENA SLEDAI - S2K, PGA and BILAG) which were scheduled to be assessed on the same calendar date. At any visit, if the dates for the three assessments differ, then the date of SRI-S2K response will be determined from the observed dates (rather than any imputed dates). If there are multiple observed dates, the latest of these dates will be taken as the date of SRI-S2K response. Please see Section 6.1 for LOCF date details.

9.4.22. SRI-S2K Durable Response from Week 44 through Week 52

A durable SRI-S2K response is defined as a sustained SRI-S2K response from Week 44 through Week 52. Specifically, a subject has a durable response if the subject is an SRI-S2K responder at Weeks 44, 48, and 52.

9.4.23. SRI-S2K Duration of Longest Response

The duration of longest SRI-S2K response among subjects with at least 1 SRI-S2K response is defined as the entire duration of a response that first occurs at or before Week 52 to the last visit in which a subject responds consecutively. It will be calculated by: 1 + number of days from the first visit to the last visit in which a subject responds consecutively.

Last consecutive response date – first consecutive response date +1

For subjects without an SRI-S2K response, the duration will be set to zero (0).

For a SRI-S2K response starting on Week 52, the duration will be set to one (1) day.

9.4.24. SRI-S2K Duration of Week 52 Response

For subjects with a SRI-S2K response at Week 52, the duration is defined as:

[Date of Week 52 Assessment – Date of the first post-baseline visit where the subject is a responder and remains a responder at all following visits up to Week 52] + 1

For subjects without any response at Week 52, the duration of response will be set to zero (0) days. If a subject misses a visit between two scheduled visits at which the subject has a response, the subject will be defined as having a response at the missing visit.

9.4.25. SLE Flare Index (SFI) Scoring

SLE Flare Index

CRF Information

SLE Flare Index

Date of Assessment: ddMMMyyyy

Has the subject experienced a flare since the last SLE Flare assessment?

☐ 1 Yes ☐ 2 No

If Yes, specify details below:

Date of first flare since last flare assessment: ddMMMyyyy

☐ 1 **Mild or Moderate Flare**	☐ 1 **Severe Flare**
☐ 1 Change in SELENA SLEDAI instrument score of 3 points or more (but not to more than 12)	☐ 1 Change in SELENA SLEDAI instrument score to greater than 12
☐ 1 New/worse: Discoid, photosensitive, profundus, cutaneous vasculitis, bullous lupus Nasopharyngeal ulcers Pleuritis Pericarditis Arthritis Fever (SLE)	☐ 1 New/worse: CNS-SLE Vasculitis Nephritis Myostis Plt < 60,000 Hemolytic anemia: Hb <70 g/L or decrease in Hb >30 g/L
☐ 1 Increase in prednisone, but not to > 0.5mg/kg/day	**Requiring:** double prednisone, or prednisone increase to >0.5 mg/kg/day, or hospitalization
☐ 1 Added NSAID or hydroxychloroquine for SLE activity	☐ 1 Increase in prednisone to > 0.5mg/kg/day
☐ 1 ≥ 1.0 Increase in PGA score, but not to more than 2.5	☐ 1 New cyclophosphamide, azathioprine, methotrexate, or mycophenolate for SLE activity
	☐ 1 Hospitalization for SLE activity
	☐ 1 Increase in PGA score to >2.5

Derivations

- SFI reports the first mild /moderate or severe flare occurrence since the last visit assessment.
- The SLEDAI criteria will be assessed programmatically to determine if the SELENA SLEDAI criteria for a flare has been met and used for the assessment of flare, irrespective of what was recorded on the SFI form.
- Although there are boxes on the form for the investigator to classify the most recent flare to mild/moderate or severe, the classification will be re-derived from the subcategory scores.
- Flares originally marked severe will be downgraded to “Not Severe” if the only reason marked is a change in SELENA SLEDAI score to >12.
    - In this case, if any of the mild/moderate reasons are checked or if the modified SELENA SLEDAI score has a change from previous visit of at least 3, then the flare will be considered Mild/Moderate; otherwise the flare will not be counted.

- Flares that are marked mild/moderate where the only reason checked is SELENA SLEDAI increase of at least 3 points but not more than 12 points will be re-derived using the modified SELENA SLEDAI score.
    - If it's found that the change is not actually ≥3, and no other reasons are checked, then the flare will not be counted.

9.4.26. Time to First SFI Flare, Censoring and Disposition Rules

The rules described in this section apply to SLE Flares and SLE Severe Flares.

Up to 52 Weeks

Time to first modified severe SLE flare is defined as the number of days from first exposure until the subject meets an event (event date – first exposure date + 1). The disposition of subjects is defined as follows in Table 5.

Only post-baseline flares are included in the analysis, therefore flares (not subjects) occurring on the treatment start date should be removed from the analysis set prior to determining the first flare. In the rare situation that a treatment failure occurs on the treatment start date, the treatment failure will be counted as a flare as the assessment of treatment failure is performed post-dose.

For the Double-Blind phase, time to first flare over 52 weeks is calculated as:

Time to first flare (days) = Date of first flare – treatment start date +1.

From Week 24 to Week 52

For the time to first flare from Week 24 to Week 52, flares (not subjects) occurring on the Week 24 visit date should be removed from the analysis set prior to determining the first flare. Time to first flare from Week 24 to Week 52 is calculated as:

Time to first flare (days) = Date of first flare – Week 24 visit date + 1.

Table 5 Subject Disposition Rules for SLE Flares (Double-Blind Phase)

Subject Disposition	Event Met	Event Date
Subject has a flare or is a treatment failure, whichever occurs first		
Subject has a flare [1]	Yes	Date of first flare
Subject is a treatment failure [1]	Yes	Treatment failure date
Subject does not have a flare and is not a treatment failure		
Subject withdraws	No	Censored at last flare assessment date
Subject dies	No	Censored at date of death
Subject completes	No	Censored at last flare assessment date

[1] If a subject has a flare and is a treatment failure then the event date is the earliest of the first flare date and the date of treatment failure.

9.4.27. Severe SFI flare rate per subject-year over 52 weeks

Within a treatment group, the unadjusted severe flare rate per subject-year from Day 1 through Week 52 (not including baseline flares) will be calculated as:

$$\text{Unadjusted Rate per subject-year} = \frac{\text{total number of severe flares}}{\text{total number of years subjects are on study}}$$

The follow-up time (years) for each subject will be calculated as (date of last available study visit (or date of first treatment failure) – treatment start date + 1) / 365.25.

If a subject is a treatment failure, the subject will be considered having a severe SFI flare at the time of treatment failure. If the subject dies or withdraws from the study without experiencing a severe flare or treatment failure, the subject will be considered as not having a severe flare. If a subject is a treatment failure prior to Day 29, the subject will be excluded from the analysis. Similarly, subjects whose last visit is the Day 1 or Day 15 visit will be excluded from the analysis. No flares will be assigned for missing visits before the exit/treatment failure date.

9.4.28. Renal Flare

The protocol definition of renal flare [Alarcón-Segovia, 2003] consists of the development of one or more of the following 3 factors: increased proteinuria, impaired renal function, and hematuria. Since the protocol was written, there is new evidence indicating that the hematuria criterion introduces more variability into the definition making it harder to identify a true renal flare. Therefore, the protocol definition has been modified to remove the hematuria criterion. Renal flare is defined as follows:

- **Increased Proteinuria (using spot urine)**

A reproducible increase in 24-hour urine protein levels (as measured in uPCR) to:

>1g if the baseline value was <0.2g

OR

>2g if the baseline value was between 0.2g and 1g

OR

More than twice the value at baseline if the baseline value was >1g

- **Impaired Renal Function**

A reproducible decrease in GFR of >20%, accompanied by proteinuria (>1), and/or cellular (RBC and WBC) casts.

The protocol definition (protocol Section 8.5.4) expresses proteinuria in mg/24-hour. The definition above converts mg/24-hours to g/24-hours by dividing by 1000.

Proteinuria is assessed using the urine protein:creatinine ratio (uPCR) as a 1:1 equivalent for urine protein in g/24-hours. For example, 1000 mg/24-hours = 1 g/24-hours.

GFR is estimated by the Cockcroft-Gault equation for creatinine clearance in mL/min.

Cockcroft-Gault Equation [Cockcroft, 1976]

$$\text{Clcr (mL/min)} = \frac{(140 - \text{age (yrs)}) \times \text{weight (kg)}}{72 \times \text{serum creatinine (mg/dL)}} \quad \text{x 0.85 if female}$$

"Reproducible" requires the criterion to be met at two consecutive visits, including unscheduled visits and the 8 week FU visit.

Table 6 identifies the lab parameters to be used to evaluate the criteria in the renal flare definition above.

Table 6 Lab Parameters used in Renal Flare Definition

Criterion	Parameter	SI Dataset LBTESTCD
Proteinuria	Urine Protein:Creatinine Ratio (mg/mg)	PRTCRT_URQ
GFR	Creatinine Clearance estimated by the Cockcroft-Gault equation (mL/min)	CRTCE_PLR
RBC cellular casts	RBC cellular casts	RBCSTM_URQ
WBC cellular casts	WBC cellular casts	WBCSTM_URQ

9.4.29. Time to First Renal Flare, Censoring and Disposition Rules

Time to first renal flare is defined as the number of days from first exposure until the subject meets an event (event date – first exposure date + 1). The disposition of subjects is defined as follows in Table 7.

Renal flares are defined as reproducible results comparative to Baseline values, so can only occur post-Baseline.

Time to first renal flare over 52 weeks is calculated as:

Time to first flare (days) = Date of first flare – treatment start date +1

NB: Patients who withdraw or are lost to follow up without a renal flare assessment will be censored at zero days. Patients who withdraw or are lost to follow up with only one renal flare assessment (which therefore cannot be reproduced) will be censored at their

flare assessment date. It is possible for a renal flare to occur at the Week 52 visit as it could be reproduced at an unscheduled visit or the 8 week follow up visit.

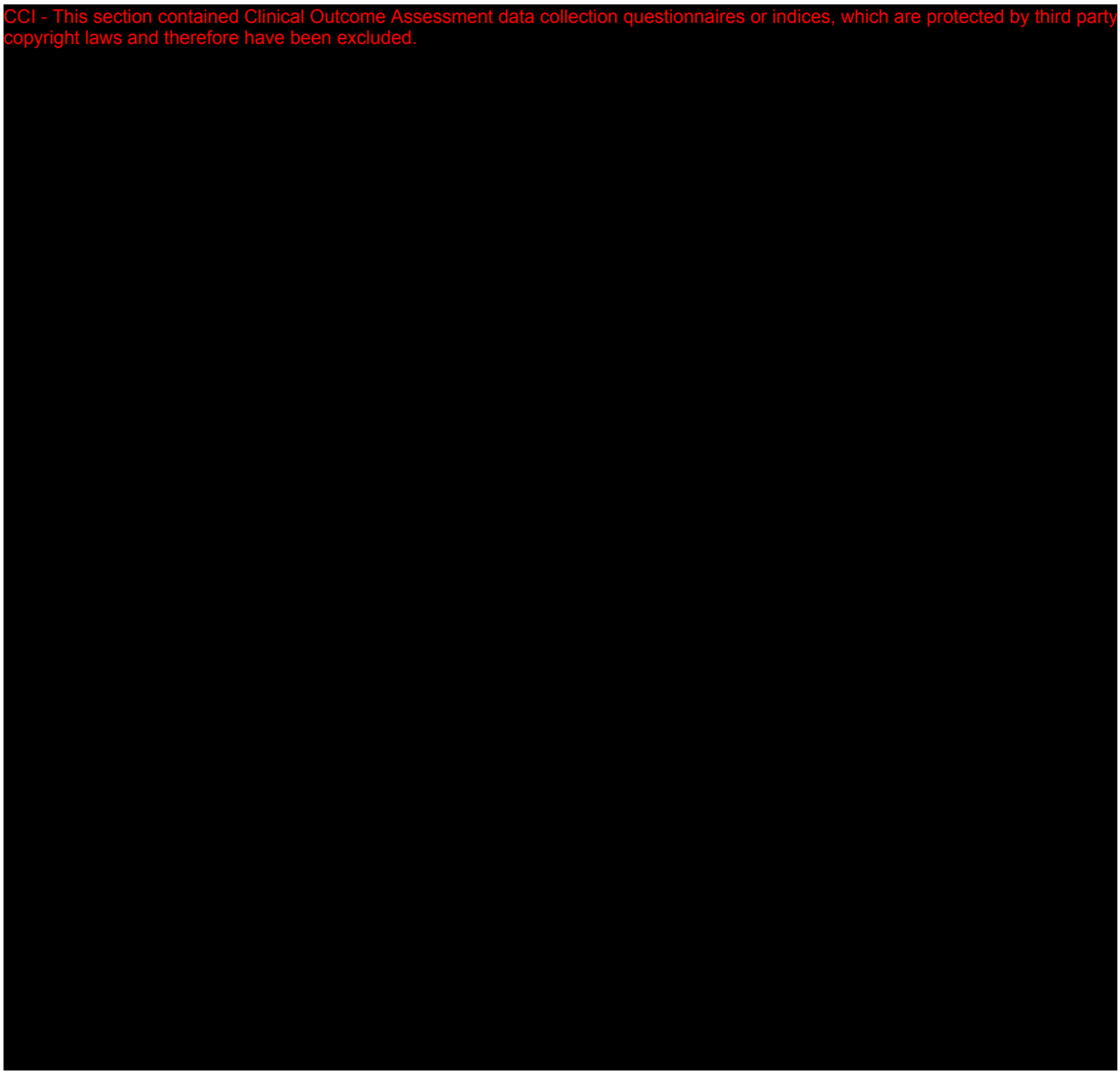
Table 7 Subject Disposition Rules for Renal Flares (Double-Blind Phase)

Subject Disposition	Event Met	Event Date
Subject has a flare		
Subject has a renal flare	Yes	Date of first renal flare which has been reproduced by a later assessment
Subject does not have a flare		
Subject withdraws	No	Censored at last flare assessment date
Subject dies	No	Censored at last flare assessment date
Subject is a treatment failure	No	Censored at last flare assessment date prior to treatment failure date
Subject completes	No	Censored at last flare assessment date

9.4.30. FACIT Fatigue Scale Scoring and MCID indicator

CCI - This section contained Clinical Outcome Assessment data collection questionnaires or indices, which are protected by third party copyright laws and therefore have been excluded.

CCI - This section contained Clinical Outcome Assessment data collection questionnaires or indices, which are protected by third party copyright laws and therefore have been excluded.



MCID ≥4

An additional indicator variable will be created by visit to identify subjects who have met or exceeded the Minimum Clinically Important Difference of greater than or equal to 4-point increase from baseline in FACIT-Fatigue score. Missing data will be imputed using the DO/TF=NR responder rule.

9.4.31. Date of Last Double-Blind Treatment Phase

For subjects continuing in the open-label phase, the last date in the double-blind treatment phase will be the **first dose date of Open Label medication**. For subjects not continuing in the open-label phase, the last **date in the double-blind** treatment phase will be the latest of the 8-week follow-up visit date, the early termination date or the Week 52/Exit visit date.

9.4.32. Date of Last Double-Blind Treatment Period

The end of the double-blind treatment period is defined as follows:

For subjects continuing in the open-label phase the end of the double-blind treatment period will be in a hierarchical fashion:

- The earliest of the Week 52 visit date and first dose date of Open Label medication.
- The earliest of the death date or early termination date
- Subjects' last treatment date of Double Blind medication

For subjects not continuing in the open-label phase the end of the double-blind treatment **period** will be (rules will be applied in a hierarchical order):

- Week 52/Exit visit date
- Earliest of death date or early termination date
- subjects' last treatment date in the double-blind phase will be used.

NB: Exit visits for subjects, in the double-blind phase, that withdrew early will be slotted to the closest visit based on planned visit day. All other visits used the CRF visit and were not slotted based on planned visit day. The general rule for the primary (Double-blind) reporting, is that it will include all records where the assessment date is less than or equals to the Open label period start date (AP02SDT). For Pharmacokinetic Concentration and Laboratory datasets, include all records where assessment date/time is less than or equals to Open label period start date/time. For Adverse Events include all records where AE start date is less than or equal to AP02SDT. For Concomitant Medications include all records where imputed start date is less than or equals to AP02SDT. Note that Adverse Events and Concomitant Medications on the first Open label treatment start date will be included for both the primary (Double-blind) and final (Open-label) reporting.

9.4.33. Extent of Exposure

Only complete dates will be used when calculating duration of exposure. First and last infusion dates will be used, regardless of any missed doses.

Duration of exposure in days for each subject will be calculated in the double-blind phase as:

Last infusion date in double-blind - first infusion date in double-blind + 28.

9.4.34. Treatment-Emergent AEs and AE Duration

Only treatment-emergent AEs will be summarized, unless otherwise stated. A treatment-emergent AE is an adverse event that emerges on or after the first treatment dose, having

been absent pre-treatment, or that worsens relative to the pre-treatment state. AEs with missing start and/or stop dates will be assumed to be treatment-emergent.

The duration of the AE will be calculated as follows:

Duration of AE (days) = Date of AE resolution – AE start date + 1.

If the AE is ongoing the duration will be left blank and no imputation will be done.

9.4.35. Windows for Assessment of Post-injection/infusion sensitivity reactions (PISR) and hypersensitivity reactions (HSR)

The windows for assessment of PISR/HSR, as described in the AESI definitions in Section 12.4, are as follows:

Date of Infusion		
Day 1	Day 2	Day 3
SMQ narrow, broad and algorithmic searches (On day of injection/infusion or within 3 days of injection/infusion)		

9.4.36. Laboratory Assessments

If a laboratory value which is expected to have a numeric value for summary purposes, has a non-detectable level reported in the database, where the numeric value is missing, but typically a character value starting with '<x' or '>x' (or indicated as less than x or greater than x in the comment field) is present, the number of decimal places in the observed values will be used to determine how much to add or subtract in order to impute the corresponding numeric value.

- Example 1: 2 Decimal Places = '< x' becomes x – 0.01
- Example 2: 1 Decimal Place = '> x' becomes x + 0.1
- Example 3: 0 Decimal Places = '< x' becomes x – 1

9.4.36.1. Clinical Laboratory Toxicity Grades

Toxicity grades are assigned by the central laboratory according to the Adverse Event Grading Tables in Appendix 11. Four analytes (electrolytes: calcium, potassium, and sodium; chemistry: glucose) have toxicities in both the high (hyper) and low (hypo) directions. Both directions will be presented in toxicity displays by the name of the toxicity. For example, calcium will have two toxicity sections, one for hypocalcaemia and one for hypercalcaemia. The hyper/hypo direction was set based on the following:

If laboratory value (AVAL) > (normal range low [ANRLO] + normal range high [ANRHI]) /2, then set as Hyper (toxicity grade=original toxicity grade);

If laboratory value(AVAL) < (normal range low [ANRLO] + normal range high [ANRHI]) /2, then set as Hypo (toxicity grade =original toxicity grade).

9.4.36.2. Laboratory Toxicity Worsening of at Least 2 Grades

Laboratory toxicity worsening of at least 2 grades will be reported for all possible shift categories even if no observed shifts occur in some of the categories, as follows:

Grade 0 to 2
Grade 0 to 3
Grade 0 to 4
Grade 1 to 3
Grade 1 to 4
Grade 2 to 4

For the four analytes (electrolytes: calcium, potassium, and sodium; chemistry: glucose) that have toxicity grade worsening in both high and low directions, ≥ 2 grade worsening will also include shifts from Hypo to Hyper and Hyper to Hypo with Grade 0 being considered a grade. Only observed shifts will be presented for shifts from Hyper to Hypo and Hypo to Hyper. An example is presented below.

Grade 0 to 2
Grade 0 to 3
Grade 0 to 4
Grade 1 to 3
Grade 1 to 4
Grade 2 to 4
Hyper 1 to Hypo 1
Hyper 1 to Hypo 2

9.4.36.3. C3 and C4 Unit Conversions

The biomarkers C3 and C4 need to be converted from G/L to mg/dL for reporting to be consistent with the Benlysta reporting standard. The formula for the unit conversion is:

$$\text{mg/dL} = \text{g/L} \times 100$$

10. STUDY POPULATION

The mITT population will be used to summarize the study population data and data will be presented by treatment and for all subjects combined, unless otherwise specified.

10.1. Disposition of Subjects

The number and percentage of subjects randomized by Country and Site ID will be summarized overall and by treatment group for the randomized population.

Using the Randomized population, the number of subjects in each population (Randomized, Safety, mITT, Per Protocol, Completers, As-Treated and PK) will be summarized overall and by treatment group

If there are any subjects who are randomized but do not receive any study drug, they will be included in the Randomized population, but not the Safety or mITT populations.

For the mITT population, the subject's completion status will be assessed to evaluate percentages of dropouts by treatment group as well as the reasons for dropout. The number and percentage of subjects who completed through Week 52 and who withdrew, including their Primary reason for withdrawal, will be displayed by treatment group and overall for the double-blind phase. Additionally, the cumulative number and percentage of subjects who withdrew by study visit will be displayed by treatment group and overall. These summaries will be repeated for the safety population.

Additionally, a Kaplan Meier (KM) plot of time to withdrawal will be generated to evaluate the pattern of dropouts over time. Subjects who complete through Week 52 will be censored at 52 weeks. To aid stage 3 review of the Kaplan Meier plot, the SAS LIFETEST output will be provided. This will not be a reported output.

A listing of subject disposition will be provided showing their completion status and whether they are included in each population. A listing of subjects who withdrew from the study, including reason for and study day of withdrawal will also be provided.

A listing of subjects that deviated from the inclusion or exclusion criteria will be provided.

10.2. Protocol Deviations

Please refer to the Protocol Deviation Management Plan (PDMP): Dated: 6Jun2017 (Version 2.0) for full details describing important deviations and important deviations which result in exclusion from the PP population.

Important protocol deviations (including deviations related to study inclusion/exclusion criteria, conduct of the trial, patient management or patient assessment) will be summarized by treatment group.

A summary of protocol deviations leading to exclusion from the Per Protocol population will be displayed. The summary will include the number of subjects excluded, category of the reason for exclusion, and the reason for exclusion.

