Official Title of Study: A Phase 2, Multi-Center, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Safety and Efficacy of BMS-931699 vs Placebo on a Background of Limited Standard of Care in the Treatment of Subjects with Active Systemic Lupus Erythematosus

NCT Number: NCT02265744

Document Date: June 28, 2016

STATISTICAL ANALYSIS PLAN

(SHORT-TERM PERIOD)

A PHASE 2, MULTI-CENTER, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED STUDY TO EVALUATE THE SAFETY AND EFFICACY OF BMS-931699 VS. PLACEBO ON A BACKGROUND OF LIMITED STANDARD OF CARE IN THE TREATMENT OF SUBJECTS WITH ACTIVE SYSTEMIC LUPUS ERYTHEMATOSUS

PROTOCOL IM128-027

VERSION 2.0

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 [redacted]

[redacted]

Research Hypothesis:

BMS-931699 (Lulizumab) will have greater clinical efficacy compared to placebo on the British Isles Lupus Assessment Group (BILAG)-based Composite Lupus Assessment (BICLA) score, when added to stable background standard of care in subjects with active manifestations of systemic lupus erythematosus (SLE) including, at a minimum, elevated antinuclear antibodies, arthritis, and/or cutaneous manifestations satisfying the BILAG "A" or "B" criteria.

Schedule of Analyses:

The following analyses will be performed:

- An interim analysis for the safety/RO at Day 29 of Part 1 when at least 6 patients per treatment arm have either reached Study Day 29 or discontinued (refer to the details in Section 7.5.5);
- An interim analysis for futility and dose adaptation will be performed when approximately 30 patients per treatment arm (including the patients from part 1) have completed at least 84 days of treatment period or discontinued the treatment;
- A final analysis will be performed after the last treated subject in the study reaches the end of the short-term period.

2 STUDY DESCRIPTION

2.1 Study Design

This is a Phase 2, parallel-arm, randomized, double-blinded, multicenter, international study, with an adaptive design.

The study will comprise a short-term period (Part 1 and Part 2), and a long-term extension (LTE) period. This SAP covers the short-term period of the study.

Part 1 of the short-term period will focus on assessing safety and RO and will be limited to a maximum of 50 patients (approximately 6-10 patients/arm) who will be included in the Safety/RO interim analysis (IA). These patients will continue study treatment beyond the IA for Safety/RO and be treated for up to 24 weeks and be followed for 6 additional weeks after treatment is completed. Treatment could be shorter if IA for Safety/RO analysis indicates one or more arms should be dropped.

Part 2 of the short-term period will start once cumulative safety profile is considered acceptable and RO data for 6-10 subjects per cohort dosed for >28 days is available and the interim analysis from Part 1 is completed. Approximately 300 subjects will be randomized into this part of the study (number of patient to be randomized into Part 2 may increase based on the results of the futility interim analysis).

All subjects will undergo screening evaluations to determine eligibility and allow down titration of prednisone (or prednisone equivalent) prior to administration of study medication. The screening period will have a duration of approximately 28 days (+/- 2 days). All eligibility criteria must be met prior to Day 1. In the event of technical difficulties leading inconclusive assessments that would go beyond the allowed window, the study medical monitor must be consulted prior to conducting Day 1. On Day 1, eligibility of subjects will be confirmed and eligible subjects will receive their first dose of study medication.

Dose may be adapted based on results of the planned interim analyses from Part 1 and Part 2, as follows:

- Safety and RO interim analysis: safety and RO analysis will be performed when approximately 6 to 10 subjects per arm have reached Study Day 29. Based on the results of this analysis, dose arms(s) may be adjusted or may be dropped.
- Interim analysis for futility and dose adaptation on BICLA response, SRI response, ACR28 and some SLE biomarkers (such as auto-antibodies, complement levels, etc.) with possible exploratory exposure response analysis, will be performed when approximately 30 subjects per arm have reached Study Day 85 (12 weeks), including subjects from Part 1 and Part 2, of treatment or have discontinued prior to reaching study Day 85~~)~~. Analysis will be performed by an unblinded Sponsor team, while maintaining blind at the site and subject level. Based on the results, the dose levels and sample size may be modified.
- Ongoing assessment of safety will be performed by an independent Data Monitoring Committee (DMC) and an internal unblinded safety monitoring team. Both entities may make recommendations to the Sponsor regarding conduct of study and dose adjustment based on safety observations.