Important protocol deviations and deviations which result in exclusion from the PP population will be listed. (See Appendix 8 for details of important protocol deviations and deviations leading to exclusion from PP population).

A subject listing of the criteria which lead to exclusion, the category and coded term for the criteria will be produced. The display will include columns to indicate from which analysis population the subject was excluded.

Protocol deviations will be tracked by the study team throughout the conduct of the study in accordance with the PDMP. Data will be reviewed prior to database release to ensure all important deviations and deviations which may lead to exclusion from the analysis are captured and categorized on the eCRF and protocol deviations dataset. This dataset will be the basis for the summaries and listings of protocol deviations.

10.3. Demographic and Baseline Characteristics

Continuous parameters will be summarized using descriptive statistics (mean, standard deviation, median, 25th percentile, 75th percentile, minimum, and maximum). Categorical parameters will be summarized using counts and percentages.

In addition to the overall mITT population, the summaries of demographic characteristics, baseline characteristics and allowable SLE medications at baseline will be repeated for the overall safety population and the following subgroups (using the mITT population):

- Baseline SS-S2K score (≤9 vs. ≥10)
- Complement level (At least one C3/C4 Low vs. No C3/C4 Low)
- Region (US/Canada, other)
- Baseline C3/C4 levels & anti-dsDNA (At least one C3/C4 low and anti-dsDNA >=30 IU/mL vs. Not (At least one C3/C4 low and anti-dsDNA >=30 IU/mL)
- Age Group (<65, ≥ 65 years)

NB: Gender summaries will not be performed due to the low number of male subjects.

Demographic and baseline characteristics will be listed for all patients by treatment group. A listing of stratifications factors used in the randomization will also be listed (using the information recorded in the IVRS) for all patients by treatment group.

A listing of medical history data will be provided by treatment group.

The following Demographic and baseline disease activity indicators will be summarized:

PARAMETER	SUMMARY TYPE
Medical History	
Medical History by Body System and Term	Categorical: presented if current or past medical history is present
Stratification Factors at Screening and as Randomized	
SELENA SLEDAI (SS)	Categorical: <=9, >=10
Complement Level	Categorical: - At least one C3/C4 Low -No C3/C4 Low
Region	Categorical: US/Canada, Rest of World
Demographic	
Age (years)	Continuous Categorical: < 65 years >= 65 years <=45 years, >45 – <65 years >=65 – <75 years >=75 years 18 - 64 65 - 84 >= 85
Height (cm)	Continuous
Weight (kg)	Continuous
Body Mass Index (BMI) (kg/m^2)	Continuous
Region	Categorical: US/Canada vs. RoW
Sex	Categorical: Female, Male
Race	Categorical: American Indian or Alaska Native Asian Black or African American Native Hawaiian or Other Pacific Islander White Multiple
Hispanic or Latino origin	Categorical: Yes, No
Baseline Disease Activity	
ACR classification criteria by symptom	Categorical: items answered 'Yes' Categorical: total criteria met (integers from 4 to 11)
SELENA SLEDAI-S2K score	Continuous Categorical: <=9, 10-11, >=12
SELENA SLEDAI score	Continuous Categorical: <=9, 10-11, >=12

PARAMETER	**SUMMARY TYPE**
Shift Table of differences between SS-S2K and SS	Categorical: <=9, >=10
BILAG organ domain scores at baseline	Categorical: -at least 1A or 2B -at least 1A -at least 1B -no A or B *subjects may be included in more than one category
SLE Flare Index	Categorical: -at least one flare -at least one severe flare
PGA score	Continuous Categorical: 0-1, >1 – 2.5, >2.5
SLICC/ACR Damage Index score	Continuous Categorical: 0, 1, >1
Proteinuria	Continuous Categorical: <=0.5, >0.5-<1, 1-<2, >=2
SLE disease duration (years)	Continuous
SELENA SLEDAI by organ domain and item at baseline	Categorical: presented if Yes
BILAG grade by organ domain at baseline	Categorical: A, B, C, D, and E scores
Baseline immunoglobulin levels (g/L): IgG, IgA, and IgM	Continuous value Categorical: - below the lower limit of normal (LLN) (IgG < 6.94 g/L, IgA < 0.81 g/L, and IgM < 0.48 g/L) - above upper limit of normal (ULN) (IgG > 16.18 g/L, IgA > 4.63 g/L, and IgM > 2.71 g/L).

PARAMETER	SUMMARY TYPE
Baseline autoantibody levels	
anti-dsDNA [IU/mL]	Continuous Categorical: -positive (≥ 30 IU/mL) -negative (<30 IU/mL)
ANA [Titer]	Continuous Categorical: -positive (≥ 1:80 Titer) -negative (<1:80 Titer)
anti-cardiolipin (aCL)	Continuous Categorical: -positive (if any of the three isotypes IgG, IgA or IgM are above the upper limit of normal (ULN)) -negative (if at least one is non-missing and none of the isotypes above the upper limit of normal) -missing (all three isotypes are missing)
anti-dsDNA and/or ANA positive	Categorical: Yes, No
Beta-2-glycoprotein [U/mL] for isotypes IgG, IgA, and IgM	Continuous Categorical: Positive (>=21 U/mL), Negative
Lupus anticoagulant	Continuous
Complement Levels and BLyS at Baseline	
Baseline levels of complement: C3 [mg/dL]	Continuous Categorical: high (>180 mg/dL), normal (90 - 180 mg/dL), low (<90 mg/dL)
Baseline levels of complement: C4 [mg/dL]	Continuous Categorical: high (>40 mg/dL), normal (10 -40 mg/dL), low (<10 mg/dL)
Complement Level	Categorical: At least one C3/C4 Low; No C3/C4 Low
Complement Level and anti-dsDNA Level	Categorical: At least one C3/C4 Low and anti-dsDNA >=30 IU/mL; Not (At least one C3/C4 Low and anti-dsDNA >=30 IU/mL)
BLyS [ng/mL]	Continuous Categorical: below limit of quantification [LOQ] (<0.02048 ng/mL), above LOQ

PARAMETER	SUMMARY TYPE
B cells at Baseline	
CD19+ CD20+ CD20+/27+ memory CD20+/27+ memory (%CD19) CD20+/27– naïve CD20+/27– naïve (%CD19) CD20+/69+ activated CD20+/138+ plasmacytoid CD19+/27^{BRIGHT}/38^{BRIGHT} SLE subset CD19+/$CD24^{HIGH}$/$CD38^{HIGH}$ transitional CD20-/138+ plasma $CD27+^{BRIGHT}$/CD20– Short-lived plasma	Continuous
Columbia Suicide Severity Rating Scale (C-SSRS) at Baseline	
C-SSRS responses by behavior and ideation components for lifetime and over the last two months Note: C-SSRS summaries use Safety Population.	Categorical: -Ideation: categories 1-5 with corresponding text -Behavior: categories 6-10 with corresponding text
FACIT Fatigue	
Baseline FACIT-Fatigue scale score	Continuous Categorical: (<20, 20-<35, ≥35)
Allowable SLE Medication Usage at Baseline	
Allowable SLE medications by class and drug	Categorical: Steroids, Anti-malarials, Immunosuppressants, Aspirin, NSAIDs
Average daily prednisone dose (mg/day) at baseline	Continuous Categorical: (0, >0 - ≤7.5, >7.5)
Steroid, Anti-malarial and Immunosuppressant Use at Baseline	Categorical: -Steroid Only -Immunosuppressant Only -Anti-malarial Only -Steroid and Immunosuppressant Only -Steroid and Anti-malarial Only -Immunosuppressant and Anti-malarial Only -Steroid and Immunosuppressant and Anti-malarial

10.4. Concomitant Medications

Concomitant medications will be coded according to drug name as defined in the GSK Drug Dictionary, and classified according to the GSK-Drug ATC classification level 1 and ATC level 4. Concomitant medications are defined as medications that start on or before the first dose date of study treatment and end on or after the first dose date of study treatment, or medications that start on or after the first dose date of study treatment. Note that medications with partial or missing start and/or stop dates will be assumed to be concomitant unless there is evidence through comparison of partial dates to suggest otherwise, for example if the day is missing, then the month and year will be compared to the month and year of the first dose date of study treatment and if the month and year are the same or later, then the medication will be considered concomitant.

A summary of the number and percentage of subjects with concomitant medications by ATC level 1 term and ATC level 4 term will be displayed. A further summary of concomitant medications by ATC level 4 term and preferred term will be provided.

A listing of all concomitant medication data will be displayed by treatment and subject.

First protocol-prohibited medication and allowable medication that results in treatment failure designation (see Section 5.5 of the protocol) will be summarized by treatment group and overall. A similar summary will be produced with subjects appearing in multiple protocol-prohibited medication and allowable medication treatment failure categories.

A listing of protocol-prohibited medication and allowable medication that result in treatment failure designation will be displayed by treatment and subject.

Concomitant medications will be summarized for the mITT and safety populations.

10.5. Extent of Exposure

The extent of exposure to study treatment through Week 52 will be assessed by examining the duration of exposure to belimumab/placebo in days and the total number of infusions a subject receives. Duration of exposure in days will be calculated as:

Duration of exposure (days) = (Last infusion date – First infusion date + 28).

Only complete dates will be used when calculating duration of exposure. First and last infusion dates will be used, regardless of any missed doses.

The duration of exposure and the total number of infusions will be summarized using descriptive statistics for the double-blind phase. The total number of infusions will also be summarized using counts and percentages using the following categories: 1 – 5 doses, 6 – 10 doses, 11 – 14 doses and ≥15 doses.

Exposure data will be summarized for the mITT and safety populations.

Exposure data will be listed for all subjects by treatment group and subject. A listing of study agent administration will also be produced for all subjects by treatment group and subject.

10.6. Subjects for Whom the Treatment Blind was Broken

A listing will be produced of subjects for whom the treatment blind was broken during the study.

11. EFFICACY ANALYSES

The efficacy analyses will be performed for the mITT population as defined in Section 6 unless otherwise stated. The data will be presented by treatment group.

Treatment failures and handling of missing data, for all efficacy analyses, will be managed as described in Section 9.

For the analysis of steroid use, all steroid dosages are converted to a prednisone equivalent in milligrams (see Section 9.4.13 for further details); therefore, analyses refer to average daily prednisone dose instead of average daily steroid dose.

Where efficacy endpoints are expected to be Normally distributed, ANCOVA is planned to be performed. However, if the normality assumptions are violated, or there are influential outliers in the dataset, then other methods of analysis will be investigated. The first alternative to be investigated will be the non-parametric Wilcoxon Rank Sum Test. In the case that a non-parametric analysis is more appropriate, then the analysis will be adapted and relevant statistics for the new analysis will be presented in the output tables. If analyses are changed to be non-parametric due to the distribution of the data, then Figures presenting the data will switch from presenting means (± SE) to medians (± quartiles).

Listings will be provided by treatment group for the following data:

- SRI-S2K Results
- SELENA SLEDAI and SS-S2K Results
- PGA Results
- BILAG Results
- Daily Prednisone Dose
- SFI Flares
- SLICC/ACR Damage Index Results

11.1. Primary Efficacy Analysis

SRI-S2K

A logistic regression model will be used to estimate the odds of an SRI-S2K response for belimumab vs. placebo. The independent variables in the model will include

- treatment group,
- baseline SS-S2K score (≤ 9 vs. ≥ 10),
- baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and

- region (US/Canada vs. rest of the world).

NB: due to the change in the primary endpoint, the SS-S2K will be used as the independent variable in the model instead of the SELENA SLEDAI to be consistent with the SLEDAI component of the primary endpoint for the primary analysis. A secondary analysis will be performed using the original endpoint SELENA SLEDAI scoring (SRI) with SS (≤9 vs. ≥10) as the independent variable in the model (Section 11.2.1).

Additionally, for the primary analysis and key secondary analyses the covariate adjustment will not occur for any of these baseline randomization factors that has <10 responders or <10 non-responders in a stratification level [Peduzzi, 1996].

The table will display the number and percentage of subjects achieving a response by treatment group and standard error for the percentage, the treatment difference versus placebo, the odds ratio, and 95% confidence interval (CI) versus placebo and a p-value for the odds ratio.

A table showing the details from the logistic regression will also be presented; including the parameter estimate, standard error and p-value for the intercept and the parameter estimate, standard error, odds ratio and 95% CI and p-value for treatment group and each covariate in the model.

The percentage of SRI-S2K responders (± SE) at Week 52 will be displayed in a bar chart, by treatment group.

NB; In the event a patient has a baseline SS-S2K<4 they will be excluded from the analysis.

11.1.1. Supportive Summaries of the Primary Efficacy Endpoint

11.1.1.1. Components of Primary Endpoint

In support of the primary endpoint, the number and percentage of subjects meeting each of the three components of the primary endpoint will be presented at Week 52 with the SRI-S2K. If any subject has missing data that excludes them from the SRI-S2K analysis, they will also be excluded from each of the component analyses in this table.

Additionally, each component will also be summarized individually by visit, with the analysis presented at Week 52. The percentage of subjects meeting each component criteria (± SE) will be displayed by treatment group in a line plot over the double-blind treatment period.

Logistic regression models will be used to estimate the odds of a response in each component for belimumab vs. placebo at Week 52. The independent variables in the models will include treatment group, baseline SS-S2K score (≤9 vs. ≥ 10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world). Baseline PGA score will also be included as a covariate in the model for the PGA No Worsening component. Similarly, baseline BILAG organ domain involvement (at least 1A/2B vs. at most 1B) will also be included as a covariate in the model for the BILAG No new 1A/2B component.

The tables will display the number and percentage of subjects achieving a response in each component by treatment group, the SE for the percentage of responders, the treatment difference vs. Placebo and also the odds ratios and 95% CIs versus placebo and p-values for the odds ratios at Week 52.

11.1.1.2. Disposition of SRI-S2K Response

For the primary SRI-S2K endpoint, the disposition of factors contributing to response or non-response will be presented as the number and percentage of subjects falling into the following categories at Week 52 by treatment group:

- SRI-S2K Responder
- SRI-S2K Non- Responder
  - Dropout (without a treatment failure)
  - Treatment failure
  - SS-S2K <4-point reduction
  - SS-S2K ≥4-point reduction with the following:
    - PGA Worsening only
    - BILAG New 1A/2B only
    - Both PGA Worsening and BILAG new 1A/2B

A bar chart showing the percentage of subjects falling into each of the categories above at Week 52 in each treatment group will also be presented.

11.1.1.3. Additional SRI-S2K Analyses

The following additional supportive analyses of the SRI-S2K primary endpoint will be performed using logistic regression at Week 52.

Sensitivity Analysis	Dependent Variable	Independent Variables[1] in Logistic Regression Model
Unadjusted	SRI-S2K	Treatment
LOCF	SRI-S2K	Treatment, SS-S2K, Complement, Region
Per Protocol	SRI-S2K	Treatment, SS-S2K, Complement, Region
Completers	SRI-S2K	Treatment, SS-S2K, Complement, Region
SS-S2K 5-point reduction	SRI5-S2K	Treatment, SS-S2K, Complement, Region
SS-S2K 6-point reduction	SRI6-S2K	Treatment, SS-S2K, Complement, Region
SS-S2K 7-point reduction	SRI7-S2K	Treatment, SS-S2K, Complement, Region
SS-S2K 8-point reduction	SRI8-S2K	Treatment, SS-S2K, Complement, Region
[1] Treatment=Belimumab vs. Placebo, SS-S2K=baseline SS-S2K score (≤9 vs. ≥10),		

Sensitivity Analysis	Dependent Variable	Independent Variables[1] in Logistic Regression Model
Complement =baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low), and Region =US/Canada vs. rest of the world. NB; In the event a patient has a baseline SS-S2K<4, they will be excluded from the analysis. Additionally, patients who have baseline SS-S2K <5, <6, <7 or <8, they will be excluded from the SRI5-S2K - SRI8-S2K respectively		

The tables will display the number and percentage of subjects achieving a response in each component by treatment group, the SE for the percentage of responders, the treatment difference vs. Placebo and also the odds ratios and 95% CIs versus placebo and p-values for the odds ratios at Week 52.

11.1.1.4. SRI-S2K by subgroup

The primary efficacy analysis at Week 52 will also be repeated for each of the subgroups listed in Section 8.3.

- The logistic regression models will be used to estimate the odds of an SRI-S2K response for belimumab vs. placebo in each subgroup category. The model will be unadjusted.
- The tables will display the number and percentage (with standard error) of subjects achieving a response by each subgroup category and treatment group, the treatment difference versus placebo, the odds ratio and 95% CI versus placebo, a p-value for the odds ratio and a p-value for the treatment-by-subgroup interaction.
- In order to obtain the treatment by sub-group interaction term a model will be fitted which includes treatment group, subgroup and treatment-by subgroup term.
- The odds ratios (and 95% CI) will be presented graphically for each category within each subgroup.

11.1.1.5. SRI-S2K by visit

An analysis of the SRI-S2K will performed by visit. The analysis will be performed as described in Section 11.1.

The percentage of responders (± SE) will be presented in a line graph over the Double-Blind treatment period.

11.1.1.6. EMA Modified SRI-S2K Response

An analysis of the EMA Modified SRI-S2K Response will performed at Week 52. The analysis will be performed as described in Section 11.1.

11.2. Secondary Efficacy Analyses

11.2.1. SRI response rate with the SELENA SLEDAI scoring of proteinuria by visit (SRI)

An analysis of the SRI using the SELENA SLEDAI scoring for proteinuria will performed by visit and the percentage of responders (± SE) will be presented in a line graph over the Double-Blind treatment period. The analysis will be performed as described in Section 11.1 (using the Baseline SELENA-SLEDAI subgroups instead of the Baseline SS-S2K subgroups as a covariate in the model).

In addition, the SRI analysis at Week 52 will be repeated for the Baseline C3/C4 levels & anti-dsDNA subgroups, the Baseline SELENA-SLEDAI subgroups and the Baseline C3/C4 level subgroups.

11.2.2. Time to first severe SFI flare over 52 weeks

Analysis of severe SFI flare will be performed on the modified SELENA SLEDAI SLE flare index in which the modification excludes severe flares that were triggered only by an increase in SELENA SLEDAI score to >12 (since this may only represent a modest increase in disease activity). Only post-baseline severe flares will be considered in these analyses.

The time to the first severe SFI flare over 52 weeks will be compared between belimumab and placebo using a Cox proportional hazards model, adjusting for baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world). Ties will be handled using approximate likelihood (EFRON).

The rules for counting events and censoring are described in Section 9.4.26.

The table will display the number and percentage of subjects with a severe SFI flare over 52 weeks, the median, 25th, and 75th percentile of days to first severe flare, the hazard ratio (and 95% CI) versus placebo and the p-value from the Cox proportional hazards model. For subjects who experience a severe flare, the study day of the flare will be summarized and the table will display the median, 25th and 75th percentiles and minimum and maximum. A Kaplan Meier plot will be produced to show the probability of experiencing a severe SFI flare over 52 weeks.

In addition, the Cox proportional hazards analysis will be repeated for the Baseline C3/C4 levels & anti-dsDNA subgroup using the SS.

11.2.3. Percent of subjects with prednisone reduction by ≥25% from baseline to ≤7.5 mg/day during week 40 through week 52 (DO/TF=NR)

For this analysis, only subjects with a baseline average daily prednisone dose >7.5mg/day will be included in the analysis. Average daily prednisone dose during Weeks 40 through 52 is defined in Section 9.4.13.4. The percent of subjects whose average prednisone dose

has been reduced by ≥25% from baseline to ≤7.5 mg/day during Weeks 40 through 52 will be compared between belimumab and placebo using a logistic regression model. Independent variables in the model will include treatment group, baseline prednisone dose, baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world).

Any subject who withdraws from the study prior to the Week 52 visit or is deemed a treatment failure prior to the Week 52 visit will be considered a non-responder for this analysis (DO/TF=NR).

The table will display the number and percentage of subjects who are responders by treatment group, the SE for the percentage of responders, the treatment difference versus placebo, the odds ratio and 95% CI versus placebo, and a p-value for the odds ratio.

This analysis will be repeated for the Baseline C3/C4 levels & anti-dsDNA subgroups.

11.2.4. Estimand Definition

This study was designed prior to the publication of the draft ICH E9 Addendum. Table 8 below shows how the primary and major secondary analyses can be aligned with the expected terminology.

Table 8 Alignment with Estimand Definition Terminology

	SRI-S2K 4-point responder	**SRI 4-point responders**	**Time to first Severe Flare**	**Prednisone reduction responder**
Population	mITT	mITT	mITT	mITT taking >7.5mg prednisone at baseline
Variable	Composite of following 1. ≥4-point reduction from baseline in SELENA SLEDAI score with the modified SLEDAI-2K scoring for proteinuria 2. No worsening (increase of <0.30 points from baseline) in Physician's Global Assessment (PGA)	Composite of following 1. ≥4-point reduction from baseline in SELENA SLEDAI 2. No worsening (increase of <0.30 points from baseline) in Physician's Global Assessment (PGA) 3. No new	Composite of following 1. Subject experiences Severe Flare 2. Subject meets treatment failure definition	Composite of following 1. prednisone reduction by ≥25% from baseline to ≤7.5 mg/day during week 40 through week 52 2. Subject does not drop out before week 52 (±28 days) 3. Subject does not meet treatment failure definition

	SRI-S2K 4-point responder	SRI 4-point responders	Time to first Severe Flare	Prednisone reduction responder
	3. No new British Isles Lupus Assessment Group of SLE Clinics (BILAG) A organ domain score or 2 new BILAG B organ domain scores compared with baseline at the time of assessment (i.e., at Week 52). 4. Subject does not drop out before week 52 (± 28 days) 5. Subject does not meet treatment failure criteria	British Isles Lupus Assessment Group of SLE Clinics (BILAG) A organ domain score or 2 new BILAG B organ domain scores compared with baseline at the time of assessment (i.e., at Week 52). 4. Subject does not drop out before week 52 (± 28 days) 5. Subject does not meet treatment failure criteria		
Intercurrent event	Subject drops out or meets treatment failure definition; incorporated in the composite endpoint as a non-responder	Subject drops out or meets treatment failure definition; incorporated in the composite endpoint as a non-responder	Subject meets treatment failure definition; incorporated in the composite endpoint Subject drops out: subject is censored at last available visit Subject dies: Subject is censored at the date of death	Subject drops out or meets treatment failure definition; incorporated in the composite endpoint as a non-responder
Measure	Adjusted odds ratio; see Section 11.1	Adjusted odds ratio; see Section 11.2.1	Adjusted hazard ratio; see Section 11.2.2	Adjusted odds ratio; see Section 11.2.3

11.3. Other Efficacy Analyses

Formal statistical analysis of the other efficacy endpoints will be performed at week 52 only. Other study visits will be limited to descriptive summaries.

11.3.1. Disease Activity

11.3.1.1. Durable SRI-S2K Response from Week 44 - Week 52

The percentage of subjects achieving a durable SRI-S2K response from Week 44 through Week 52 will be presented for belimumab and placebo. A logistic regression model will be used to estimate the odds of an SRI-S2K response for belimumab vs. placebo. The independent variables in the model will include treatment group, baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world).

The table will display the number and percentage of subjects achieving a response by treatment group, the SE of the percentage of responders, the treatment difference versus placebo, the odds ratio and 95% CI versus placebo and a p-value for the odds ratio.

11.3.1.2. Time to first SRI-S2K response that is maintained through week 52

The time to the first SRI-S2K response that is maintained through Week 52 will be compared between belimumab and placebo using a Cox proportional hazards model, adjusting for baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world). Ties will be handled using approximate likelihood (EFRON).

The table will display the number and percentage of subjects with a SRI-S2K response that is maintained through Week 52, the median, 25th, and 75th percentile of days to first SRI response, the hazard ratio (and 95% CI) versus placebo and the p-value from the Cox proportional hazards model. A Kaplan Meier plot for time to first SRI response that is maintained through Week 52 will also be produced.

11.3.1.3. Duration of longest SRI-S2K response

The duration of longest SRI-S2K response among subjects with at least one SRI-S2K response will be compared between belimumab and placebo using an analysis of covariance (ANCOVA) model. The covariates in the model will include treatment group, baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world). The table will display the number of responders, the mean, standard deviation and SE of the mean, median, minimum, maximum, 25^{th} and 75^{th} percentiles, least squares (LS) mean and associated standard error, treatment difference versus placebo, 95% CI for the treatment difference, and p-value from the ANCOVA model.

In addition, another table will display the number and percentage of subjects with a SRI-S2K response, the median, 25th, and 75th percentile, the hazard ratio (and 95% CI) versus placebo and the p-value from the Cox proportional hazards model.

11.3.1.4. Duration of week 52 SRI-S2K response by month

The duration of the primary response evident at Week 52 will be grouped into monthly (where a month is defined as 30 days) intervals.

The table will display the number and percentage of subjects falling into each duration of response category (≥1 day, ≥1 month, ≥2 months, etc.). Subjects will be presented in multiple categories, e.g. if their duration of response is 2.5 months, they will be presented in the first 3 categories.

11.3.1.5. SLICC/ACR Damage Index change from baseline at Week 52/Exit Visit (Observed Case, Corrected)

The change from baseline in SLICC/ACR Damage Index score at Week 52/Exit visit will be compared between belimumab and placebo using the Exact Wilcoxon Rank Sum test with continuity correction. Only subjects with a baseline and post-baseline assessment will be included in the analysis.

The table will display the summary statistics for the SLICC/ACR Damage Index score at baseline and at Week 52/Exit visit. The table will also display the mean change from baseline, standard deviation and SE of the mean, median, minimum, maximum, 25th and 75th percentiles, plus the p-value from the Wilcoxon Rank Sum Test.

This analysis will be repeated by baseline SLICC score (0, ≥1).

11.3.1.6. SLICC/ACR Damage Index Worsening at Week 52 (WOCF)

The number and percentage subjects with worsening in their SLICC/ACR Damage Index score compared with baseline at Week 52 will be presented. Worsening is defined as a change in score >0, i.e. (Week 52/Exit visit minus Baseline) >0.

The percentage of subjects with worsening in their SLICC/ACR Damage Index score at Week 52 will be compared between belimumab and placebo using a logistic regression model. The independent variables in the model will include treatment group, baseline SLICC/ACR Damage Index score, baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world).

Missing data will be imputed using WOCF (NB: in the instance of SLICC/ACR this would mean baseline). The table will display the number and percentage of subjects with worsening in their SLICC/ACR Damage Index score by treatment group, the SE of the percentage of subjects, the treatment difference versus placebo, the odds ratio and 95% CI versus placebo, and a p-value for the odds ratio.

This analysis will be repeated by baseline SLICC score (0, ≥1).

11.3.1.7. SELENA SLEDAI Organ System Improvement by Organ System and Visit among Subjects with Organ System Involvement at Baseline (DO/TF=NR)

Details of the grouping of the SELENA SLEDAI organ systems can be found in Section 9.4.15.2.

The number and percentage of subjects with an improvement in their SELENA SLEDAI score compared to baseline by each organ system will be presented for each visit. Within each organ domain, only subjects with organ system involvement (i.e. SELENA SLEDAI organ system score >0) at baseline will be included. The renal domain will use the SLEDAI-2K proteinuria scoring as the primary method. An improvement is defined as a decrease (compared to baseline) in the SELENA SLEDAI S2K organ system score within the same organ system at a post-baseline visit.

For each organ system and visit, the table will display the number and percentage of subjects with an improvement in their SELENA SLEDAI score by treatment group and the treatment difference vs. Placebo.

11.3.1.8. SELENA SLEDAI Organ System Worsening by Organ System and Visit among Subjects with No Organ System Involvement at Baseline (LOCF)

The number and percentage of subjects with worsening in their SELENA SLEDAI score compared to baseline by each organ system and visit will be presented. Within each organ domain, only subjects with no organ system involvement (i.e. SELENA SLEDAI domain score = 0) at baseline will be included. The renal domain will use the SLEDAI-2K proteinuria scoring. Worsening is defined as an increase (compared to baseline) in the SELENA SLEDAI S2K score within the same organ system at a post-baseline visit.

Missing data due to treatment failure or study withdrawal will be imputed using LOCF. For each organ system and visit, the table will display the number and percentage of subjects with worsening in their SELENA SLEDAI score by treatment group and the treatment difference versus placebo.

11.3.1.9. SS-S2K Percent Change from Baseline by Visit (LOCF)

The percent change from baseline in SS-S2K score at each visit will be calculated as the visit score minus the baseline score divided by the baseline score then multiplied by 100. Subjects with a baseline score of zero will be excluded from the analysis due to division by zero. Belimumab and placebo will be compared using an ANCOVA model with treatment group, baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world) as covariates. The LOCF method will be employed for subjects with missing data on SS-S2K score at each visit, see Section 9.3.4 for further details.

For the baseline visit, the tables will display the mean SS-S2K score, standard deviation and SE of the mean, median, 25th and 75th percentiles, minimum and maximum. At each post-baseline visit, the tables will display the mean percent change from baseline, standard deviation and SE of the mean, median, 25th and 75th percentiles, minimum and

maximum. At week 52 the table will also include the LS mean and associated standard error, treatment difference versus placebo in LS means, 95% CI for the LS means treatment difference, and p-value from the ANCOVA model.

The mean percent change from baseline in SS-S2K score (± SE) at each visit will be presented graphically using a line graph by treatment group.

This analysis will also be performed on the SELENA SLEDAI for comparison to other trials.

11.3.1.10. SS-S2K Change from Baseline by Visit (LOCF)

The change from baseline in SS-S2K score at each visit will be calculated as the visit score minus the baseline score. Belimumab and placebo will be compared using an ANCOVA model with treatment group, baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world) as covariates. The LOCF method will be employed for subjects with missing data on SS-S2K score at each visit, see Section 9.3.4 for further details.

For the baseline visit, the tables will display the mean SS-S2K score, standard deviation and SE of the mean, median, 25th and 75th percentiles, minimum and maximum. At each post-baseline visit, the tables will display the mean change from baseline, standard deviation and SE of the mean, median, 25th and 75th percentiles, minimum and maximum. At week 52 the table will also include the LS mean and associated standard error, treatment difference versus placebo in LS means, 95% CI for the LS means treatment difference, and p-value from the ANCOVA model.

The mean change from baseline in SS-S2K score at each visit (± SE) will be presented graphically using a line graph by treatment group.

This analysis will also be performed on the SELENA SLEDAI for comparison to other trials.

11.3.1.11. PGA Percent Change from Baseline by Visit (LOCF)

The percent change from baseline in PGA at each visit will be calculated as the visit score minus the baseline score divided by the baseline score then multiplied by 100. Subjects with a baseline score of zero will be excluded from the analysis due to division by zero. Belimumab and placebo will be compared at Week 52 using an ANCOVA model with treatment group, baseline PGA score, baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world) as covariates. The LOCF method will be employed for subjects with missing data on PGA score at each visit, see Section 9.3.6 for further details.

For the baseline visit, the table will display the mean PGA score, standard deviation and SE of the mean, median, 25th and 75th percentiles, minimum and maximum. At each post-baseline visit, the tables will display the mean percent change from baseline, standard deviation and SE of the mean, median, 25th and 75th percentiles, minimum and

maximum. At week 52 the table will also include the LS mean and associated standard error, treatment difference versus placebo in LS means, 95% CI for the LS means treatment difference, and p-value from the ANCOVA model.

The mean percent change from baseline in PGA score at each visit (± SE) will be presented graphically using a line graph by treatment group.

11.3.1.12. PGA change from baseline by visit (LOCF)

The change from baseline in PGA score at each visit will be calculated as the visit score minus the baseline score. Belimumab and placebo will be compared at Week 52 using an ANCOVA model with treatment group, baseline PGA score, baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world) as covariates. The LOCF method will be employed for subjects with missing data on PGA score at each visit, see Section 9.3.6 for further details.

For the baseline visit, the tables will display the mean PGA score, standard deviation and SE of the mean, median, 25th and 75th percentiles, minimum and maximum. At each post-baseline visit, the tables will display the mean change from baseline, standard deviation and SE of the mean, median, 25th and 75th percentiles, minimum and maximum. At week 52 the table will also include the LS mean and associated standard error, treatment difference versus placebo in LS means, 95% CI for the LS means treatment difference, and p-value from the ANCOVA model.

The mean change from baseline in PGA score at each visit (± SE) will be presented graphically using a line graph by treatment group.

11.3.1.13. PGA percent of subjects with ≥0.3 point improvement from baseline by visit (DO/TF=NR)

The percent of subjects with ≥0.30 points improvement (i.e. reduction) from baseline in PGA will be compared at Week 52 between belimumab and placebo using a logistic regression model. The independent variables in the model will include treatment group, baseline PGA score, baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world). Treatment failures and handling of missing data will be managed as described in Section 9.1, Section 9.2 and Section 9.3. Subjects with baseline PGA score <0.3 will be treated as having no improvement.

At each visit, the table will display the number and percentage of subjects with ≥0.30 points improvement in their PGA score by treatment group, the SE for the percentage and the treatment difference versus Placebo. At week 52 the table will also include the odds ratio and 95% CI versus placebo, and a p-value for the odds ratio.

The percentage of subjects with ≥0.30 points improvement in their PGA score at each visit (± SE) will be presented graphically using a line graph by treatment group.

11.3.1.14. BILAG no 1A/2B organ domain scores by visit (supportive of EMA SRI-S2K)

The percent of subjects with no BILAG 1A/2B organ domain scores will be compared at Week 52 between belimumab and placebo using a logistic regression model. The independent variables in the model will include treatment group, baseline BILAG organ domain involvement (at least 1A/2B vs. at most 1B), baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world). Treatment failures and handling of missing data will be managed as described in Section 9.1, Section 9.2 and Section 9.3.

At each visit, the table will display the number and percentage of subjects with no BILAG 1A/2B organ domain scores by treatment group, the SE for the percentage and the treatment difference versus Placebo. At week 52 the table will also include the odds ratio and 95% CI versus placebo and a p-value for the odds ratio.

The percentage of subjects with no BILAG 1A/2B organ domain scores at each visit (± SE) will be presented graphically using a line graph by treatment group.

11.3.1.15. BILAG improvement by organ domain and visit in subjects with an A or B domain score at baseline (DO/TF=NR)

The number and percentage of subjects with an improvement in their BILAG score compared to baseline will be summarized for each organ domain and visit. Only subjects with an A or B domain score at Baseline will be included.

If a subject withdraws early or is deemed a treatment failure, the subject will be considered as having no improvement.

For each organ domain and visit, the table will display the number and percentage of subjects with an improvement in their BILAG score by treatment group and the treatment difference versus placebo.

11.3.1.16. BILAG worsening by organ domain and visit among subjects with no A domain score at baseline (LOCF)

The number and percentage of subjects with worsening in their BILAG score compared to baseline by each organ system domain and visit will be presented. Only subjects with no A domain score at Baseline will be included.

Missing data due to treatment failure or study withdrawal will be imputed using LOCF.

For each organ domain and visit, the table will display the number and percentage of subjects with worsening in their BILAG score by treatment group and the treatment difference versus placebo.

11.3.2. Flares

11.3.2.1. Time to first severe SFI flare over 52 weeks

See Section 11.2.2 for details of the time to first severe SFI flare analyses.

11.3.2.2. Time to first severe SFI flare from week 24 to week 52

The time to first severe SFI flare after 24 weeks is defined in Section 9.4.26. The analysis of time to first severe SFI flare after 24 weeks will be identical to Section 11.2.2.

This analysis will be performed on subjects who have at least 1 visit after the Week 24 visit. Data observed at or prior to Week 24 visit will not be included in this analysis.

11.3.2.3. Time to first SFI flare over 52 weeks

The time to first SFI flare (defined as mild, moderate, or severe) analyses will be performed using SELENA SLEDAI. The analysis of time to first SFI flare will be identical to Section 11.2.2.

A Kaplan Meier plot will be produced to show the probability of experiencing a SFI flare over 52 weeks.

11.3.2.4. Time to first SFI flare from week 24 to week 52

The time to first SFI flare (defined as mild, moderate, or severe) after 24 weeks is defined in Section 9.4.26. The analysis of time to first SFI flare after 24 weeks will be identical to Section 11.2.2.

This analysis will be performed on subjects who have at least 1 visit after Week 24 visit. Data observed at or prior to Week 24 visit will not be included in this analysis.

11.3.2.5. Severe SFI flare rate per subject-year over 52 weeks

Severe SFI flare rate per subject year is defined in Section 9.4.27.

The rate of severe flares per subject-year over 52 Weeks will be compared between treatment groups using a negative binomial regression with the number of severe flares as the dependent variable and adjusting for baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world). Adjustment will also be made for the subject's follow-up time by including log follow-up time (years) as an offset variable in the model. The tables will display the total number of severe flares and total subject-years of study follow-up in the time interval by treatment group, the unadjusted rate per subject-year, the adjusted rate per subject-year, and the adjusted rate ratio (95% CI) and p-value vs. placebo (the adjusted rates and ratio, along with the 95% CI, are a back-transformation of the values on the logarithmic scale).

11.3.2.6. SFI flare rate per subject-year over 52 weeks

Analyses of SFI flare rate (defined as mild, moderate, or severe) will be performed as defined in Section 11.3.2.5 for the severe SFI flare rate.

11.3.3. Organ-specific Measures

11.3.3.1. Time to first renal flare over 52 weeks

A SLE renal flare is defined in Section 9.4.28. Time to first renal flare is defined in Section 9.4.29.

The time to the first renal flare over 52 weeks will be compared between belimumab and placebo using a Cox proportional hazards model, adjusting for baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world). Ties will be handled using approximate likelihood (EFRON).

The table will display the number and percentage of subjects with a flare in the interval, the median days to first flare, the 25th and 75th percentile of days to first flare, and the hazard ratio and 95% CI versus placebo and the p-value from the Cox proportional hazards model. For subjects who experience a renal flare, the study day of the flare will be summarized and the table will display the median, 25th and 75th percentiles and minimum and maximum. A Kaplan Meier plot of time to first renal flare over 52 weeks will also be produced.

The analysis for time to first renal flare over 52 weeks will be repeated for subjects with baseline proteinuria >0.5 g/24hr. A Kaplan Meier plot will be produced to show the probability of experiencing a renal flare for subjects with baseline proteinuria >0.5 g/24hr over 52 weeks.

11.3.3.2. Proteinuria percent change from Baseline by visit among subjects with baseline proteinuria >0.5 g/24hr (Observed)

The percent change from baseline in proteinuria at each visit among subjects with baseline proteinuria >0.5 g/24hr will be calculated as the visit value minus the baseline value divided by the baseline value then multiplied by 100. The analysis will be performed based on the observed data in subjects with proteinuria >0.5 g/24 hour at baseline. Belimumab and placebo will be compared at Week 52 using a Wilcoxon rank sum test with continuity correction. No imputation will be done for missing data.

For the baseline visit, the tables will display the mean proteinuria, standard deviation and SE of the mean, median, 25th and 75th percentiles, minimum and maximum. At each post-baseline visit, the tables will display the mean percent change from baseline, standard deviation and SE of the mean, median, 25th and 75th percentiles, minimum and maximum. At week 52 the table will also include the p-value from the Wilcoxon rank sum test.

A line graph for the median percent change from baseline (± quartiles) in proteinuria at each visit will be presented by treatment group.

11.3.3.3. Proteinuria change from baseline by visit (Observed)

The change from baseline in proteinuria will be analyzed by treatment group at Week 52. Belimumab and placebo will be compared using a Wilcoxon rank sum test with continuity correction. This analysis will be performed based on the observed data. No imputation will be done for missing data.

For the baseline visit, the table will display the mean proteinuria, standard deviation and SE of the mean, median, 25th and 75th percentiles, minimum and maximum. At each post-baseline visit, the tables will display the mean change from baseline, standard deviation and SE of the mean, median, 25th and 75th percentiles, minimum and maximum. At week 52 the table will also include the p-value from the Wilcoxon rank sum test.

A line graph for the median change from baseline (± quartiles) in proteinuria at each visit will be presented by treatment group.

11.3.3.4. Proteinuria shifts from Baseline by visit (Observed)

Baseline proteinuria data will be summarized as the number and percent of subjects who are normal (≤0.5 g/24 hour) or high (>0.5 g/24 hour). For each post-baseline visit the data will be summarized by baseline status defined as normal or high. Among subjects normal at baseline the shifts presented will be 'No change' or 'Normal to High'. Among subjects high at baseline, the shifts presented will be 'No change' or 'High to Normal'.

Additionally, the proteinuria values will be summarized based on shifts occurring any time while on treatment. Among subjects with normal proteinuria at baseline, the percentage of subjects with at least one high post-baseline value will be presented as 'Normal to High'; subjects who never experience a high proteinuria value post-baseline will be presented as 'No change'. Among subjects with high baseline proteinuria, subjects with at least one normal post-baseline value will be presented as 'High to Normal'; subjects who never experience a normal post-baseline value will be presented as 'No change'. No statistical tests will be performed.

11.3.3.5. Percent of subjects with doubling from baseline of serum creatinine by visit (Observed)

The percent of subjects with doubling of serum creatinine from baseline to each scheduled visit will be presented by treatment group, and will, at week 52, be compared between belimumab and placebo using a logistic regression model. The independent variables in the model will include treatment group, serum creatinine level at baseline, baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world). This analysis will be performed based on the observed data only. No imputation will be done for missing data.

For each visit, the table will display the number and percentage of subjects with doubling of serum creatinine by treatment group, the SE of the percentage and the treatment difference versus Placebo. At week 52 the table will also include the odds ratio and 95% CI versus placebo, and a p-value for the odds ratio.

11.3.4. Prednisone

For analyses, all corticosteroids are converted to a prednisone equivalent average daily dose (mg/day), therefore analyses refer to average daily prednisone dose instead of average daily steroid dose. The definition and derivation of this can be found in Section 9.4.13. For these analyses (except cumulative dose), average daily prednisone dose is based on the 7–day average prednisone dose. This represents a departure from the protocol which indicates the prednisone analyses will be conducted using dosing on all days between visits. The departure is to be consistent with the reporting of prior studies for comparison purposes.

11.3.4.1. Cumulative Prednisone dose

Cumulative prednisone dose (area under the curve [AUC]) is defined as the sum of daily prednisone equivalent dose from Day 1 to Day 365 (Week 52) Visit.

The daily prednisone equivalent dose after the last visit day or the day of treatment failure, if prior to Day 365 (Week 52) Visit, will be imputed using the average of the last 28 daily prednisone equivalent doses prior to the day of last visit/treatment failure (i.e., not including day of visit/treatment failure).

For subjects who drop out or have treatment failure before Day 28, the daily prednisone equivalent dose after early dropout or treatment failure will be imputed using the average post-baseline daily doses available prior to dropout/treatment failure.

The AUC will be compared between Belimumab and placebo using a rank ANCOVA model. The covariates in the model will include treatment group, baseline prednisone dose, baseline SS-S2K score (≤ 9 vs. ≥ 10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world).

The table will display the unadjusted mean, standard deviation and SE of the mean, median, 25th and 75th percentiles, minimum and maximum of AUC and p-value from the rank ANCOVA.

11.3.4.2. Prednisone Change from Baseline by Visit (Observed)

The change from baseline in average daily prednisone dose will be summarized by treatment group and visit. At week 52 Belimumab and placebo will be compared using an ANCOVA model with treatment group, baseline prednisone dose, baseline SS-S2K score (≤ 9 vs. ≥ 10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world) as covariates. This analysis will be performed based on the observed data. No imputation will be done for missing data.

For the baseline visit, the table will display the mean, standard deviation and SE of the mean, median, 25^{th} and 75^{th} percentiles, minimum and maximum. At each post-baseline visit, the tables will display the mean change from baseline, standard deviation and SE of the mean, median, 25^{th} and 75^{th} percentiles, minimum and maximum. At week 52 the table will also include the LS mean and associated standard error, treatment difference

versus placebo in LS means, 95% CI for the LS means treatment difference, and p-value from the ANCOVA model.

A line graph for the mean change from baseline (± SE) in average daily prednisone dose will be presented by treatment group.