Following interim analyses and DMC reviews, dose arm(s) may be modified based on the results, as follows:

- If one or more arms are added, future randomizations will include the new arm(s).
- If one or more arms are dropped for safety reasons, randomization into that arm will be terminated and all subjects currently randomized to that or those arm(s) will be discontinued from receiving study medication.
- If one or more arms are dropped due to inadequate RO, there will be no new randomizations into that dose arm; subjects currently randomized may continue at the Investigator discretion.

The **long-term extension (LTE) period** is optional and will include eligible subjects who have completed Day 169 (week 24) of treatment and consent to participate. This period of the study will remain blinded but will no longer have a placebo arm. Eligible subjects will remain on their originally-assigned dose arm, unless they were on the placebo arm during the short-term period. Placebo-arm subjects will be re-randomized into one of the existing active arms at Day 169 (24 weeks). Re-randomization of the placebo subjects will be done by IVRS and only the unblinded pharmacist/drug preparer will know the new randomization arm. The LTE will remain blinded to the study team and study personnel.

The long-term extension period will not have interim analyses for dose adaptation. However, the BMS unblinded safety committee will continue to review the safety data and discuss it with the DMC (also unblinded), as needed. During their evaluation of adverse events, if a significant safety concern is identified, these two committees may propose that one or more dose arms are

dropped. If one or more dose arms are dropped for safety reasons, all subjects currently receiving that or those dose(s) will be discontinued from receiving study medication.

A separate document (LTE SAP) describes the analysis plan for the long-term extension period.

The study design schematics are presented in Figure 2.1-1, Figure 2.1-2, Figure 2.1-3, and Figure 2.1-4 below.