11.3.4.3. Prednisone Reduction by ≥50% from Baseline by Visit among subjects using Prednisone at Baseline (DO/TF=NR)

The percent of subjects with ≥50% reduction in daily prednisone dose compared to baseline will be presented by treatment group and by visit. At week 52, belimumab and placebo will be compared using a logistic regression model. Independent variables in the model will include treatment group, baseline prednisone dose, baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world). A responder is defined as having ≥50% reduction in prednisone compared to baseline. All subjects who use prednisone (dose>0) at baseline will be included in the analysis.

If a subject withdraws from the study and/or is a treatment failure prior to a scheduled visit, the subject will be considered a non-responder (i.e. no decrease in prednisone) for that and subsequent visits (DO/TF=NR).

For each visit, the table will display the number and percentage of subjects achieving a response by treatment group, the SE for the percentage and the treatment difference versus placebo. At week 52 the table will also include the odds ratio and 95% CI versus placebo and a p-value for the odds ratio.

A line graph will present the percentage of responders (± SE) at each visit.

11.3.4.4. Prednisone Increase by ≥50% from Baseline and by a Minimum of ≥5 mg/day by Visit (LOCF)

The percent of subjects with ≥50% increase in daily prednisone dose compared to baseline and by a minimum of ≥5 mg/ day will be presented by treatment group and by visit. At week 52, belimumab and placebo will be compared using a logistic regression model. Independent variables in the model will include treatment group, baseline prednisone dose, baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world).

Missing data due to treatment failure or study withdrawal will be imputed using LOCF. Subjects who are not on prednisone at baseline (dose=0) are considered to have an increase if their average daily dose increases by at least 5 mg/day.

For each visit, the table will display the number and percentage of subjects with an increase in prednisone by treatment group, the SE for the percentage and the treatment difference versus Placebo. At week 52 the table will also include the odds ratio and 95% CI versus placebo and a p-value for the odds ratio.

A line graph will present the percentage of subjects with ≥50% increase in prednisone from Baseline and by a minimum of ≥5 mg/day (± SE) at each visit in each treatment group.

11.3.4.5. Prednisone Any Increase from Baseline by Visit (LOCF)

The percent of subjects with any increase in daily prednisone dose compared to baseline will be presented by treatment group and by visit. At week 52, belimumab and placebo will be compared using a logistic regression model. Independent variables in the model will include treatment group, baseline prednisone dose, baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world).

Missing data due to treatment failure or study withdrawal will be imputed using LOCF.

For each visit, the table will display the number and percentage of subjects with an increase in prednisone by treatment group, the SE for the percentage and the treatment difference versus placebo. At week 52 the table will also include the odds ratio and 95% CI versus placebo and a p-value for the odds ratio.

A line graph will present the percentage of subjects with any increase in prednisone (± SE) at each visit in each treatment group.

11.3.4.6. Percent of Subjects with Daily Prednisone Dose Reduced to ≤7.5 mg/day from >7.5 mg/day at Baseline by Visit (DO/TF=NR)

The percent of subjects with daily prednisone dose reduced to ≤7.5 mg/day from >7.5 mg/day at baseline will be presented by treatment group and by visit. At week 52, belimumab and placebo will be compared using a logistic regression model. Independent variables in the model will include treatment group, baseline prednisone dose, baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world). A responder is defined as subject who decreased their daily prednisone dose to ≤7.5 mg/day from a baseline dose >7.5 mg/day. This analysis will be performed on subjects who used prednisone >7.5 mg/day at baseline.

If a subject withdraws from the study and/or receives a protocol-prohibited medication or a dose of allowable (but protocol-restricted) medication that results in treatment failure designation prior to a scheduled visit, the subject will be considered a non-responder (i.e. no reduction in prednisone) for that and subsequent visits (DO/TF=NR).

For each visit, the table will display the number and percentage of subjects achieving a response by treatment group, the SE for the percentage and the treatment difference versus placebo. At week 52 the table will also include the odds ratio and 95% CI versus placebo and a p-value for the odds ratio.

A line graph will present the percentage of responders (± SE) at each visit in each treatment group.

11.3.4.7. Percent of Subjects with Daily Prednisone Dose Increased to >7.5 mg/day from ≤7.5 mg/day at Baseline by Visit (LOCF)

The percent of subjects with average daily prednisone dose increased to >7.5 mg/day at each visit from ≤7.5 mg/day at baseline will be presented by treatment group and by visit. At week 52, belimumab and placebo will be compared using a logistic regression model. Independent variables in the model will include treatment group, baseline prednisone dose, baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world). This analysis will be performed on subjects who had a prednisone dose ≤7.5mg/day at Baseline.

Missing data due to treatment failure or study withdrawal will be imputed using LOCF.

For each visit, the table will display the number and percentage of subjects with an increase by treatment group, the SE for the percentage and the treatment difference versus placebo. At week 52 the table will also include the odds ratio and 95% CI versus placebo and a p-value for the odds ratio.

11.4. Missing Data Sensitivity Analyses

The following analyses will be performed if pre-specified criteria are met. As these analyses are contingent on the study results they are not considered part of the Statistical Analysis Complete (SAC) delivery but will be delivered within 2 weeks of SAC.

Within this study LOCF imputation is restricted to missing visit (but not withdrawn/treatment failure) and item level imputation within the primary and major secondary analyses (Section 9.3) and endpoint imputation for some of the other endpoints. Within the primary and major secondary responder analyses, study drop outs and treatment failures are imputed using DO/TF=NR. It is not expected that a subject, who has not been withdrawn, will be missing a week 52 (± 28 days) assessment for the primary endpoint. Therefore, no further sensitivity analyses are planned for the other secondary analyses. This is consistent with previously reported Benlysta studies.

11.4.1. Observed Case Analyses

This study was not designed to collect off treatment efficacy data. Per protocol subjects who became treatment failures should have been withdrawn from the study. However, it is likely that a small number of subjects who became treatment failures remained in the study. If more than 5% of subjects contribute off-treatment efficacy data the primary and three major secondary analyses will be repeated using the observed data (i.e., subject's response status will be assessed using their week 52 observed data, without regard to their treatment failure status). In this analysis subjects who drop out from the study will be imputed as non-responders. These analyses will only be performed for statistically significant primary and key secondary endpoints.

11.4.2. Tipping Point Analyses

If the primary analysis is statistically significant (P<0.05) then tipping point analysis will be performed for the SRI-S2K week 52 responder analysis. Additionally, tipping point analysis will be performed for each of the key secondary endpoints if it is plausible that missing data could alter the conclusions of the analysis. Plausible is defined as the p-value for the key secondary endpoint being ≤ 0.05 and > 0.001. Tipping point will only be performed for secondary endpoints considered to be statistically significant according to the prespecified testing hierarchy. All tipping point analysis will be unadjusted.

11.4.2.1. Tipping Point Methodology for Responder Endpoints

The methodology specified in this section will be applied to the following endpoints:

Primary and Key Secondary Responder Endpoints

- SRI -S2K Responder Analysis
- SRI Responder Analysis
- Percent of subjects whose average prednisone dose has been reduced by ≥25% from baseline to ≤7.5 mg/day during Weeks 40 through 52, in subjects receiving greater than 7.5 mg/day at baseline.

In the DO/TF=NR endpoint analyses, withdrawals who did not have an assessment within 28 days of week 52 or received prohibited medications will be considered non-responders. Tipping Point sensitivity analyses will be conducted to investigate the impact of the above assumptions on the study conclusions.

Additional analyses will be conducted on the response at Week 52 using an 'observed' dataset. This dataset will include all subjects who complete the Week 52 visit (including subjects who withdraw at day 365 assessment ± 28 days) without applying any imputation for subjects with completely missing data at Week 52. However, if components of the endpoint are missing, those items will have LOCF applied. Unlike the primary analysis, treatment failures will not be imputed as non-responders, but will be categorized as responder or non-responder if the subject has a Week 52 visit, or categorized as a dropout if the subject withdraws prior to the Week 52 visit window. Subjects who withdrew from the study and did not have an observed value in the Week 52 window will be excluded from this analysis.

As no covariates will be used in the tipping point analysis it is possible to enumerate all possible combinations of responder/non-responder status for dropouts on each arm. For both treatment groups the response rate will be incrementally increased by one subject (with a corresponding reduction in non-response rate), up to a maximum extreme whereby all dropouts were assumed to be responders. The statistical significance of treatment difference will be assessed by the Mantel-Hantzel chi-square test. A listing of all combinations of placebo and belimumab response rates for dropouts up to the maximum of all dropouts as responders, along with p-values will be produced, Results will be displayed graphically in a heatmap for each endpoint with the percentage of additional placebo responders on the Y-axis and belimumab responders on the X-axis.

A plausible assumption is that the response rate for dropouts in both the belimumab and placebo groups is the same as the observed response rate in the placebo group. This provides an anchor point to aid interpretation of the results of the tipping point analyses.

11.4.2.2. Tipping Point Methodology for Time to First Severe Flare endpoint

To assess the sensitivity to the censoring-at-random assumption, the missing event times for subjects who dropped out or died will be imputed under the censoring-not-at-random assumption implemented using a range of values for the true event rate for these subjects [Jackson, 2014; Zhao, 2014].

The following methods and assumptions will be applied:

The assumed event rates for censored subjects will be varied independently for each treatment group, and range from assuming that all prematurely censored subjects on a particular treatment arm experienced a severe flare at the time of censoring (worst case scenario corresponding to an infinite event rate for subjects who dropped out/died) to assuming that all prematurely censored subjects on a particular treatment arm had completed the study without experiencing a severe flare (best case scenario corresponding to zero event rate for subjects who dropped out/died). Assumed event rates for each arm will be varied incrementally between these two extreme scenarios.

A multiple imputation approach will be used to account for uncertainty about imputed event times. The multiple imputation process will follow Jackson et al [Jackson, 2014]. For a particular assumption about the post-censoring step-change in the event rate for censored subjects on each arm, this will involve the following steps:

1. Bootstrap (i.e. sample with replacement the subjects in the original dataset). This accounts for uncertainty in the estimated baseline hazard function used to impute the event times for censored subjects. The seed to be used for the bootstrap will be 115471. The number of replications will be at least 50.
2. Fit a Cox proportional hazards model (unadjusted for covariates) to the bootstrapped dataset. To inform on the validity of the adjusted Cox proportional hazards model, the proportional hazards assumption will be assessed by plotting the logarithm of the negative logarithm of the estimated survivor function against the logarithm of time, for each treatment group. If the hazards are proportional, the lines should be approximately parallel. With the fitted model
    - Extract the baseline cumulative hazard function.
    - Predict the hazard risk for subjects who discontinued prematurely.
    - Increase or decrease the predicted hazard of the subject by the pre-specified step-change amount relative to those in the same arm who stay longer in the trial. The hazard rate of subjects will be increased/decreased multiplicatively with separate multiplying factors for experimental and control

3. Impute time-to-event for subjects who dropped out or died. For a given subject, if the imputed time was below the maximum follow-up time (taken here to be 365 days), the subject will be considered to have experienced a severe SFI flare at the imputed event time. Otherwise the subject was administratively censored at 365 days (i.e., assumed to have completed the study event-free). The resulting imputed dataset assumes that all subjects have been followed until the end of the study.
4. Repeat steps 1 to 3 m times to obtain m imputed datasets.
5. Fit a Cox proportional hazards model to each imputed dataset to estimate the hazard ratio for severe SFI flare for belimumab versus placebo. The proportional hazards assumption will be assessed for a random subset of the fitted models. The analysis will allude to an average hazard ratio for interpretation of result if appropriate.
6. Combine the results of the m imputed datasets using Rubin's rules [Rubin, 1987].

These analyses assume a step change in the hazard even though there could be a more gradual change after dropout. The sensitivity analyses will not distinguish between reasons for dropout, i.e., they assume dropouts for any reason have the same step change in the event rate.

The results will be displayed in a heat map.

12. SAFETY ANALYSES

Safety will be evaluated by adverse events (AEs), changes in laboratory parameters, vital signs and immunogenicity. Safety analyses will be performed on the safety population. However, if there are more than 5% of subjects who receive a study treatment that is different from the randomized treatment through the entire study, safety analyses will be performed on the As-Treated population.

12.1. Adverse Events

- All subjects will be followed for safety through at least 8 weeks post-treatment (unless continuing treatment in the open label extension).
- A table summarizing AEs that occurred prior to treatment start date will be presented, for each system organ class (SOC) and preferred term (PT) by treatment group.
- All AEs will be classified using the standard GSK Medical Dictionary for Regulatory Activities (MedDRA) dictionary, and grouped by SOC and PT, unless otherwise stated. The investigator will evaluate all AEs with respect to seriousness, severity, and causality. The severity of an AE is to be evaluated according to the Adverse Event Severity Grading Tables in Appendix 12 and Section 7.7 of the Protocol, if a grade is defined for the AE of interest.
- All treatment-emergent AEs will be summarized for the double-blind phase.
- An overall summary of AEs will be presented showing the number and percent of subjects with at least one: AE, related AE, serious AE (SAE), severe AE, Serious and/or severe AE, AE resulting in study agent discontinuation, and deaths.
- The number and percentage of subjects experiencing an AE and the total number of AEs will be summarized for each of the following AE categories:
    - All AEs (by SOC; by SOC and PT; by PT only)
    - Serious AEs (by SOC; by SOC and PT; by PT only)
    - Severe AEs (by SOC and PT)
    - Related AEs (by SOC and PT; by PT only)
    - AEs leading to permanent discontinuation of study agent (by SOC and PT)
    - Deaths (by Category and PT)
    - Common Non-Serious AEs (by SOC and PT)
    - Common related SAEs (by SOC and PT)
    - Related Serious AEs (by SOC and PT)
    - Fatal Serious AEs (by SOC and PT)
- The tabular summary for each category of AE listed above will include the number of events, number of subjects who reported at least one event, and percentage of subjects who reported at least one AE (incidence) by treatment group for each SOC (where applicable), each PT, and overall. By default, adverse events will be sorted by MedDRA SOCs, in descending order from the SOC with the highest total incidence (i.e., summed across all treatment groups) for any adverse event within the class, to

the SOC with the lowest total incidence. If the total incidence for any two or more adverse events is equal, the events will be presented in alphabetical order. Only SOCs with observed AE PTs will be presented. Repeat sort order for MedDRA PTs within each SOC.

- Common AEs will be defined as ≥5% incidence in any treatment group.
- The table for all AEs by SOC and PT will be repeated for age group (<65, ≥65).
- A summary of AEs by SOC and severity will also be provided by treatment group. For this display, the number and percentage of subjects will be summarized as mild, moderate, or severe, based on the maximum severity observed across all PTs within the SOC for a given subject.
- A summary of AEs by SOC, PT and severity will also be provided by treatment group. The number and percentage of subjects will be summarized as mild, moderate, or severe, based on the maximum severity observed within each PT, and within each SOC.
- The hierarchical relationship between MedDRA SOCs, PTs, and verbatim text will be displayed in a table for all AEs.
- A listing of all AEs will be presented, including duration and study day of onset/resolution.
- A listing that displays which subjects reported each AE will also be produced. AEs will be grouped and sorted by SOC and PT.

12.2. Deaths, Serious Adverse Events, and Survival Status

In addition to the tabular summaries of AEs described in Section 12.1, listings for all SAEs and all deaths will be produced. The categorization of the cause of death will be adjudicated by GSK.

Survival status will also be summarized at Week 52. The number and percentage of subjects will be summarized for each of the following categories:

- Alive ('date last known alive' is >= study day 351 **or** entered open label extension)
- Unknown (Including subjects where the 'date last known alive' was < study day 351 (Week 50))
- Lost to Follow-up
- All Deaths
    - Fatal SAEs that started during the Double-Blind phase although death occurred in a later study phase (i.e. alive at Week 52)
    - Based on stop date of Fatal SAE (i.e. date of death) being in the double-blind phase.

A listing of survival status at Week 52 will also be presented.

12.3. Adverse Events Leading to Discontinuation of Investigational Product

In addition to the tabular summaries described in Section 12.1, a listing of all AEs leading to permanent discontinuation of study treatment will be produced.

12.4. Adverse Events of Special Interest

The Benlysta Program Safety Analysis Plan (PSAP) has been developed to include an adverse event of special interest (AESI) analysis for consistent reporting across belimumab studies. Categorizations for the AESIs will be defined in the PSAP and reporting of AESIs for these analyses is defined in Attachment 3. An overall summary of AESIs will be presented and each specific category of AESI will be presented separately by PT. The number and percentage of subjects with at least one occurrence and the number of events of the following AESI will be provided.

A listing of AESI will be produced by treatment group and Special Interest category.

The following AESIs will be identified using the list of preferred terms in Attachment 3 and adjudicated by GSK.

Adverse Events
Adverse Events of Special Interest (AESI)
AESI will be defined per the version of the PSAP/MedDRA in effect at the time of DBR. Malignant Neoplasms • Malignancies Excluding non-melanoma skin cancer (NMSC) • Malignancies Including NMSC • Solid Tumour • Hematologic • Skin (All) • NMSC • Excluding NMSC • Tumours of unspecified malignancy adjudicated as malignant per GSK Post-Infusion Systemic Reactions (PISR) • PISR per Anaphylactic Reaction Customized MedDRA Query (CMQ) narrow search • PISR per Anaphylactic Reaction CMQ broad search • PISR per Anaphylactic Reaction CMQ algorithmic search • Serious Anaphylaxis per Sampson Criteria per GSK adjudication • Serious Acute PISR/Hypersensitivity Per GSK adjudication – Serious Acute PISR Excluding Hypersensitivity per GSK adjudication – Serious Acute Hypersensitivity Reactions per GSK adjudication • Serious Delayed Acute Hypersensitivity Reactions per GSK adjudication • Serious Delayed Non-Acute Hypersensitivity Reactions per GSK adjudication All Infections of Special Interest (Opportunistic Infections (OI), Herpes Zoster (HZ), Tuberculosis (TB), And Sepsis; All and Serious, separately) • All opportunistic infections (OI) per GSK adjudication • OI per GSK adjudication excluding Tuberculosis and Herpes Zoster

Adverse Events
Adverse Events of Special Interest (AESI)
• Active Tuberculosis – Non-Opportunistic – Opportunistic • Herpes Zoster – Non-Opportunistic – Opportunistic ▪ Recurrent ▪ Disseminated • Sepsis <u>Depression (including mood disorders and anxiety)/suicide/self-injury (All and Serious, separately)</u> • Depression (including mood disorders and anxiety) (excluding suicide and self-injury) • Suicide/self-injury • Serous suicide/self-injury per GSK adjudication – Suicidal Behavior ▪ Completed Suicide – Suicidal Ideation – Self-injurious Behavior without Suicidal Intent • Deaths

For malignant neoplasms events identified as “tumours of unspecified malignancy” will be reviewed for classification as malignant per GSK adjudication.

Post-infusion systemic reactions will be presented using nine different definitions as indicated above. These will also be presented by category and PT.

Post-infusion systemic reactions per Anaphylactic Reaction CMQ narrow search are defined as at least one event from the list of preferred terms in Category A as listed in Attachment 3.

Post-infusion systemic reactions per Anaphylactic Reaction CMQ broad search are defined as at least one event from the list of preferred terms in either Category A, B, C or D as listed in Attachment 3. The SAEs that are identified by the Anaphylactic Reaction CMQ broad search will be reviewed further for potential classification as serious post-infusion systemic reactions per GSK adjudication.

For CSR reporting, all post-infusion/injection systemic reaction AESIs defined via narrow, broad, or algorithmic search, the AEs need to have occurred on the day of an infusion/injection or within 3 days after an infusion/injection. See Section 9.4.35 for the definition of the 3 day assessment window. GSK will review all serious events identified via the broad search occurring within 21 days after an infusion/injection, and adjudicate these events as post-infusion/injection systemic reactions or hypersensitivity reactions per the criteria in Section 9.4.35 Adverse events with partial or missing start dates will be included unless there is evidence through comparison of partial dates to suggest otherwise

Possible cases of serious anaphylaxis per Sampson criteria will be identified as defined in Attachment 3.

Infection AESIs will also be presented by category and PT. Infections AESI leading to discontinuation will be presented by category and PT.

Depression, Suicide and Self-injury Adverse Events of Special Interest will be presented by Category and PT. Depression, suicide and self-injury are defined using terms from the 'Depression (excluding suicide and self injury)' CMQ plus additional terms added by MAH and 'Suicide/self-injury' CMQ.

12.4.1. Post-infusion systemic reactions by infusion

Summaries of post-infusion systemic reactions that occur on the day of an infusion or within 3 days after an infusion will be presented by the first six infusions and PT for the following:

- Post-Infusion Systemic Reactions per Anaphylactic Reaction CMQ broad search (on the day of the infusion or within 3 days after the infusion)
- Serious Post-Infusion Systemic Reactions per Anaphylactic Reaction CMQ broad search (on the day of the infusion or within 3 days after the infusion)
- Serious Acute Post-Infusion Systemic Reactions/Hypersensitivity per GSK adjudication
- Serious Delayed Acute Hypersensitivity Reactions per GSK adjudication
- Serious Delayed Non-Acute Hypersensitivity Reactions per GSK adjudication
- Post-Infusion Systemic Reactions Adverse Events of Special Interest by Category and PT
- Serious Post-Infusion Systemic Reactions Adverse Events of Special Interest by Category and PT

12.5. Columbia-Suicide Severity Rating Scale (C-SSRS)

Suicidality assessments are completed at every visit during the double-blind phase, but not completed during the open-label phase. Assessments are done using the C-SSRS. If a "yes" response is given to any suicidal behavior or a "yes" response to suicidal ideation questions 3, 4 or 5 on the C-SSRS, the investigator will be prompted to complete the Possible Suicidality Related Questionnaire (PSRQ). A listing of the PSRQ will be presented.

Listings will be generated for the following:

- Suicidal ideation and behavior data for subjects who have any suicidal ideation or behavior recorded at any point on the study (including screening)
- Behavior details for subjects who have any suicidal behavior recorded at any point on the study (including screening)

- The most severe suicidal ideation details for subjects who have any suicidal ideation recorded at any point on the study (including screening).

12.5.1. C-SSRS suicidal ideation or behavior during treatment

The number and percentage of subjects with each category of suicidal ideation or behavior during treatment (including all assessments post Day 1) will be presented for the double-blind phase, selecting the worst record a subject has for each category. The categories of suicidal ideation and behavior are presented in increasing order of severity from 1 to 10. For the rows pertaining to suicidal behavior, the number of subjects who have the specified behavior at least once during treatment is presented. For the rows pertaining to suicidal ideation, the number of subjects whose maximum ideation at any on-treatment assessment is the specified ideation is presented. Within each category, subjects may have more than one type of suicidal ideation and behavior.

12.5.2. C-SSRS suicidal ideation or behavior relative to pre-treatment

The number and percentage of subjects with treatment-emergent suicidal ideation or behavior during treatment (including all assessment post Day 1) will be presented. A subject must have at least one pre-treatment assessment and at least one on-treatment assessment in order to be included in this display. Any assessments performed on Day 1 will be considered pre-treatment. A subject may have more than one treatment-emergent suicidal ideation and/or behavior.

12.5.3. C-SSRS shift changes in categories from pre-treatment to on-treatment

A summary of the shift from maximum pre-treatment C-SSRS (Day 1 or before) category to maximum on-treatment (up to and including Week 52) category will be produced. The pre-treatment period is based on the lifetime evaluation up to and including Day 1. A subject must have at least one pre-treatment assessment and at least one on-treatment assessment in order to be included in this display. The table will display the number and percentage of subjects within the specific shift categories: No suicidal ideation or behavior, suicidal ideation, and suicidal behavior.

12.6. Clinical Laboratory Evaluations

For laboratory analyses, only analytes with a numeric normal range will be analyzed and summaries and analyses will be performed based on the observed data. No imputation will be done for missing data. Baseline is defined as described in Section 9.4.1. See Appendix 10 of the Protocol for a list of laboratory parameters.

Listings will be generated for all laboratory results and for Grade 3 or Grade 4 laboratory toxicity results, by laboratory category (hematology, liver function, electrolytes, other chemistries, urinalysis and immunoglobulins) and treatment group.

12.6.1. Laboratory descriptive statistics by visit

Descriptive statistics for each analyte will be displayed by treatment group for each visit. The tables will display the mean value, standard error, standard deviation, median, 25th and 75th percentiles, minimum and maximum. No statistical tests will be performed.

A line graph will be produced for each analyte which displays the mean value (±SE) by visit and treatment group.

12.6.2. Worst laboratory toxicity grade post-baseline

Laboratory toxicity will be graded using Adverse Event and Laboratory Toxicity Grading Tables (Section 18.11) when possible. The worst laboratory toxicity grade post-baseline for each laboratory parameter within each laboratory category (hematology, liver function, electrolytes, other chemistries, urinalysis and immunoglobulins) will be presented. Data for both scheduled and unscheduled visits will be included throughout the double-blind phase of the study (including follow-up).

12.6.3. Laboratory toxicity ≥2 grade shift post-baseline

Toxicity grade shifts from baseline of ≥2 grades will be summarized for each laboratory parameter within each laboratory category (hematology, liver function, electrolytes, other chemistries, urinalysis and immunoglobulins). The table will display the number and percentage of subjects with at least one ≥2 grade shift as well as the specific shift categories: Grade 0 to 2, Grade 0 to 3, Grade 0 to 4, Grade 1 to 3, Grade 1 to 4, Grade 2 to 4 and hyper to hypo Laboratory reference range shifts from baseline. Data for both scheduled and unscheduled visits will be included throughout the double-blind phase.

For laboratory tests without toxicity grades, shifts relative to the normal range will be summarized for each analyte as shifts 'to low' and shifts 'to high.' For the 'to low category' the percentage of subjects with at least one low post-baseline value relative to the baseline will be displayed using the categories: no shift to low and normal/high to low. For the 'to high category' the percentage of subjects with at least one high post-baseline value relative to baseline will be displayed using the categories: no shift to high and normal/low to high. No statistical tests will be performed.

A laboratory value that is above the testing laboratory's normal range will be considered a high abnormal laboratory value. A laboratory value that is below the testing laboratory's normal range will be considered a low abnormal value.

12.6.4. Immunoglobulin reference range shifts from baseline by visit

For immunoglobulins (IgG, IgA, and IgM) reference range shifts will be summarized across all visits based on the baseline normal range category. For subjects with immunoglobulin values below the LLN, the number and percentage of subjects who 'remained low' or went 'to normal/high' post-baseline will be summarized. Similarly, for subjects with immunoglobulin values within the normal range or above the ULN, the number and percentage of subjects who 'remained normal/high' or went 'to low' post-baseline will be summarized.

12.6.5. Immunoglobulin below LLN by visit

The number and percentage of subjects with immunoglobulin values (IgG, IgA, and IgM) below the LLN at each visit will also be presented for all subjects and then repeated for subjects ≥LLN at baseline. No statistical tests will be performed.

12.6.6. Immunoglobulin above LLN by visit

The number and percentage of subjects with immunoglobulin values (IgG, IgA, and IgM) above the LLN at each visit will also be presented for all subjects and then repeated for subjects <LLN at baseline. No statistical tests will be performed.

12.6.7. Immunogenicity

For the immunogenicity assessment, two types of antibody assays will be performed, i.e. a binding assay and neutralizing assay. For the binding assay, there will be 3-testing steps. A screening assessment is performed which produces a result of positive or negative. For samples with a positive screening result, a confirmation assay is then carried out, which also produces a result of positive or negative. For samples with a positive confirmation result, a titer value will also be obtained to quantify the degree of binding in a titration assay step. Patients will be viewed as positive for the binding assay if the confirmation assay was positive. Subjects, who tested positive for the binding assay, will be tested for the neutralizing assay, which again produces a result of positive or negative.

For the incidence of patients with positive binding antibody, two tables will be produced summarizing results for the binding antibody assay by (i) treatment group and visit and (ii) a shift table. The table will include the number and proportion of subjects in each results category for each visit (including early withdrawal visit). Binding confirmatory assay results will be categorized as negative, persistent positive (defined as a positive immunogenic response on at least 2 consecutive assessments or a single result at the final assessment) or transient positive (defined as a single positive immunogenic response that does not occur at the final assessment).

A listing of titer values will be produced for anyone who was positive at any time during the study.

12.7. Other Safety Measures

12.7.1. Vital Signs

A summary of change from baseline of vital signs will be presented by visit and by treatment group. The tables will display the mean change from baseline, standard deviation, median, 25th and 75th percentiles, minimum and maximum.

A listing of all subjects' vital signs will be presented by treatment group.

12.7.2. Concomitant Procedures/Surgery

A listing of all concomitant procedures/surgery will be presented.

13. HEALTH OUTCOMES ANALYSES

13.1. FACIT-Fatigue Scale score change from baseline by visit (LOCF)

A composite fatigue score will be created from the FACIT-Fatigue questionnaires (http://www.facit.org). The change from baseline in FACIT-Fatigue Scale score at each visit will be calculated as the visit score minus the baseline score. Belimumab and placebo will be compared, at week 52, using an ANCOVA model with treatment group, baseline FACIT-Fatigue Scale score, baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (U.S./Canada vs Rest of World). The LOCF method will be employed for subjects with missing data on FACIT-Fatigue Scale score at each visit, see Section 9.4.29 and Section 6.1 (LOCF) for further details.

For the baseline visit, the tables will display the mean FACIT-Fatigue Scale score, standard deviation, SE of the mean, median, 25th and 75th percentiles, minimum and maximum. At each post-baseline visit, the tables will display the mean change from baseline, standard deviation, SE of the mean, median, 25th and 75th percentiles, minimum and maximum, and at Week 52 LS mean and associated standard error, treatment difference versus placebo in LS means, 95% CI for the LS means treatment difference, and p-value from the ANCOVA model.

A listing of all subjects' FACIT-Fatigue scale scores will be presented.

The mean change from baseline in FACIT-Fatigue Scale score (±SE) at each visit will be presented graphically using a line graph by treatment group.

13.2. FACIT-Fatigue Scale Score Improvement exceeding the MCID (≥4 points) by visit (DO/TF=NR)

A composite fatigue score will be created from the FACIT-Fatigue questionnaires (http://www.facit.org). The percent of subjects with ≥4 points improvement (i.e. increase) from baseline in FACIT-Fatigue Scale score will be presented by treatment group and by visit, and will be compared, at week 52, between belimumab and placebo using a logistic regression model. The independent variables in the model will include treatment group, baseline FACIT-Fatigue Scale score, baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world). Treatment failures and handling of missing data will be managed as described in Section 9.1, Section 9.2 and Section 9.3. Subjects with baseline FACIT-Fatigue Scale score >48 will be treated as having no improvement.

At each visit, the table will display the number and percentage of subjects with ≥4 points improvement in their FACIT-Fatigue Scale score by treatment group, and the treatment difference versus placebo. At week 52, the table will also display the odds ratio and 95% CI versus placebo, and a p-value for the odds ratio.The percentage of subjects with ≥4 points improvement in their FACIT-Fatigue Scale score (±SE) at each visit will be presented graphically using a line graph by treatment group.

14. CLINICAL PHARMACOLOGY DATA ANALYSES

14.1. Pharmacokinetic Analyses

The PK summaries will be performed using the PK population, defined in Section 6.

Descriptive statistics for serum belimumab concentrations will be displayed for each visit for the double-blind phase. Summary statistics will be calculated including mean, geometric mean, as well as SD, %CV, median, 25th and 75th percentiles, minimum, maximum and 95% confidence intervals for mean and geometric means. No statistical tests will be performed.

A line graph will be produced which displays the median value along with 25th and 75th percentiles by visit for the double-blind phase.

To assess the effect of body size and immunogenicity on belimumab exposure the table and line graph described above will also be repeated for the subgroups of baseline BMI category and baseline body weight quartiles and may be repeated for immunogenicity status (persistent positive vs. transient persistent vs. negative) if the number of subjects with a positive status warrant it.

A listing of serum belimumab PK concentration-time data will be presented.

15. BIOMARKER DATA ANALYSIS

The biomarker analyses will be performed for the mITT population as defined in Section 6 unless otherwise stated.

For the duration of the study, biomarker data (serum immunoglobulin isotypes IgA and IgM and B cell results) that have the potential to unblind the study team will not be transferred to the blinded study team. Instead, blinded datasets will be required which contain dummy results. These blinded datasets will be exact models of the real datasets which will be received following unblinding. This will ensure that programs written using blinded data will still run on the real treatment codes and real unblinded data following the first database lock.

A listing of Biomarker data will be presented.

The intention is to analyze the Biomarker data using ANCOVA. However, if the normality assumptions are violated, or there are a significant proportion of outliers in the dataset, then other methods of analysis will be investigated The first alternative to be investigated will be the non-parametric Wilcoxon Rank Sum Test. If a non-parametric analysis is more appropriate, then the analysis will be updated and the relevant statistics will be presented in the output tables. If analyses are changed to non-parametric due to the distribution of the data, then Figures presenting the data will switch from presenting means (± SE) to medians (± quartiles).

15.1. Immunoglobulins, Autoantibodies, and Complement

Immunoglobulin isotypes (IgG, IgA, and IgM), autoantibodies (anti-dsDNA, ANA, aCL (IgA, IgG, IgM), lupus anticoagulant, beta-2-glycoprotein, anti-sm, anti-RNP, anti-SS-A, anti-SS-B, anti-ribosomal P) and complement (C3 and C4) will be assessed.

15.1.1. Percent change from baseline in immunoglobulins, autoantibodies, and complement by visit (Observed)

The percent change from baseline for immunoglobulins, autoantibodies, and complement will be summarized by treatment group and visit. Belimumab and placebo will be compared, at week 52, using an ANCOVA model with treatment group, baseline biomarker value, baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world) as covariates. These analyses will be performed based on the observed data. No imputation will be done for missing data.

For the baseline visit, the table will display the mean, standard deviation and SE of the mean, median, 25th and 75th percentiles, minimum and maximum. At each post-baseline visit, the tables will display the mean percent change from baseline, standard deviation and SE of the mean, median, 25th and 75th percentiles, minimum and maximum. At week 52 the table will also display the LS mean and associated standard error, treatment difference versus placebo in LS means, 95% CI for the LS means treatment difference, and p-value from the ANCOVA model.

A line graph for the mean percent change from baseline (±SE) in each of these biomarkers will be presented by treatment group.

Additionally, for autoantibodies (anti-dsDNA, ANA), analyses will be performed on subjects who were positive at baseline (anti-dsDNA ≥30 IU /mL, ANA index ≥0.80). For complement (C3 and C4), analyses will be performed on subjects with low values at baseline (C3 <90 mg/dL and C4 <10 mg/dL).

A line graph for the mean percent change from baseline (±SE) in each of these biomarkers in these subgroups will be presented by treatment group.

15.1.2. Change from baseline in immunoglobulins, autoantibodies, and complement by visit (Observed)

The change from baseline for immunoglobulins, autoantibodies, and complement will be summarized by treatment group and visit. Belimumab and placebo will be compared, at week 52, using an ANCOVA model with treatment group, baseline biomarker value, baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world) as covariates. These analyses will be performed based on the observed data. No imputation will be done for missing data.

For the baseline visit, the table will display the mean, standard deviation and SE of the mean, median, 25th and 75th percentiles, minimum and maximum. At each post-baseline visit, the tables will display the mean change from baseline, standard deviation and SE of the mean, median, 25th and 75th percentiles, minimum and maximum, At Week 52 the table will also display LS mean and associated standard error, treatment difference versus placebo in LS means, 95% CI for the LS means treatment difference, and p-value from the ANCOVA model.

A line graph for the observed mean percent change from baseline (±SE) in each of these biomarkers will be presented by treatment group.

Additionally, for autoantibodies (anti-dsDNA, ANA), analyses will be performed on subjects who were positive at baseline (anti-dsDNA ≥30 IU /mL, ANA index ≥0.80). For complement (C3 and C4), analyses will be performed on subjects with low values at baseline (C3 <90 mg/dL and C4 <10 mg/dL).

A line graph for the mean percent change from baseline (±SE) in each of these biomarkers in these subgroups will be presented by treatment group.

15.1.3. Shifts in immunoglobulins, autoantibodies, and complement by visit

Shift tables will be used to summarize the changes in immunoglobulins, autoantibodies, and complement by visit.

For IgG, IgA, and IgM baseline data will be summarized as the number and percent of subjects who are low or normal/high at baseline. For post-baseline visits the data will be

summarized by baseline status defined as low (IgG <6.94 g/L, IgA <0.81 g/L, and IgM <0.48 g/L) or normal/high (IgG ≥6.94 g/L, IgA ≥0.81 g/L, and IgM ≥0.48 g/L). Among subjects low at baseline the shifts presented will be low to normal/high and low to low. Among subjects normal/high at baseline, the shifts presented will be normal/high to normal/high and normal/high to low. At Week 52 the shifts will be evaluated using Fisher's exact test within each baseline category.

For anti-dsDNA, baseline data will be summarized as the number and percent of subjects who are positive and negative at baseline. For post-baseline visits the data will be summarized by baseline status defined as positive (≥30 IU/mL) or negative (<30 IU/mL). Among subjects positive at baseline the shifts presented will be positive to negative and positive to positive. Among subjects negative at baseline, the shifts presented will be negative to negative and negative to positive. At week 52, the shifts will be evaluated using Fisher's exact test within each baseline category.

For C3 and C4, baseline data will be summarized as the number and percent of subjects who are low or normal/high at baseline. For post-baseline visits the data will be summarized by baseline status defined as low (C3 <90 mg/dL, C4 <10 mg/dL) or normal/high (C3 ≥90 mg/dL, C4 ≥10 mg/dL). Among subjects low at baseline the shifts presented will be low to normal/high and low to low. Among subjects normal/high at baseline, the shifts presented will be normal/high to normal/high, and normal/high to low. At week 52 the shifts will be evaluated using Fisher's exact test within each baseline category.

15.2. B cell Analyses

A listing of B cell results will be presented.

15.2.1. B cell percent change from baseline by visit (Observed)

The following biomarkers will be summarized by treatment group and visit:

- Percent change in absolute B cell subsets (CD19+, CD20+, CD20+/27+ memory, CD20+/27– naïve, CD20+/69+ activated, CD20+/138+ plasmacytoid, CD27+ ^{BRIGHT}/CD20– Short-lived plasma cells, CD19+/27^{BRIGHT}/38^{BRIGHT} SLE subset, CD19+/CD24^{HIGH}/CD38^{HIGH} transitional and CD20-/138+ plasma cells) at each visit.

Belimumab and placebo will be compared, at week 52, using an ANCOVA model with treatment group, baseline B cell value, baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world) as covariates. These analyses will be performed based on the observed data. No imputation will be done for missing data.

For the baseline visit, the table will display the mean, standard deviation and SE of the mean, median, 25th and 75th percentiles, minimum and maximum. At each post-baseline visit, the tables will display the mean percent change from baseline, standard deviation and SE of the mean, median, 25th and 75th percentiles, minimum and maximum. At week 52 the table will also display the LS mean and associated standard error, treatment

difference versus placebo in LS means, 95% CI for the LS means treatment difference, and p-value from the ANCOVA model.

A line graph for the mean percent change from baseline (±SE) in each of these B cells will be presented by treatment group.

15.2.2. B cell change from baseline by visit (Observed)

The following biomarkers will be summarized by treatment group and visit:

- Change in absolute B cell subsets (CD19+, CD20+, CD20+/27+ memory, CD20+/27– naïve, CD20+/69+ activated, CD20+/138+ plasmacytoid, CD27+ ^{BRIGHT}/CD20– Short-lived plasma cells, CD19+/27^{BRIGHT}/38^{BRIGHT} SLE subset, CD19+/CD24^{HIGH}/CD38^{HIGH} transitional and CD20-/138+ plasma cells) at each visit.
- Change in CD20+/27– naïve and CD20+/27+ memory B cell subsets expressed as a percentage of CD19

Belimumab and placebo will be compared, at week 52, using an ANCOVA model with treatment group, baseline B cell value, baseline SS-S2K score (≤9 vs. ≥10), baseline complement levels (At least one C3/C4 Low vs. No C3/C4 Low) and region (US/Canada vs. rest of the world) as covariates. These analyses will be performed based on the observed data. No imputation will be done for missing data.

For the baseline visit, the table will display the mean, standard deviation and SE of the mean, median, 25th and 75th percentiles, minimum and maximum. At each post-baseline visit, the tables will display the mean percent change from baseline, standard deviation and SE of the mean, median, 25th and 75th percentiles, minimum and maximum. At week 52 the table will also display the LS mean and associated standard error, treatment difference versus placebo in LS means, 95% CI for the LS means treatment difference, and p-value from the ANCOVA model.

A line graph for the mean percent change from baseline (±SE) in each of these B cells will be presented by treatment group.

16. PHARMACOGENETIC DATA ANALYSES

Any pharmacogenetic analyses will be described in a separate pharmacogenetic analysis plan and will be reported separately from the main clinical study report.

17. REFERENCES

Alarcón-Segovia D, Tumlin JA, Furie RA, et al. LJP 394 for the prevention of renal flare in patients with systemic lupus erythematosus. Arthritis Rheum 2003;48(2):442–54.

Buyon JP, Petri MA, Kim MY, et al. The effect of combined estrogen and progesterone hormone replacement therapy on disease activity in systemic lupus erythematosus: A randomized trial. Ann Intern Med 2005;142(12 Pt 1):953-62.

Cockcroft DW and Gault MH. Prediction of creatinine clearance from serum creatinine. Nephron 1976;16:31-41.

Cohen J. 1960. A coefficient of Agreement for Nominal Scales Educational and Psychology.

Fleiss JL, Cohen J, Everitt B.S. 1969. Large Sample Standard Errors of Kappa and Weighted Kappa Psychological Bulletin 72,323-327.

Fleiss JL, Cohen J. 1973. The Equivalence of Weighted Kappa and Intraclass Correlation.

Gladman DD, Ibanez D, Urowitz MB. Systemic lupus Erythematosus disease activity index 2000. J Rheumatol 2002;29(2):288-91.

Gladman EM, Ginzler E, Goldsmith C et al. SLICC/ACR damage index for SLE. Arthritis Rheum. 1996a;39(3):363-9.

Hay EM, Bacon PA, Gordon C, et al. The BILAG index: a reliable and valid instrument for measuring clinical disease activity is systemic lupus erythematosus. Q J Med 1993;86:447-58.

Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum. 1997;40(9):1725.

Isenberg DA, Gordon C. From BILAG to BLIPS--disease activity assessment in lupus past, present and future. Lupus 2000;9(9):651-4.

Jackson D, White IR, Seemain S, et al. Relaxing the independent censoring assumptions in the Cox proportional hazards model using multiple imputation. Stat Med. 2014;33:4681-94.

Neergard. Reducing the risk of blood thinners. Associated press, September 2006.

Peduzzi P, Concato J et al. A simulation of the number of events per variable in logistic regression analysis, J Clin Epidemiol. 1996 Dec;49(12):1373-9.

Petri M, Buyon J, Kim M. Classification and definition of major flares in SLE clinical trials. Lupus. 1999;8:685-91.

Petri M, Kim MY, Kalunian KC, et al. Combined oral contraceptives in women with systemic lupus erythematosus. N Engl J Med 2005;353(24):2550-8.

Rubin DB. Multiple Imputations for Nonresponse in Surveys. John Wiley & Sons. 1987.

Tan EM, Cohen AS, Fries JF, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum. 1982;25(11):1271-7.

Yee CS, Development and Validation of the BILAG 2004 Index for the Assessment of Disease Activity in Systemic Lupus Erythematosus. Doctoral Thesis submitted to the University of Birmingham, Rheumatology Research Group, Division of Infection and Immunology, School of Medicine, February 2008.

Yee CS, Farewell V, Isenberg DA, et al. British Isles Lupus Assessment Group 2004 index is valid for assessment of disease activity in systemic lupus erythematosus. Arthritis Rheum. 2007;56(12):4113-9.

Yee CS, Isenberg DA, Prabu A, et al. BILAG-2004 index captures systemic lupus erythematosus disease activity better than SLEDAI-2000. Ann Rheum Dis. 2008;67(6):873-6.

Zhao Y, Herring AH, Zhou H, et al. A multiple imputation method for sensitivity analyses of time-to-event data with possibly informative censoring. J Biopharm Stat. 2014;24:229-53.