Figure 2.1-1: Overall Study Design

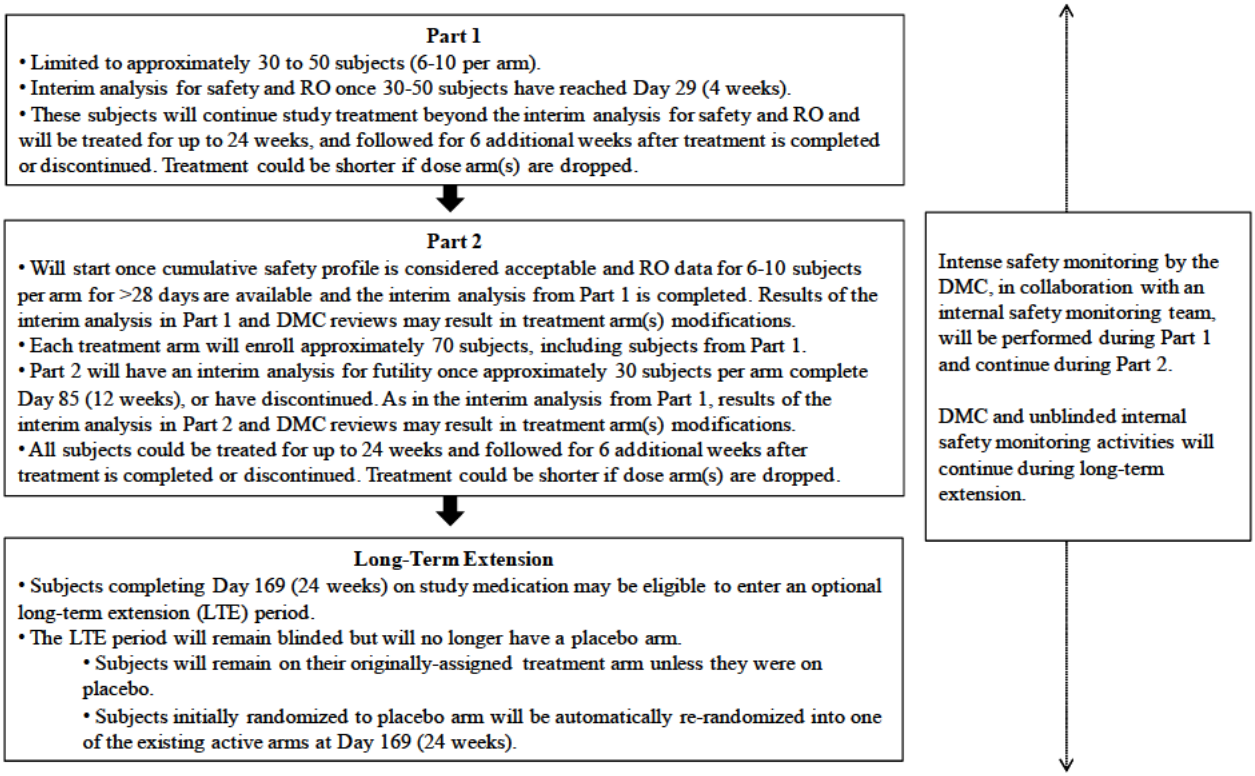
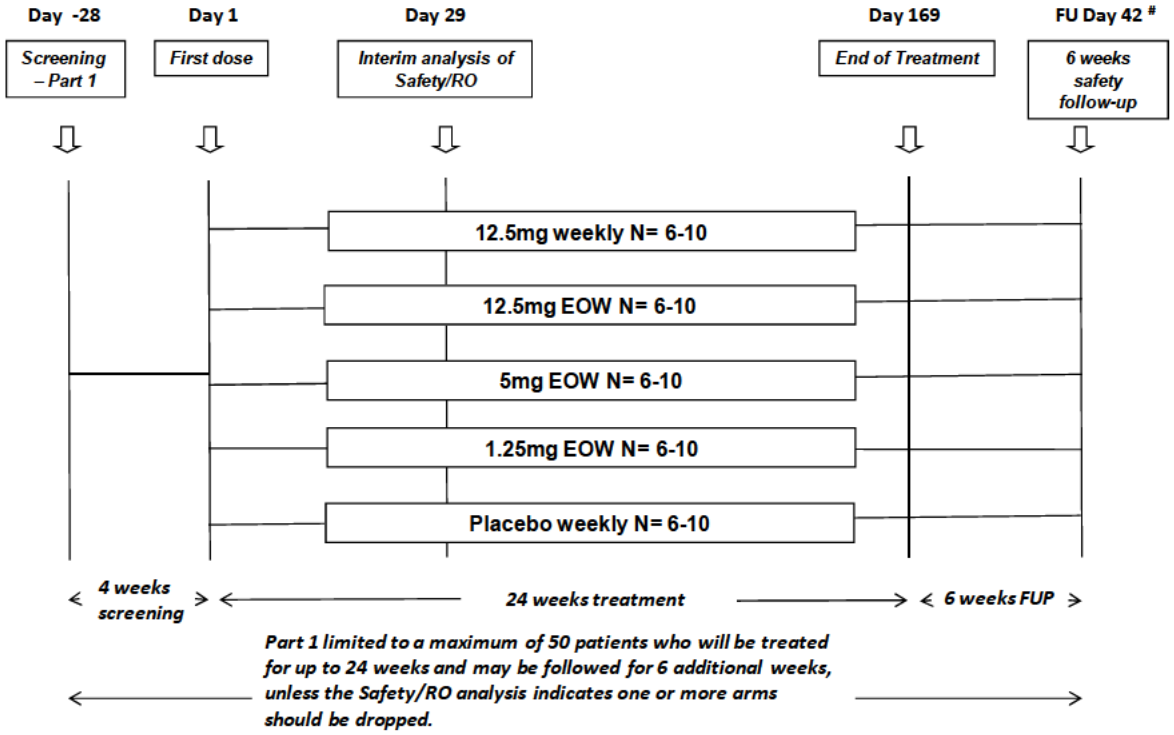
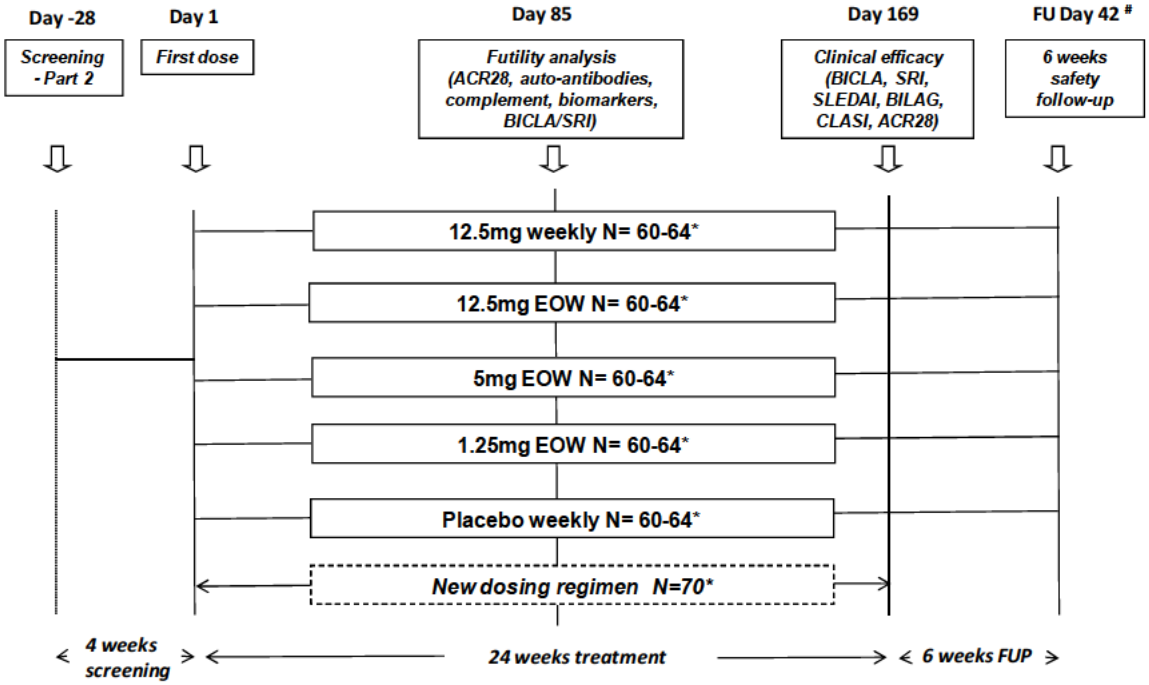