18. APPENDICES

18.1. Appendix 1 Allowable Medication Categories

Medication Category	Rule
Anti-malarials	Set to "ANTIMALARIALS" if the preferred term begins with "QUINACRINE", "QUININE", "HYDROXYCHLOROQUINE", "MEPACRINE", or CHLOROQUINE" AND the route of administration is not 'TOPICAL', 'VAGINAL', 'CONJUNCTIVAL', 'INTRANASAL', 'INHALATION', 'INTRA-OCULAR', 'INTRATRACHEAL', 'EPIDURAL', 'INTRA-ARTICULAR', or 'OTHER'.
Steroids	Set to 'STEROIDS' if at least one associated ATC code (ATCCD1 – ATCCD6) begins with 'H02' AND Route of administration is "INTRADERMAL", "INTRAMUSCULAR", "INTRAVENOUS", "ORAL", "SUBCUTANEOUS", or "INTRA-ARTICULAR".
Immunosuppressants	Set to 'IMMUNOSUPPRESSANTS' if at least one associated ATC code (ATCCD1 – ATCCD6) begins with 'L04A' or the preferred term begins with "CYCLOPHOSPHAMIDE" (oral and parenteral routes) or "MERCAPTOPURINE" (oral route) or "METHOTREXATE" AND route of administration is not "TOPICAL" or "CONJUCTIVAL".
NSAIDs	Set to NSAIDs if at least one associated ATC code (ATCCD1 – ATCCD6) begins with 'M01A'.
Aspirin	Set to "ASPIRIN" if preferred term contains "ACETYLSALICYLIC ACID" or "ACETYLSALICYLATE LYSINE".
Prohibited	Set to "PROHIBITED" if any of the following conditions are met, if preferred term equals "INVESTIGATIONAL DRUG", "IMMUNOGLOBULIN", "ADALIMUMAB", "BELIMUMAB", "ETANERCEPT", "INFLIXIMAB", "RITUXIMAB", "ABATACEPT", "ANAKINRA", "GOLIMUMAB", "CERTOLIZUMAB", "PLASMAPHERESIS", "TOCILIZUMAB" or "CYCLOPHOSPHAMIDE" (IV route).
NOTE: Plasmapheresis in the 'Prohibited' category would be recorded in the Surgery/Procedures form, not the concomitant medication form.	

18.2. Appendix 2 Prednisone Equivalent Conversion

- A concomitant medication is identified as a steroid if at least one associated ATC code (ATCCD1 – ATCCD6) begins with 'H02.' Mineralocorticords are included in the group of ATC codes beginning with 'H02' but do not have sufficient anti-inflammatory properties to be considered as a prednisone equivalent. For this reason, the conversion factor has been set to 0.
- The following routes are considered to provide systemic exposure: oral, subcutaneous, intramuscular, intradermal, and intravenous. Although not systemic, intra-articular steroids are also identified for treatment failure rules. Topical routes of administration are excluded (e.g., topical, conjunctival, intranasal).
- At data base release, all preferred terms identified with an ATC code beginning with 'H02' will be reviewed to ensure a conversion factor exists for all terms with a systemic route of administration.
- Similarly, all routes of administration for preferred terms with an ATC code beginning with 'H02'will be reviewed to ensure all systemic routes have been identified in the list above.
- In order to be converted the frequency and dose of the steroid must be present with the unit dose in milligrams (mg) or grams (g). Doses recorded in grams will be converted to milligrams by multiplying the dose in grams by 1000 prior to applying the conversion factor.
- Reported dose for systemic steroid is converted to prednisone equivalent dose using conversion factor for each particular medication (refer to online calculator http://www.globalrph.com/corticocalc.htm).

Prednisone Equivalent Daily Dose (mg)	=	Collected Dose (mg)	x	Conversion Factor	x	Frequency Factor

Prednisone Conversion Factors (mg)	
Preferred term	**Conversion factor for converting to a prednisone-equivalent dose**
BETAMETHASONE	8.3333
BETAMETHASONE DIPROPIONATE	8.3333
BETAMETHASONE SODIUM PHOSPHATE	8.3333
BETROSPAM	8.3333
BUDESONIDE	0.3333
CELESTONA BIFAS	8.3333
CORTISONE	0.2
CORTISONE ACETATE	0.2
CRONOLEVEL	8.3333
DEFLAZACORT	0.8333
DEPO-MEDROL MED LIDOKAIN	1.25
DEXAMETHASONE	6.6667

Prednisone Conversion Factors (mg)	
Preferred term	**Conversion factor for converting to a prednisone-equivalent dose**
DEXAMETHASONE ACETATE	6.6667
DEXAMETHASONE SODIUM PHOSPHATE	6.6667
FLUDROCORTISONE	0
FLUOCORTOLONE	3
HYDROCORTISONE	0.25
HYDROCORTISONE ACETATE	0.25
HYDROCORTISONE SODIUM SUCCINATE	0.25
MEPREDNISONE	1.25
METHYLPREDNISOLONE	1.25
METHYLPREDNISOLONE ACETATE	1.25
METHYLPREDNISOLONE SODIUM SUCCINATE	1.25
PARAMETHASONE	2.5
PREDNISOLONE	1
PREDNISOLONE ACETATE	1
PREDNISOLONE SODIUM PHOSPHATE	1
PREDNISOLONE SODIUM SUCCINATE	1
PREDNISONE	1
PREDNISONE ACETATE	1
TRIAMCINOLONE	1.25
TRIAMCINOLONE ACETATE	1.25
TRIAMCINOLONE ACETONIDE	1.25

Combination Products: Prednisone Conversion Factors (mg)		
Preferred term	**Ingredients**	**Conversion factor for converting to a prednisone-equivalent dose**
CELESTONA BIFAS	Betamethasone acetate + Betamethasone sodium phosphate	8.3333
DEPO-MEDROL MED LIDOKAIN	Methylprednisolone + Lidocaine	1.25
CELESTAMINE	Betamethasone + Dexchlorpheniramine maleate	8.3333
SYNBETAMINE	Betamethasone + Dexchlorpheniramine maleate	8.3333

Frequency Factors	
Frequency	**Factor**
BID	2
BIW	2/7
HS	1
OAM	1/30
ONCE	1
PRN	null
Q2H	12
Q2W	1/14
Q3D	1/3
Q3H	8
Q3MO	1/84
Q3W	1/21
Q4D	1/4
Q4H	6
Q4W	1/28
Q6H	4
Q8H	3
QAM	1
QD	1
QH	24
QHS	1
QID	4
QOD	1/2
QPM	1
QW	1/7
QWK	1/7
TID	3
TIW	3/7
UNK	null
2 TIMES PER WEEK	2/7
3 TIMES PER WEEK	3/7
EVERY 2 WEEKS	1/14
EVERY 3 WEEKS	1/21
EVERY 4 WEEKS	1/28
EVERY WEEK	1/7

18.3. Appendix 3 Treatment Failure Rules

General Conventions

- Assessment of dose is based on analysis dose. Analysis dose is daily dose adjusted for dosing frequency, and in the case of steroid is converted to prednisone equivalents. Exception is intra-articular dose where analysis dose is not populated.
- In the event there are multiple visit dates within a visit window, the date for determining Treatment Failure, will be the earliest date at which efficacy data (from PGA, BILAG, SLEDAI and SLE Flare) is recorded. This may be different than the earliest visit within that window
- If a critical visit (Day 113 [Week 16], 169 [Week 24], 309 [Week 44]) is missing from efficacy data (from PGA, BILAG, SLEDAI and SLE Flare) then the date is imputed e.g., date of Day 169 visit is imputed as the target day for Day 169 (Week 24).
- Actual visit date, not target visit date, is used to assess treatment failures. For example, the Day 169 (Week 24) visit can occur on study day 169 ± 7 days. If the subject's Day 169 (Week 24) study visit occurs on Day 171, the date for Day 171 is used when applying the treatment failure rules.
- Prohibited medications/dosages started on the day the subject **completes** the double-blind treatment phase do not result in TF designation.
- Prohibited medications/dosages started on the date of early withdrawal are considered a TF (see clarification below for steroids). If the prohibited medication/dose starts after the date of withdrawal it will not be part of the TF assessment.
- Generally, only the **first** instance of each unique TF rule violation type is output programmatically. If this instance is adjudicated as a not being a TF then the clinical adjudicators will review the entirety of the relevant concomitant medication records to assess if the subject subsequently became a TF for the same violation type (e.g., steroid dose does not return to within 25% or 5mg, whichever is higher, above baseline dose by Day 169 [Week 24] visit).
- Clinical may amend the date of TF during adjudication if, for instance, a subject did not meet the criterion on the date identified by the program (as may be the case if their steroid usage was short term and not SLE-related) but did meet it later.

Steroids

- SLE-related steroids are steroids where medication type is marked as 'Systemic Lupus Erythematosus (SLE)' (CMTYPCD=156), regardless of reason for medication (CMREAS).
- Total steroids include steroids for SLE and non-SLE reasons.
- Baseline dose is the 7-day average based on the 7 days **prior** to, but not including, treatment start date.

- The Day 309 (Week 44) steroid dose is the sum of steroid dose over 7 consecutive days leading to and including the Day 309 (Week 44) visit, divided by 7. The Day 309 steroid dose is used to determine if there is a new increase in steroids above the baseline (Day 1) or Week 44 (Day 309) visit within 8 weeks of the Day 365 (Week 52) visit. Within 8 weeks of the Day 365 (Week 52) visit is defined as occurring in the window starting the day after the Week 44 (Day 309) visit through the Week 52 (Day 365) visit. Note there is no check that the Day 309 (Week 44) and Day 365 (Week 52) visits are within 8 weeks of each other.
- In all instances in which the protocol states that a subject's steroid dose must return to a specified level (e.g., within 5 mg or 25% of baseline whichever is higher) by a specific visit day (e.g., Day 169 [Week 24] visit), the calculation of the 7-day average steroid dose to determine whether a subject is a treatment failure will begin on the day after the visit. This does not apply to the calculation of the Day 309 (Week 44) visit dose (see above).
- When assessing dose at critical visits the average dose is based on the 7 days after the visit e.g., Day 169 (Week 24) 7-day average dose is the average of day 170 -176 (if Day 169 occurred on the actual target date; or 7 days after the date of the Day 169 [Week 24] visit otherwise).
- The final week interval for subjects who complete the double-blind treatment period will be 7 days prior to the exit visit date; [exit visit -7 days] to [exit visit -1 day]. The day of the exit visit will not be included.
- The final interval for subjects who withdraw early from the double-blind phase will be the 7 days up to and including the exit visit date; [exit visit/early withdrawal date -6] to [exit visit/ early withdrawal date]. The day of withdrawal will be included.
- If a subject meets the criterion for TF based on a 7-day average, the date of TF will be the **last day** of the 7-day interval.
- The above rules may miss a subject who starts/increases a steroid close to the day of withdrawal. In this situation it is possible that the subject may have withdrawn prior to the end of the interval for computing their 7-day average exceeding the TF threshold. Hence all subjects whose dose of steroid increased within 7 days (including day of withdrawal) of withdrawal, and have not already crossed the dose threshold for TF based on the 7-day average rule, will be output for clinical adjudication. If clinical adjudication determines these to be treatment failures the date of treatment failure will be set to the date of withdrawal.
- A subject who is receiving 0mg of steroid at baseline will be allowed to take ≤5mg of steroid at critical assessments without being considered a treatment failure.
- Intra-articular steroids are not included in average steroid dose calculations.
- All assessments relating to within 8 weeks of Day 365 (Week 52) are based off the interval from Day 309 (Week 44) visit date to the Day 365 (Week 52) visit date; not Day 365 (Week 52) visit date – 64 days.
- QOD dosing regimens (and regimens with frequency <once/day) will be reviewed to ensure the analysis dose is calculated correctly. Consider an example of a subject taking 5mg QOD and 7.5mg QOD. To calculate an analysis average daily dose for a

QOD (every other day) regimen, half of the dose is attributed to each day in the dosing interval. In this example, 2.5mg would be assigned as the analysis dose for each day of the 5mg QOD dosing interval and 3.75mg would be the analysis dose for each day of the 7.5mg QOD dosing interval. The analysis dose for a given day is the sum of all steroid doses for the day. If the 5mg QOD dose is recorded as starting one day prior to the 7.5mg QOD dose and no other steroids were taken on that day, then the analysis dose for the first day of the 5mg QOD will be 2.5mg; for subsequent days when the 5mg and 7.5mg dosing intervals overlap, the analysis dose will be 2.5mg + 3.75mg = 6.25mg.

Systemic Steroids for SLE-related Disease Activity

- A subject who fails to return to within 25% or 5 mg over the baseline (Day 1) dose, whichever is higher, by the Day 169 (Week 24) visit will be considered a **treatment failure**.
- After the Day 169 (Week 24) visit, an increase >25% or >5 mg over the baseline (Day 1) dose, whichever is higher, for SLE activity will deem the subject a **treatment failure**.
- Within 8 weeks before the Day 365 (Week 52) visit, no new increase over the baseline (Day 1) or Day 309 (Week 44) visit dose, whichever is higher, is allowed. A new increase would deem the subject a **treatment failure**.

Intra-articular (IA) injections

- Subjects may receive intra-articular injections between baseline (Day 1) and the Day 309 (Week 44) visit. Intra-articular (IA) injections after the Day 309 (Week 44) visit and before the Day 365 (Week 52) visit will result in **treatment failure**.

Steroids for Reasons Other Than SLE Disease Activity

<u>From Day 1 to the Day 169 (Week 24) Visit:</u>

- Steroids may be given for reasons other than SLE disease activity (such as asthma, contact dermatitis) as clinically indicated until Day 169 (Week 24) visit.

<u>From Days 169 to 309 (Weeks 24 to 44) Visits:</u>

- Steroids may be given for reasons other than SLE disease activity from the Day 169 (Week 24) visit until the Day 309 (Week 44) visit at any dose/duration that does not result in a total steroid dose (for SLE and non-SLE reasons) >25% or >5 mg, whichever is higher, over the baseline dose. The total baseline steroid dose will be calculated from steroids given for both SLE and non-SLE reasons. Any total steroid dose exceeding this rule will deem the subject a **treatment failure**.
- In addition, steroids for non-SLE reasons may be given short term at higher doses according to the following guidelines:
  - Up to 750 mg (prednisone) for 1 day,

  and/or

- Up to 100 mg/day (prednisone) for 2-3 days,

and/or

- Up to 40 mg/day (prednisone) for 4-7 days.

The duration of high dose steroid use for reasons other than SLE must not exceed 7 days, after which time, tapering should begin. The total steroid dose must be tapered to within 25% or 5 mg over the baseline dose, whichever is higher, within 30 days of the 1st dose of a course of steroids. In addition, the steroid dose must be tapered to within 25% or 5 mg over the baseline dose, whichever is higher, by the Day 309 (Week 44) visit. Otherwise the subject will be deemed a **treatment failure**.

From the Day 309 to 365 (Week 44 to 52) Visits:

After the Day 309 (Week 44) visit through the Day 365 (Week 52) visit, no new steroids are allowed for reasons other than SLE activity that result in a total daily steroid dose >25% or >5 mg, whichever is higher, over the baseline total steroid dose. A subject will be considered a **treatment failure** for any steroid use 8 weeks before the Day 365 (Week 52) visit that does not meet this criterion.

Anti-Malarials

- Dose of anti-malarial at baseline is the dose the subject received on the treatment start date. If a patient is receiving an anti-malarial at baseline it will be assumed that dose is stable and not a loading dose
- Dose of anti-malarial at Day 113 is the dose of received on the Day 113 (Week 16) visit date.
- Treatment failure date will be the anti-malarial start date that resulted in treatment failure.
- Clinical loading dose is permitted for initiation or replacement. Whether or not the dose was a loading dose will be assessed by clinical adjudication.
- A new anti-malarial (e.g., hydroxychloroquine, chloroquine, quinacrine) may be started between Day 1 and the Day 113 (Week 16) visit.
- The dose of an anti-malarial may be reduced during the course of the study. The dose of an anti-malarial may be increased as clinically required, up to the Day 113 (Week 16) visit.
- After the Day 113 (Week 16) visit, any increase in dose of an anti-malarial over the baseline (Day 1) or Day 113 (Week 16) visit dose, whichever is higher, will declare the subject a **treatment failure**.
- Starting any new anti-malarial treatment after the Day 113 (Week 16) visit will declare the subject a **treatment failure**.
- An anti-malarial treatment will be considered new if the subject did not receive an anti-malarial at any time during the Day 1 to Day 113 (Week 16) treatment interval.

- An anti-malarial may be replaced by another anti-malarial due to documented toxicity or lack of availability at any time during the study. Replacement due to toxicity/lack of availability will be assessed during clinical adjudication.

Immunosuppressant/Immunomodulatory agents

- Baseline dose is the dose received on the treatment start date.
- Whether the dose was a loading dose will be assessed by clinical adjudication.
- Replacement due to toxicity/lack of availability will be assessed during clinical adjudication.
- No check is made for medications not listed in protocol which required medical monitor approval. Conjunctival agents will not be considered in treatment failure assessment

NSAIDs and Aspirin

- NSAIDs may be given as clinically indicated until the Day 309 (Week 44) visit.
- For subjects who never received an NSAID between the Day 1 and Day 309 (Week 44) visit, starting a new NSAID after the Day 309 (Week 44) visit will declare the subject a **treatment failure** unless the NSAID is given for <1 week. The programming algorithm will need to check if an NSAID was taken between the Day 1 and Day 309 (Week 44) visit.
- An NSAID may be replaced with another NSAID due to documented toxicity or lack of availability. Replacement due to toxicity/lack of availability will be assessed during clinical adjudication; therefore, programming should identify first unique NSAID terms that were taken post-baseline that were not present at baseline.
- Daily doses of aspirin up to 1000 mg/day are allowed at any time during the study.
- Daily doses of aspirin above 1000 mg/day may be initiated at any time up to Day 309 (Week 44) visit and may continue through the end of the study.
- For subjects who never received an aspirin regimen at a dose >1000 mg/day between the Day 1 and Day 309 (Week 44) visit, starting a new aspirin regimen at a dose >1000 mg/day after the Day 309 (Week 44) visit will declare the subject a **treatment failure** unless the aspirin is given for <1 week. The programming algorithm will need to check if an aspirin was received at a dose >1000 mg/day between the Day 1 and Day 309 (Week 44) visit.
- Baseline is defined as the NSAID/aspirin received on the treatment start date. Program to checks if subject received NSAID from Baseline to day 309.If so, they are not output for adjudication.
- Treatment failure date will be the date the medication/dose was started resulting in treatment failure designation.

Prohibited medications/non-Drug Therapies

- Date of treatment failure is date subject started Prohibited medications/non-Drug Therapy.

- The following medications and therapies are forbidden at any time during the study:
  - Other investigational agents (biologic or non-biologic). Investigational applies to any drug not approved for sale in the country in which it is being used. [No check for investigational agents not approved for sale in country is being made.]
  - Co-enrollment into another study of an investigational agent or another study that may interfere with the conduct of this protocol. [No check available for this criterion.]
  - Anti-TNF therapy (e.g., adalimumab, certolizumab pegol, etanercept, golimumab, infliximab).
  - Other biologics (e.g., rituximab, abatacept, interleukin-1 receptor antagonist [anakinra], tocilizumab).
  - Intravenous immunoglobulin (IVIG).
  - IV cyclophosphamide.
  - Plasmapheresis

Live Vaccines

- Receiving a live vaccine is prohibited due to safety reasons but is not a treatment failure criterion.

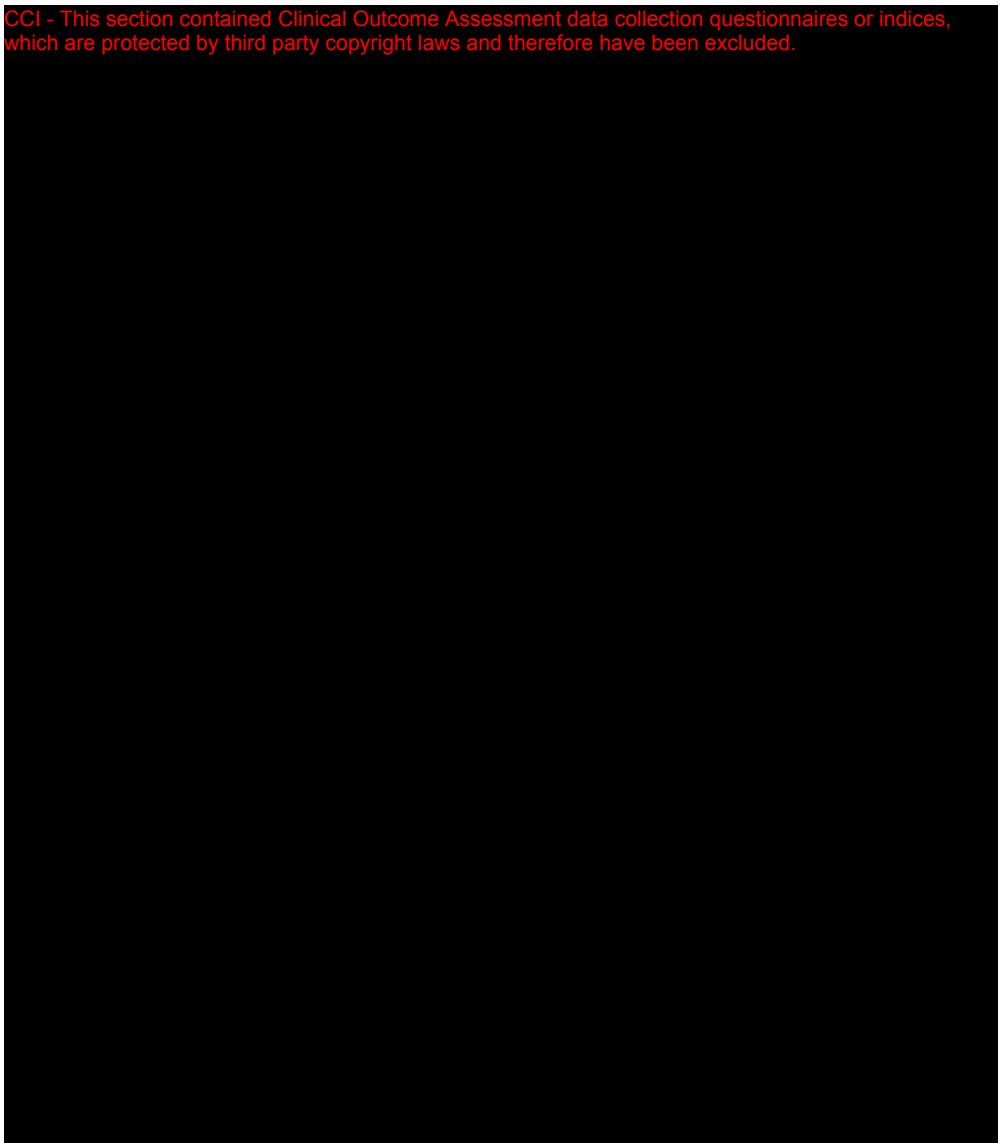
18.4. Appendix 4 ACR Criteria at Baseline

CCI - This section contained Clinical Outcome Assessment data collection questionnaires or indices, which are protected by third party copyright laws and therefore have been excluded.

18.5. Appendix 5 SLICC/ACR Damage Index

CCI - This section contained Clinical Outcome Assessment data collection questionnaires or indices, which are protected by third party copyright laws and therefore have been excluded.

18.6. Appendix 6 BILAG Assessment



18.7. Appendix 7: SLE Flare Index

Mild-moderate flare	Severe flare
☐ Change in SELENA SLEDAI instrument score of 3 points or more (but not to more than 12), or	☐ Change in SELENA SLEDAI instrument score to greater than 12a, or
☐ New/worse: Discoid, photosensitive, profundus, cutaneous vasculitis, bullous lupus Nasopharyngeal ulcers Pleuritis Pericarditis Arthritis Fever (SLE) Or	☐ New/worse: CNS-SLE Vasculitis Nephritis Myositis Plt<60,000 Hemolytic anemia with Hb<70g/L or decrease in Hb>3g/L **Requiring**: double prednisone, or prednisone increase to >0.5mg/kg/day, or hospitalization, or
☐ Increase in prednisone, but not to >0.5mg/kg/day, or	☐ Increase in prednisone to >0.5mg/kg/day, or
☐ Added NSAID or hydroxychloroquine for SLE activity, or	☐ New cyclophosphamide, azathioprine, methotrexate, or mycophenolate for SLE activity, or
☐ ≥ 1.0 Increase in PGA score, but not to more than 2.5	☐ Hospitalization for SLE activity, or
	☐ Increase in a PGA score to ≥ 2.5

a. The modified SLE Flare index excludes severe flares that are triggered only by an increase is SLEDAI Score to greater than 12.

(Adapted from [Buyon, 2005; Petri, 2005; Neergard, 2006. Petri, 1999])

18.8. Appendix 8 Important Protocol Deviations

A summary of important protocol deviations is given below. Further detail is given in the Protocol Deviation Management Plan (PDMP): Dated: 25Feb2018 (Version 3.0).

Category	GSK Important Deviations
ICF Process	Subject never signed ICF or amendment ICF signed after study procedure done
Eligibility Criteria	Any "no" response to inclusion criteria and/or any "yes" response to exclusion criteria and subject was subsequently randomized and dosed Study drug ADMINISTERED at Day 0 PRIOR to confirmation that subject meets all eligibility criteria AND subject subsequently determined to be ineligible
IP assignment in IXRS	Study drug assigned in IXRS prior to confirming subject meets eligibility on Day 0; then determined not to be eligible and subject NEVER dosed
Misstratification	Subject misstratified in IXRS
IP Infusion Time	Less than 50 minutes
Study Procedures Not Done	Missing Screening AND Day 0 SRI components (includes SS or BILAG or PGA)
Missing Laboratory Values at Randomization	Subjects randomized without results available for review and, upon availability of lab results, subject ineligible for inclusion
Not Withdrawn from Study after developing Protocol Specific Withdrawal Criteria	Subject developed withdrawal criteria specified in the protocol but were not withdrawn and continued dosing/study **Note: Determine if GSK medical monitor allowed patient to continue if this criterion was met after patient completed Week 48 of IP dosing. If GSK medical monitor granted permission, then NOT Important
Prohibited Medication	Subject used prohibited medication(s) and/or therapies at any time during the study.
Incorrect IP or IP dose administered	Site administered wrong IP to subject Dose of >=20mg/kg/day given at more than 2 consecutive visits

Category	GSK Important Deviations
Failure to report SAE	SAE discovered as never reported
Liver Stopping Criteria	Patient met liver stopping criteria and not reported to GSK OR Liver stopping criteria reported but patient received additional doses of IP without GSK approval
Blinding and Unblinding	An event that led to improper unblinding to Sponsor, CRO, site and/or subject

18.9. Appendix 9 Per Protocol Population Exclusions

According to the RAP and the PDMP, if a subject meets the following major study criteria, then they will be excluded from the Per Protocol population:

- Received an incorrect treatment most of the time (>50% of the time).
- Did not self-identify as black race (Inclusion Criterion 2).
- Did not have a clinical diagnosis of SLE according to the American College of Rheumatology (ACR) criteria (Inclusion Criterion 3)
- Did not have active SLE disease defined as a SELENA SLEDAI score >= 8 at screening (Inclusion Criterion 4)
- Did not have unequivocally positive anti-nuclear antibody (ANA) test results from 2 independent time points as defined in the inclusion criteria (Inclusion Criterion 5)
- Was not on a stable SLE treatment regimen at baseline as defined in the protocol (Inclusion Criterion 6)
- Received an excluded medication prior to Day 0 (Exclusion Criteria 1-6)
- Missed 3 or more consecutive study agent infusions
- Study blind/unblind procedures Investigator/site staff/GSK Clinical team did not remain blinded to treatment assignment through Week 52/Exit visit efficacy evaluation
- Other, a deviation that does not satisfy the above criteria, however, in the judgment of the clinical team, including the medical monitor, constitutes an exclusion from the Per Protocol population

Specific Adjudications: All violations will be discussed and adjudicated as important or not important and for exclusion from the Per Protocol population.

18.10. Appendix 10 Laboratory Parameters

<table>
<tr><th>Hematology</th><th>Urinalysis</th><th>Modified Chem-20</th></tr>
<tr><td>Total white blood cell count
Differential:
• Absolute Neutrophils
  ○ Segmented Neutrophils
  ○ Band Neutrophils
  ○ Myelocytes
  ○ Metamyelocytes
  ○ Promyelocytes
• Lymphocytes
• Monocytes
• Eosinophils
• Basophils
Hemoglobin
Hematocrit
Red blood cell (RBC) count
Platelet count
Prothrombin time (PT)
Partial thromboplastin time (PTT)</td><td>Protein
Glucose
Ketones
Occult blood
Microscopic examination including:
• WBC per hpf
• RBC per hpf
• Dysmorphic RBC
• Casts (specified by type e.g., RBC, WBC)
Spot Urine
• Protein:creatinine ratio
Urine Pregnancy</td><td rowspan="3">Electrolytes:
• Sodium
• Potassium
• Magnesium
• Chloride
• Carbon dioxide
• Calcium adjusted for Albumin
• Inorganic Phosphate
Enzymes:
• AST
• ALT
• Alkaline Phosphatase
• GGT
• LDH
Other:
• Creatinine
• Blood urea nitrogen
• BUN/creatinine ratio
• Bilirubin, total
• Protein, total
• Albumin
• Uric acid
• Glucose
• Estimated Creatinine Clearance/GFR (Cockcroft-Gault)</td></tr>
<tr><th>Biological Markers</th><th>Autoantibodies</th></tr>
<tr><td rowspan="2">FACS of peripheral lymphocytes:
B lymphocytes
• CD19+
• CD20+
• CD20+/27+memory
• CD20+/27+ memory (%CD19)
• CD20+/27– naïve
• CD20+/27– naïve (%CD19)
• CD20+/69+ activated
• CD20+/138+plasmacytoid
• CD27+^{BRIGHT}/CD20– Short-lived plasma cells
• CD19+/27^{BRIGHT}/38^{BRIGHT} SLE subset
• CD19+/CD24^{HIGH} /CD38^{HIGH} regulatory B cell
• CD20-/138+plasma cells
Complement
• C3
• C4
BLyS Protein</td><td>ANA
Anti-dsDNA
aCL
Lupus anti-coagulant
±beta-2-glycoprotein-1**
ENAs**</td></tr>
<tr><th>Immunoglobulins</th><th>PK/Immunogenicity</th></tr>
<tr><td>Serum immunoglobulin isotypes:
IgG, IgM, IgA</td><td>PK assessment
Immunogenicity assessment</td></tr>
</table>

18.11. Appendix 11 Adverse Event and Laboratory Toxicity Grading Tables

HEMATOLOGY	GRADE 1 MILD	GRADE 2 MODERATE	GRADE 3 SEVERE	GRADE 4 POTENTIALLY LIFE-THREATENING
Hemoglobin	> 9.5 - 11.0 g/dL	> 8.0 – 9.5 g/dL	6.5 - 8.0 g/dL	< 6.5 g/dL
Leukocytes	3000-3999/mm3	2000-2999/mm3	1000-1999/mm3	< 1000/mm3
Absolute Neutrophil Count	1500-1999/mm3	1000-1499/mm3	500-999/mm3	< 500/mm3
Platelets	75,000 - 99,999/mm3	50,000 – 74,999/mm3	25,000 - 49,999/mm3	< 25,000/mm3
Prothrombin Time (PT)	> 1.0-1.25 x ULN*	> 1.25-1.5 x ULN	> 1.5-3.0 x ULN	> 3.0 x ULN
Partial Thromboplastin Time (PTT)	> 1.0-1.66 x ULN	> 1.66-2.33 x ULN	> 2.33-3.0 x ULN	> 3.0 x ULN
Methemoglobin	5.0-10.0 %	10.1-15.0 %	15.1-20.0 %	> 20%

(continued)

*ULN = Upper Limit of Normal.

Modified from DMID Adult Toxicity Tables, 2001

CHEMISTRIES	GRADE 1 MILD	GRADE 2 MODERATE	GRADE 3 SEVERE	GRADE 4 POTENTIALLY LIFE-THREATENING
Sodium				
Hyponatremia	130-135 meq/L	123-129 meq/L	116-122 meq/L	< 116 meq/L
Hypernatremia	146-150 meq/L	151-157 meq/L	158-165 meq/L	> 165 meq/L
Potassium				
Hypokalemia	3.0-3.4 meq/L	2.5-2.9 meq/L	2.0-2.4 meq/L	< 2.0 meq/L
Hyperkalemia	5.6-6.0 meq/L	6.1-6.5 meq/L	6.6-7.0 meq/L	> 7.0 meq/L
Phosphate				
Hypophosphatemia	2.0-2.4 mg/dL	1.5-1.9 mg/dL	1.0-1.4 mg/dL	< 1.0 mg/dL
Calcium- (Corrected For Albumin)				
Hypocalcemia	7.8-8.4 mg/dL	7.0-7.7 mg/dL	6.1-6.9 mg/dL	< 6.1 mg/dL
Hypercalcemia	10.6-11.5 mg/dL	11.6-12.5 mg/dL	12.6-13.5 mg/dL	>13.5 mg/dL
Magnesium				
Hypomagnesemia	1.2-1.4 meq/L	0.9-1.1 meq/L	0.6-0.8 meq/L	< 0.6 meq/L
Albumin				
Hypoalbuminemia	3.00-3.49 g/dL	2.50-2.99 g/dL	2.00-2.49 g/dL	< 2.00 g/dL
Bilirubin (Total)				
Hyperbilirubinemia (Total)	> 1.0-1.5 x ULN	> 1.5-2.5 x ULN	> 2.5-5 x ULN	> 5 x ULN
Glucose				
Hypoglycemia	55-64 mg/dL	40-54 mg/dL	30-39 mg/dL	< 30 mg/dL
Hyperglycemia (nonfasting & no prior diabetes)	116-160 mg/dL	161-250 mg/dL	251-500 mg/dL	> 500 mg/dL
Triglycerides	151-399 mg/dL	400-750 mg/dL	751-1200 mg/dL	> 1200 mg/dL
Creatinine	> 1.0-1.5 x ULN	> 1.5-3.0 x ULN	> 3.0-6.0 x ULN	> 6.0 x ULN

(continued)

Modified from DMID Adult Toxicity Tables, 2001

CHEMISTRIES (continued)	GRADE 1 MILD	GRADE 2 MODERATE	GRADE 3 SEVERE	GRADE 4 POTENTIALLY LIFE-THREATENING
Uric Acid				
Hyperuricemia	7.5-10.0 mg/dL	10.1-12.0 mg/dL	12.1-15.0 mg/dL	> 15.0 mg/dL
Liver Transferases (AST, ALT, and GGT)	1.25-2.5 x ULN	> 2.5-5.0 x ULN	> 5.0-10.0 x ULN	> 10.0 x ULN
Alkaline Phosphatase	1.25-2.5 x ULN	> 2.5-5.0 x ULN	> 5.0-10.0 x ULN	> 10.0 x ULN
Pancreatic Enzymes				
Amylase	> 1.0-1.5 x ULN	> 1.5-2.0 x ULN	> 2.0-5.0 x ULN	> 5.0 x ULN
Pancreatic amylase	> 1.0-1.5 x ULN	> 1.5-2.0 x ULN	> 2.0-5.0 x ULN	> 5.0 x ULN
Lipase	> 1.0-1.5 x ULN	> 1.5-2.0 x ULN	> 2.0-5.0 x ULN	> 5.0 x ULN
*Hypoglobulinemia (IgG)**	550-700 mg/dL	400-549 mg/dL	250-399 mg/dL	< 250 mg/dL

(continued)

*(Goldfarb et al, 2001; Yamani et al, 2001; Eibl and Rosen, 1995).

Modified from DMID Adult Toxicity Tables, 2001

URINALYSIS	GRADE 1 MILD	GRADE 2 MODERATE	GRADE 3 SEVERE	GRADE 4 POTENTIALLY LIFE-THREATENING
Proteinuria:				
Dipstick: Protein	1 +	2-3 +	4 +	Nephrotic syndrome
Spot Urine: Protein:Creatinine Ratio mg/mg	0.2-1.0	> 1.0-2.0	> 2.0-3.5	> 3.5
24 hour Urine: Protein	200 mg – 1 g loss/day	> 1-2 g loss/day	> 2-3.5 g loss/day	Nephrotic syndrome OR > 3.5 g loss/day
Hematuria	Microscopic only < 10 RBC/hpf	Gross, No clots ≥ 10 RBC/hpf	Gross plus clots OR RBC casts	Obstructive OR transfusion required

(continued)

RBC = red blood cell; hpf = high power field.

Modified from DMID Adult Toxicity Tables, 2001

CARDIOVASCULAR	GRADE 1 MILD	GRADE 2 MODERATE	GRADE 3 SEVERE	GRADE 4 POTENTIALLY LIFE-THREATENING
Cardiac Arrhythmia	-	Asymptomatic/transient; dysrhythmia; no treatment req	Recurrent/persistent dysrhythmia. Symptomatic; treatment req	Unstable dysrhythmia hospitalization and treatment required
Hypotension	Transient orthostatic hypotension, no treatment	Symptoms correctable with oral fluid treatment	IV fluid req, no hospitalization req	Hospitalization req
Hypertension	Transient, increase > 20 mm/Hg; no treatment	Recurrent; chronic increase > 20 mm/Hg, treatment req	Acute treatment req; out patient hospitalization possible	Hospitalization req
Pericarditis	Minimal effusion	Mild/moderate asymptomatic effusion, no treatment	Symptomatic effusion, pain, ECG changes	Tamponade OR pericardiocentesis OR surgery req
Hemorrhage, Blood Loss	-	Mildly symptomatic; no treatment required	Gross blood loss OR 1-2 units transfused	Massive blood loss OR > 2 units transfused

(continued)

Modified from DMID Adult Toxicity Tables, 2001

GASTROINTESTINAL	GRADE 1 MILD	GRADE 2 MODERATE	GRADE 3 SEVERE	GRADE 4 POTENTIALLY LIFE-THREATENING
Nausea	Mild OR transient; reasonable intake maintained	Mod discomfort OR intake decreased for < 3 days	Severe discomfort OR minimal intake for ≥ 3 days	Hospitalization required
Vomiting	Mild OR transient; 2-3 episodes/day OR mild vomiting lasting < 1 week	Mod OR persistent; 4-5 episodes per day; OR vomiting lasting ≥ 1 week	Severe vomiting of all foods/fluids in 24 hours OR orthostatic hypotension OR IV treatment req	Hypotensive shock OR hospitalization required for IV treatment req
Diarrhea	Mild or transient; 3-4 loose stools per day OR mild diarrhea lasting < 1 week	Mod OR persistent; 5-7 loose stools per day or diarrhea lasting ≥ 1 week	Bloody diarrhea; OR orthostatic hypotension OR > 7 loose stools/day OR IV treatment req	Hypotensive shock OR hospitalization req
Oral Discomfort/Dysphagia	Mild discomfort, no difficulty swallowing	Difficulty swallowing but able to eat and drink	Unable to swallow solids	Unable to drink fluids; IV fluids req
Constipation	Mild	Moderate	Severe	Distention with vomiting

(continued)

Modified from DMID Adult Toxicity Tables, 2001

RESPIRATORY	GRADE 1 MILD	GRADE 2 MODERATE	GRADE 3 SEVERE	GRADE 4 POTENTIALLY LIFE-THREATENING
Cough (for aerosol studies)	Transient; no treatment	Treatment associated cough; inhaled bronchodilator	Uncontrolled cough; systemic treatment req	-
Bronchospasm Acute	Transient; no treatment; FEV1 70% to < 80% (or peak flow)	treatment req; normalizes with bronchodilator; FEV1 50% to < 70% (or peak flow)	No Normalization with bronchodilator; FEV 25% to < 50% (or peak flow), retractions	Cyanosis; FEV1 < 25% (or peak flow) OR intubated
Dyspnea	Dyspnea on exertion	Dyspnea with normal activity	Dyspnea at rest	Dyspnea requiring O2 therapy

(continued)

MISCELLANEOUS	GRADE 1 MILD	GRADE 2 MODERATE	GRADE 3 SEVERE	GRADE 4 POTENTIALLY LIFE-THREATENING
Fever (oral > 12 hours)	37.7-38.5°C or 100.0-101.5°F	38.6-39.5°C OR 101.6-102.9°F	39.6-40.5°C OR 103-105°F	> 40.5°C OR > 105°F
Headache	Mild; No treatment req	Mod; or non-narcotic analgesia treatment	Severe; OR responds to initial narcotic treatment	Intractable; OR requiring repeated narcotic treatment
Allergic Reaction	Pruritus without rash	Localized urticaria	Generalized urticaria angioedema	Anaphylaxis
Cutaneous/Rash/ Dermatitis	Erythema, pruritus rash OR dry desquamation	Diffuse maculopapular OR dry desquamation	Vesiculation OR moist desquamation ulceration	ANY ONE: mucous membrane involvement, suspected Stevens-Johnson (TEN), erythema multiforme, necrosis req surgery, exfoliative dermatitis
Local Reaction (secondary to parenteral treatment- not vaccination or skin test)	Erythema	Induration < 10 mm OR inflammation OR phlebitis	Induration > 10 mm OR ulceration	Necrosis of skin
Fatigue	Normal activity Reduced < 25%	Normal activity Reduced 25-50%	Normal activity reduced > 50%; cannot work	Unable to care for self

(continued)

Modified from DMID Adult Toxicity Tables, 2001

NEUROLOGIC	GRADE 1 MILD	GRADE 2 MODERATE	GRADE 3 SEVERE	GRADE 4 POTENTIALLY LIFE-THREATENING
Neuro-cerebellar	Slight incoordination OR dysdiadochokinesia	Intention tremor OR dysmetria OR slurred speech OR nystagmus	Ataxia requiring assistance to walk or arm incoordination interfering with ADLs	Unable to stand
Neuro-psych/ mood	-	none	Severe mood changes requires medical intervention	Acute psychosis requiring hospitalization
Paresthesia (burning, tingling, etc)	Mild discomfort; no treatment needed	Mod discomfort non-narcotic analgesia req	Severe discomfort; OR narcotic analgesia req with symptomatic improvement	Incapacitating; OR not responsive to narcotic analgesia
Neuro-motor	Mild weakness in muscle of feet but able to walk and/or mild increase or decrease in reflexes	Mod weakness in feet (unable to walk on heels and/or toes), mild weakness in hands, still able to do most hand tasks and/or loss of previously present reflex or development of hyperreflexia and/or unable to do deep knee bends due to weakness	Marked distal weakness (unable to dorsiflex toes or foot drop), and mod proximal weakness ie, in hands interfering with ADLs and/or requiring assistance to walk and/or unable to rise from chair unassisted	Confined to bed or wheelchair because of muscle weakness
Neuro-sensory	Mild impairment sensations, (ie, vibratory, pinprick, hot/cold in great toes) in focal area or symmetrical distribution	Mod impairment mod de-sensation, (ie, of vibratory, pinprick, hot/cold to ankles) and/or joint position or mild impairment that is not symmetrical.	Severe impairment (dec or loss of sensation to knees or wrists) or loss of sensation of at least mod degree in multiple different body areas (ie, upper and lower extremities)	Sensory loss involves limbs and trunk

(concluded)

Modified from DMID Adult Toxicity Tables, 2001

18.12. Appendix 12 Adverse Events of Special Interest

Prior to RAP finalization this section will be updated to reflect the latest version of the Benlysta PSAP.

The primary source for rules governing identification, adjudication, and reporting of Adverse Events of Special Interest is the Program Safety Analysis Plan (PSAP). AESI are defined using preferred terms from the current version of MedDRA. The intent is to update these definitions semi-annually using the newest MedDRA version. Preferred terms used in the current version of MedDRA can be found in Attachment 3.

Malignant neoplasms

Malignant neoplasms are identified using the sub-SMQs of Malignant or unspecified tumours (20000091) and Malignancy related conditions (20000092) under the current version of MedDRA. The sub-SMQ of Malignant or unspecified tumours contains two further subcategories: "Malignant Tumours" and "Tumours of unspecified malignancy." Tumours of unspecified malignancy will be reviewed by GSK and identified as malignant or non-malignant for reporting.