Figure 2.1-2: Study Design Schematic - Part 1



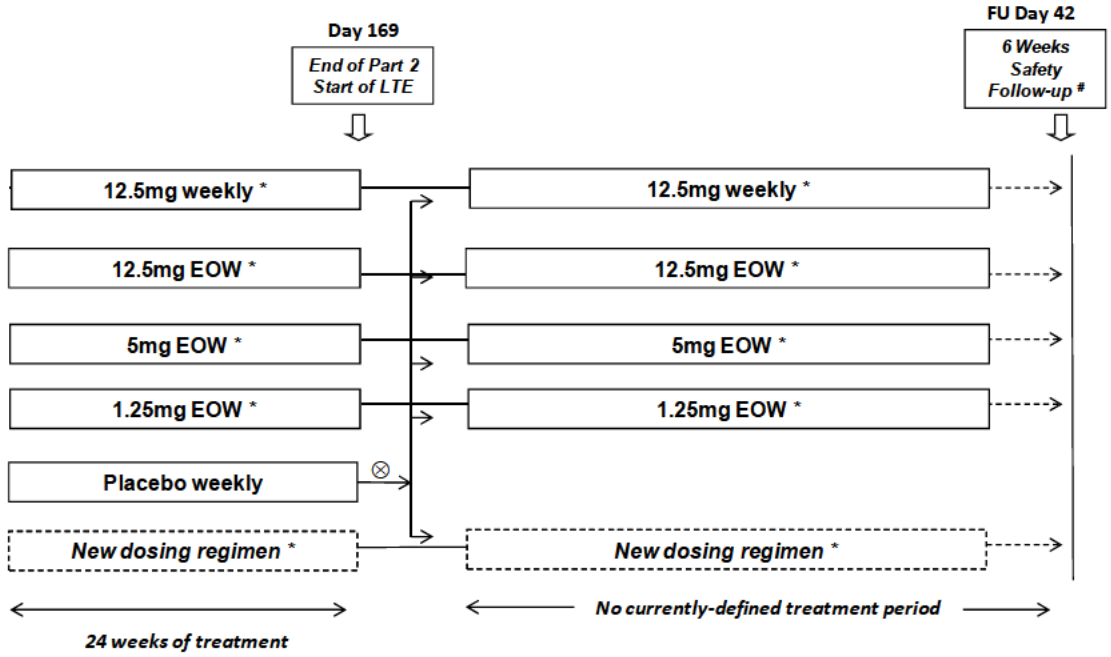
Will not apply to subjects entering LTE.

Figure 2.1-3: Study Design Schematic - Part 2



* Treatment arm may or may not be included based on Part 1 Day 29 Interim Analysis.
# Will not apply to subjects entering LTE.