Malignancies other than those in the "Tumours of unspecified malignancy" category will be categorized as hematologic, skin, or solid, based on a CMQ developed by the MAH (Attachment 3). In addition, the following customizations have been made:

- The term "Paraneoplastic glomerulonephritis" has been removed from the SMQ as it is a complication of malignancy.
- The term "Mismatch repair cancer syndrome" has not been assigned a tumor type.
- The term "Malignant meningioma metastatic" has been added as a solid tumor type.
- The term "Marginal zone lymphoma recurrent" has been added as a hematological tumor type.
- The term "Skin neoplasm bleeding" is added as a tumour of unspecified malignancy.

Non-melanoma skin cancer (NMSC) will be categorized using a CMQ developed by the Marketing authorization holder (MAH) (Attachment 3).

Post-infusion/injection systemic reactions

Post-infusion/injection systemic reactions will be identified using a customization of the Anaphylactic Reaction SMQ (20000021). This SMQ includes a broad list of preferred terms including symptoms of systemic injection/infusion reactions and hypersensitivity reactions and anaphylaxis. For the Anaphylactic Reaction query, 4 categories of preferred terms are considered, including a set of core anaphylactic terms (Category A), upper airway/respiratory terms (Category B), angioedema/urticaria/pruritus/flush terms (Category C), and cardiovascular/hypotension terms (Category D).

The customizations of the SMQ involve terms in Category A. Category A has been modified to include the following additional terms: "Infusion-related reaction," "Drug

hypersensitivity", "Hypersensitivity"," and "Urticarial vasculitis". Category B has been modified to include the following additional terms: "Oropharyngeal oedema" and "Pharyngeal oedema". Category C has been modified to include the following additional term: "Fixed eruption". GSK has also removed three terms that are not relevant for an analysis of hypersensitivity reactions to belimumab ("Anaphylactic transfusion reaction", "Dialysis membrane reaction", and "First use syndrome"). Anaphylactic transfusion reaction is an adverse event associated with a blood transfusion, not related to study medication. First use syndrome and dialysis membrane reaction are associated with adverse events related to kidney transplants and dialysis, not related to study medication

Algorithmic Search Criteria

The post-infusion/injection systemic reactions per Anaphylactic Reaction SMQ algorithmic search are defined as follows:

Subjects must have the following associated with the same infusion/injection:

a. at least 1 AE coding to a Category A preferred term *or*

b. 2 AEs, 1 coding to a Category B preferred term and the other coding to a Category C preferred term *or*

c. 2 AEs, 1 coding to a Category D preferred term and the other coding to either a Category B preferred term or to a Category C preferred term.

For the algorithmic search, if any event at a given infusion/injection meets the definition under criteria a, b or c, then all events in Categories A, B, C and D associated with that injection/infusion will be considered AESI.

For CSR reporting, all post-infusion/injection systemic reaction AESIs defined via narrow, broad, or algorithmic search, the AEs need to have occurred on the day of an infusion/injection or within 3 days after an infusion/injection. See Section 9.4.35 or the definition of the 3 day assessment window. GSK will review all serious events identified via the broad search occurring within 21 days after an infusion/injection, and adjudicate these events as post-infusion/injection systemic reactions or hypersensitivity reactions per the criteria in Section 9.4.35. Adverse events with partial or missing start dates will be included unless there is evidence through comparison of partial dates to suggest otherwise.

Sampson Criteria

Sampson et al define anaphylaxis as a severe, potentially fatal, systemic allergic reaction that occurs suddenly after contact with an allergy-causing substance. In addition, one of the following 3 criteria must be met: (1) acute onset of illness with involvement of skin or mucosal tissue, accompanied with either respiratory compromise, reduced blood pressure, or hypotension-related symptoms of end-organ dysfunction (2) reduced blood pressure associated with a known allergen or (3) two or more of the following that occur rapidly after exposure to an allergen: a) involvement of skin-mucosal tissue b) respiratory compromise c) reduced blood pressure d) persistent GI symptoms.

With the exception of GI symptoms, all symptoms required to assess anaphylaxis per Sampson criteria would be identified by Broad Anaphylaxis SMQ or the Anaphylactic Reaction SMQ algorithmic. Therefore, any events falling under the below criteria will be adjudicated by GSK prior to database release to determine if serious anaphylaxis per Sampson criteria is met.

Possible cases of serious anaphylaxis per Sampson criteria will be identified as follows:

a. Any Injection-related Reaction per Anaphylactic Reaction SMQ broad search SAE which occurs on the day of an injection.
b. Any AE or SAE in the "Gastrointestinal disorders" SOC that occurs on the day that criterion in a) above is met.
c. Any anaphylaxis and hypersensitivity reactions per Anaphylactic Reaction SMQ algorithmic search SAE which occurs on the day of an injection.

Infections

The infections of special interest are described below.

Opportunistic Infections

Opportunistic infections will be identified using a broad CMQ developed by the MAH (Attachment 3). Any events falling under these preferred terms will be adjudicated by GSK prior to database release to determine if criteria are met for an opportunistic infection, per the criteria in PSAP Appendix 7: Potential Opportunistic Infections.

Mycobacterium Tuberculosis

Tuberculosis events will be identified using a CMQ developed by the MAH (Attachment 3). Any events falling under these preferred terms will be adjudicated by GSK prior to database release to determine if criteria are met for an opportunistic infection PSAP (PSAP Appendix 7: Potential Opportunistic Infections).

Herpes Zoster

Herpes Zoster events will be identified using a CMQ developed by the MAH (Attachment 3). Additional manual adjudication by GSK prior to database release will identify events that are recurrent or disseminated (PSAP Appendix 7: Herpes Zoster).

Pneumonia

Pneumonia events will be identified using a CMQ developed by the MAH (Attachment 3). Pneumonia events will not be reported separately, but are being flagged in the event further evaluation is necessary.

Sepsis

Sepsis events will be identified using a CMQ developed by the MAH (Attachment 3).

Depression/suicide/self-injury

Depression (excluding suicide and self-injury)

Depression events will be identified using a CMQ including the preferred terms from the depression (excluding suicide and self injury) SMQ (20000035) plus additional terms added by the MAH (Attachment 3).

Suicide and Self-Injury

Suicide and self-injury events will be identified using the SMQ (20000035) preferred terms (Attachment 3).

Fatalities

All fatalities will be identified in the clinical database and subsequently adjudicated by the GSK SRT into a general category of death (PSAP Appendix 7: Fatalities).

Post-study fatalities that are captured in ARGUS prior to CSR approval, but are not captured in the clinical database, will be described within the CSR text but cannot be included in statistical post-text displays.

GSK SRT Adjudication of Adverse Events of Special Interest

Adverse events of special interest (AESI) are identified per the preferred terms (Attachment 3) and other criteria described in PSAP Appendix 6. The following AESI are adjudicated at the subject level by the GSK SRT during regular SRT meetings or during quarterly adjudication. The adjudication occurs prior to database release and is performed for reporting purposes, per the criteria described below.

Assignment of adjudication flags in the clinical database will occur ongoingly as part of the quarterly SRT review process. In addition, as part of individual study close-out procedures, the adjudications should be finalized as follows:

- Just preceding data base release (DBR), allowing time to send queries or update the eCRF/database as necessary prior to DBR.
- After DBR to provide final confirmation of adjudications and ensure there are no new AESI or relevant data changes to adjudicated events since the pre-DBR adjudication. This would be a requirement for declaring data a freeze (DBF).

Malignancies

All malignancies identified via the terms in Attachment 3 will be reviewed by GSK SRT. The classification of malignancies as solid tumor, hematological, and skin will be reviewed against the verbatim term to confirm an appropriate and accurate preferred term has been assigned, or to recommend follow-up with the investigator for additional specificity on the verbatim term. In addition, malignancies that are flagged more than once, e.g., based on a term for both a diagnostic procedure and a diagnosis, will be adjudicated as one event.

Tumors of unspecified malignancy, as identified per the terms in Attachment 3, will be reviewed clinically by the GSK SRT for reporting. In general, non-serious events in the tumours of unspecified malignancy with insufficient information will be categorized as not malignant. Serious adverse events with insufficient information will be categorized as either not malignant or malignant based on the type of tumor and likelihood the tumor type is malignant (e.g., thyroid nodules are common in SLE patients and are generally not malignant; tumor types with higher likelihood for malignancy would be assumed to be malignant).

Serious Hypersensitivity and Post-Infusion/Injection Systemic Reactions

Before the data base is released, GSK SRT will review all serious cases identified from the Broad Anaphylaxis SMQ as described in PSAP Appendix 6 and Attachment 3, applying clinical judgment to determine if the preferred terms are indicative of a hypersensitivity or infusion/injection reaction. Time to onset after an infusion/injection and details provided in the clinical narratives with respect to the nature and likely cause of the events are taken into consideration. Time to onset within 24 hours is generally applied to post infusion/injection reactions. The GSK SRT adjudicates serious hypersensitivity reactions into a category based primarily on time to onset: acute (onset ≤1 day), delayed acute (onset 2-3 days), or delayed, non-acute (onset 4-21 days). In addition to time to onset, descriptions of associated symptoms are taken into account for this categorization. In studies where subjects are receiving weekly injections, any delayed, non-acute reactions will typically occur in the interval 4-7 days later, but may occur up to 21 days later following a missed injection or after the last injection.

In addition, possible cases of serious anaphylaxis per Sampson criteria will be identified per the criteria in PSAP Appendix 6. Any events falling under these criteria will be adjudicated by GSK prior to database release to determine if serious anaphylaxis per Sampson criteria is met.

Potential Opportunistic Infections

Opportunistic infections (OIs) will be identified using a list of preferred terms (Attachment 3), designed to cast a wide net for events potentially indicative of an opportunistic infection. Any identified events will be adjudicated by the GSK SRT prior to database release to determine if criteria are met for an opportunistic infection. Targeted follow-up is sought for events with insufficient information. In general, potential OIs that are non-serious with insufficient information to adjudicate will be considered non-opportunistic. Potential OI SAEs with insufficient information to adjudicate will be

considered opportunistic. See below for a list of agreed upon pathogens and infections considered to be opportunistic for the purpose of adjudication.

Pathogens and Infections Considered Opportunistic:

- Acinetobacter infection
- Aspergillosis
- Blastomycosis, extrapulmonary
- Candidiasis of esophagus, bronchi, trachea or lungs
- Coccidioidomycosis, disseminated or extrapulmonary
- Cryptococcosis, extrapulmonary
- Cryptosporidiosis infection, chronic intestinal (greater than 1 month duration)
- CMV disease other than liver, spleen, or nodes
- Herpes simplex – bronchitis, pneumonitis, or esophagitis
- Herpes Zoster (adjudication details are below)
- Histoplasmosis disseminated or extrapulmonary
- Human polyomavirus infection
- Isosporiasis, chronic intestinal (greater than one month duration)
- Listeriosis
- Mycobacterium avium complex or M. Kansasii, disseminated or extrapulmonary
- Nocardiosis
- Other non-tuberculous mycobacterium (NTM) infections (other species or unidentified species), disseminated or extrapulmonary*
- Polyomavirus (JC virus or BK virus) associated nephropathy (including PML)
- Pneumocystis jiroveci infection
- Toxoplasmosis of brain

* Extra pulmonary NTM infections are generally considered an OI unless the affected extra pulmonary area followed a wound or trauma.

In addition, GSK SRT will review all SAEs, including those coded under preferred terms not listed in Attachment 3, and utilizing the supplemental/narrative information, will adjudicate the SAEs as OI if warranted based on medical judgment.

Other Infections of Interest but not generally considered opportunistic:

- Mycobacterium tuberculosis (adjudication details are below)

Herpes Zoster

Herpes Zoster events will be identified per terms in Attachment 3. Adjudication by GSK prior to database release will identify events that are recurrent or disseminated. Herpes Zoster is considered disseminated if there is involvement of other organs other than the skin or if skin lesions (1) cross the midline of the body or (2) are in non-adjacent dermatomes or (3) are located in more than three adjacent dermatomes. Herpes zoster is considered an opportunistic infection if it is adjudicated as recurrent or disseminated. However, there may be some uncommon occurrences of a herpes zoster case that is adjudicated as an OI but is neither recurrent nor disseminated.

Mycobacterium Tuberculosis

Tuberculosis (TB) cases are reviewed clinically by the GSK SRT before data base release to determine if a case is an OI. The following principles are applied: Pulmonary TB in an endemic area is not considered an OI. Pulmonary TB in a non-endemic area would be considered an OI unless the subject had close contact with a person infected with TB. Extra pulmonary TB is generally considered an OI unless the affected extra pulmonary area followed a wound or trauma.

Suicide/self-injury

Suicide and self-injury SAEs will be identified using the preferred terms identified in Attachment 3 and subsequently adjudicated into the following categories:

Adjudicated Category
Suicidal Behavior
Completed Suicide
Suicidal Ideation
Self-Injurious Behavior without Suicidal Intent

In addition, GSK SRT will review all SAEs, including those coded under preferred terms not listed in Attachment 3, and utilizing the supplemental/narrative information, will adjudicate the SAEs as suicide/self-injury if warranted based on medical judgment.

Fatalities

All fatalities will be identified in the clinical database and subsequently adjudicated by the GSK SRT into a general category of death.

All fatalities will be adjudicated into one of the following categories:

Adjudicated Category of Death
SLE-Related
Infectious
Vascular
Gastrointestinal
Respiratory

Adjudicated Category of Death
Malignancy
Hypersensitivity
Suicide
Surgical Complication
Unknown
Hematologic
Trauma

Additional ‘categories of death’ may be added in the future should a fatality not clearly fit into one of the ‘categories’ listed above. The ‘categories’ will not change unless agreed upon by the GSK SRT.

18.13. Appendix 13 B Cells

B Cell Panel (BIMETHCD)	Biomarker Category Code [BICATCD]	Biomarker Category [BICAT]	Lab Test Code (LBTESTCD)	Lab Test Code (LBTEST)	Units of Measurement (BIORRESU/ LBORRESU)
FLWTBNK	CD19	CD19	CD19LY	CD19/Lymphocytes	%
FLWTBNK	CD19	CD19	CD19	CD19	GI/L
FLWPLSM	CD19	CD19	CD19E	CD19 Number of Events	EVENTS
FLWTRANS	CD19	CD19	CD19E	CD19 Number of Events	EVENTS
FLWPLSM	CD20	CD20	CD20CD19	CD20/CD19+	%
FLWPLSM	CD20	CD20	CD20	CD20 Number of Events	EVENTS
FLWPLSM	CD20	CD20	CD20E	CD20	GI/L
FLWPLSM	CDX136	CD20+ CD27-	CDX13619	CD20+ CD27-/CD19+	%
FLWPLSM	CDX136	CD20+ CD27-	CDX136E	CD20+ CD27- Number of Events	EVENTS
FLWPLSM	CDX136	CD20+ CD27-	CDX136	CD20+ CD27-	GI/L
FLWPLSM	CDX137	CD20+ CD27+	CDX13719	CD20+ CD27+/CD19+	%
FLWPLSM	CDX137	CD20+ CD27+	CDX137E	CD20+ CD27+ Number of Events	EVENTS
FLWPLSM	CDX137	CD20+ CD27+	CDX137	CD20+ CD27+	GI/L
FLWPLSM	CDX141	CD20+ CD69+	CDX14119	CD20+ CD69+/CD19+	%
FLWPLSM	CDX141	CD20+ CD69+	CDX141	CD20+ CD69+	GI/L
FLWPLSM	CDX141	CD19+CD20+ CD69+	CDX141E	CD20+CD69+_Number of events	EVENTS
FLWPLSM	CDX143	CD20- CD138+	CDX14319	CD20- CD138+/CD19+	%
FLWPLSM	CDX143	CD20- CD138+	CDX143E	CD20- CD138+ Number of Events	EVENTS
FLWPLSM	CDX143	CD20- CD138+	CDX143	CD20- CD138+	GI/L
FLWPLSM	CDX145	CD20+ CD138+	CDX14519	CD20+ CD138+/CD19+	%
FLWPLSM	CDX145	CD20+ CD138+	CDX145E	CD20+ CD138+ Number of Events	EVENTS
FLWPLSM	CDX145	CD20+ CD138+	CDX145	CD20+ CD138+	GI/L
FLWPLSM	CDX154	CD27+b CD20-	CDX15419	CD27+b CD20-/CD19+	%
FLWPLSM	CDX154	CD27+b CD20-	CDX154E	CD27+b CD20- Number of Events	EVENTS
FLWPLSM	CDX154	CD27+b	CDX154	CD27+b CD20-	GI/L

B Cell Panel (BIMETHCD)	Biomarker Category Code [BICATCD]	Biomarker Category [BICAT]	Lab Test Code (LBTESTCD)	Lab Test Code (LBTEST)	Units of Measurement (BIORRESU/ LBORRESU)
		CD20-			
FLWPLSM	CDX156	CD27+CD38+ CD19+	CDX15619	CD27+CD38+CD19+/CD19+	%
FLWPLSM	CDX156	CD27+CD38+ CD19+	CDX156E	CD27+CD38+CD19+ Number of Events	EVENTS
FLWPLSM	CDX156	CD27+CD38+ CD19+	CDX156	CD27+CD38+CD19+	GI/L
FLWTRANS	CDX199	CD19+ CD24b+ CD38b+ CD27	CDX19919	CD19+ CD24b+ CD38b+ CD27/CD19+	%
FLWTRANS	CDX199	CD19+ CD24b+ CD38b+ CD27	CDX199	CD19+ CD24b+ CD38b+ CD27	CELLS/CUMM
FLWTRANS	CDX199	CD19+ CD24b+ CD38b+ CD27	CDX199E	CD19+ CD24b+ CD38b+ CD27 Number of Events	EVENTS
FLWTRANS	CDX200	CD19+CD24b +CD38b+ CD27-IgD+ CD10+	CDX20019	CD19+CD24b+CD38b+CD27-IgD+CD10+/CD19+	%
FLWTRANS	CDX200	CD19+CD24b +CD38b+ CD27-IgD+ CD10+	CDX200	CD19+CD24b+CD38b+CD27-IgD+CD10+	CELLS/CUMM
FLWTRANS	CDX200	CD19+CD24b +CD38b+ CD27-IgD+ CD10+	CDX200E	CD19+24b+38b+27-IgD+10+ Num of Events	EVENTS

B cell subsets to be reported:

Lab Test Code (LBTESTCD)	Lab Test (LBTEST)	Units of Measurement[1] (LBORRESU)	Display Label for B cell
Common B cells			
CD19	CD19	GI/L	CD19 (/uL)
CD20	CD20	GI/L	CD20 (/uL)
CDX136	CD20+ CD27-	GI/L	Naive CD19+CD20+CD27- (/uL)
CDX13619	CD20+ CD27-/CD19+	%	Naive CD19+CD20+CD27- (%CD19)
CDX137	CD20+ CD27+	GI/L	Memory CD19+CD20+CD27+ (/uL)
CDX13719	CD20+ CD27+/CD19+	%	Memory CD19+CD20+CD27+ (%CD19)
Rare B cells[2]			
CDX141N	CD20+ CD69+	GI/L	Activated CD19+CD20+CD69+ Normalised (COUNT/mL)
CDX143N	CD20- CD138+	GI/L	Plasma CD19+CD20-CD138+ Normalised (COUNT/mL)
CDX145N	CD20+ CD138+	GI/L	Plasmacytoid CD19+CD20+CD138+ Normalised (COUNT/mL)
CDX154N	CD27+b CD20-	GI/L	Short-lived Plasma CD19+CD20-CD27b+ Normalised (COUNT/mL)
CDX156N	CD27+CD38+CD19+	GI/L	SLE Subset CD19+CD38b+CD27b+Lymph Normalised (COUNT/mL)
CDX199N	CD19+ CD24b+ CD38b+ CD27-	GI/L	Transitional CD19+CD24b+CD38b+CD27- Normalised (COUNT/mL)

[1]GI/L=10^9/L

[2]The lab test code for the new record containing the normalized value will be the same as the corresponding absolute B cell concentration record prior to normalization, suffixed with N. The display label corresponds to the normalized value that is to be reported in the displays.

18.14. Appendix 14 Headline Results

The following tables will be included as headline results:

Table numbers will be populated when numbering is finalized.

1.02	Populations (Double-Blind Phase) (randomized)
1.03	Summary of Subject Disposition for the Double-Blind Subject Conclusion Record (mITT)
1.08	Summary of Demographic Characteristics (Double-blind phase) (mITT)
1.13	Baseline Disease Activity (Double-blind phase) (mITT)
2.01	SRI-S2K Response at Week 52 (Double-blind phase) (mITT)
2.02	Logistic Regression Analysis of SRI-S2K Response at Week 52 (Double-blind phase) (mITT)
2.03	SRI-S2K Response at Week 52 and the 3 Components (Double-blind phase) (mITT)
2.09	SRI-S2K Response by Visit (Double-blind phase) (mITT)
2.21	SRI response by visit using SELENA SLEDAI (Double-blind phase) (mITT)
2.25	Time to First Severe SFI Flare over 52 Weeks (Double-blind phase) (mITT)
2.27	Prednisone Reduction by >= 25% from Baseline to <= 7.5 mg/day during Week 40 through Week 52 (DO/TF=NR) (Double-blind phase) (mITT)
3.15	Serious Adverse Events by SOC and PT (Double-Blind phase) (Safety)
3.19	Adverse Events Resulting in Study Agent Discontinuation (Double-Blind phase) (Safety)
3.23	Adverse Events of Special Interest by Category (Double-Blind phase) (Safety)
3.59	Laboratory Toxicity Grade Worsening of at Least 2 Grades from Baseline: Immunoglobulins (Double-Blind phase) (Safety)

19. ATTACHMENTS

19.1. Attachment 1: Table of Contents for Data Display Specifications

See separate TFL document.

19.2. Attachment 2: Data Display Specifications

See separate TFL document.

19.3. Attachment 3: AESI Preferred Term Definitions under MedDRA v20.1

The AESI definitions are stored at the project level for GSK1550188.