Figure 2.1-4: Long-Term Extension Schematic



** Treatment arm may or may not be included based on Interim Analyses.*
⊗ Subjects in the placebo arm will be re-randomized to existing active treatment arms at Day 169.
 # FU Day 42 to occur after last dose of study medication, whether last dose is in short term or long term periods.

Physical examinations, vital sign measurements, 12-lead electrocardiograms (ECG), and clinical laboratory evaluations will be performed at selected times throughout the study. Subjects will be closely monitored for adverse events throughout the study. Blood will be collected for pharmacokinetic (PK) and pharmacodynamic measurements as well as for safety monitoring.

The approximate duration of the short-term period (Parts 1 and 2) of the study is 238 days (34 weeks), including: 28 days (4 weeks) of screening, 168 days (24 weeks) of treatment, and 42 days (6 weeks) of safety follow-up. If subject is eligible and opts to continue into LTE, the 42-day follow-up visit will not be performed after the short-term period is completed and subject will enter LTE directly. If subject opts not to enter LTE then a follow-up visit will be completed 42 days after end of treatment. At the time of writing there is no defined end date to the long-term extension period, however, the LTE provision may be further adjusted based on results from the ongoing lulizumab development program. Subjects discontinuing treatment during the LTE period will complete the follow-up visit approximately 6 weeks after receiving their last dose of study medication.

Subjects randomized in the short-term period (either Part 1 or Part 2) will be treated for up to 24 weeks, will have the same procedures performed and will follow the same visit schedule. All subjects will account for the approximately 350 subjects planned to be randomized in this study.

Approximately 125 sites will participate in this study, which will be conducted globally and should include, but will not be limited to, participation of countries located in North America and Europe.

Subjects who discontinue from the study will have one follow-up visit, approximately 42 days (6 weeks) after receiving their last dose of study medication, to perform safety assessments. During this period, when subjects are no longer receiving study drug, it is recommended that they are not treated with other biologic therapy, due to the compound's half-life. However, this decision remains at the investigator's discretion. If the study drug becomes commercially available, and if the subject chooses to receive treatment with the commercial product, then the subsequent post-dose follow-up visits are not required.

For more details, refer to the study protocol.

2.2 Treatment Assignment

After completion of all screening evaluations and confirmation of eligibility, subjects will be randomized on Day 1 to one of the following dosing arms in a 1:1:1:1:1 ratio in Part 1 of the study design: BMS-931699 at 12.5mg weekly sub-cutaneous (SC) injection, BMS-931699 at 12.5mg every other week (EOW) SC injection, BMS-931699 at 5mg EOW SC injection, BMS-931699 at 1.25mg EOW SC injection, or Placebo SC injection. The randomization ratio and the treatment arms may be modified based on the results from the two interim analyses defined in Section 7.6.

In this study, subjects will be randomized to receive one of the treatments according to a computer-generated randomization scheme prepared by a Randomization Coordinator within the Drug Supply Management Department of BMS Research and Development.

During the Screening visit, the investigative site will call into the enrollment option of the Interactive Voice Response System (IVRS) designated by BMS for assignment of a 5-digit subject number that will be unique across all sites. Enrolled subjects, including those not dosed, will be assigned sequential subject numbers starting with 00001, 00002, 00003.... 00010. The patient identification number (PID) will ultimately be comprised of the site number and subject number. For example, the first subject screened (i.e., enrolled) at site number 1, will have a PID of 0001 00001.Once it is determined that the subject meets the eligibility criteria following the Screening visit, the investigative site will call the IVRS to randomize the subject into the treatment period.

Every randomized subject will be required to come to the clinic/research center weekly to be dosed. This will ensure double-blind is maintained despite the variability of regimens. Subjects randomized to weekly subcutaneous injections of either placebo or BMS-931699 will be dosed weekly as per schedule and subjects randomized to one of the every other week arm will be alternating between receiving a subcutaneous injection of BMS-931699 one week and one of placebo the following week.

2.3 Blinding and Unblinding

2.3.1 Blinding at the Study Sites

The subjects and clinical assessor(s) will not be aware of which treatment is being administered to the subjects enrolled in the study. The pharmacist (or qualified drug preparation person) will be unblinded to study medication (BMS-931699 or placebo).

The pharmacist (or qualified drug preparation person) will know the randomization assignments and prepare the appropriate dose of active BMS-931699 or placebo accordingly. The prepared drug must be supplied to study personnel in a manner such that neither study personnel nor subjects will be aware of whether they receive active drug or placebo.

2.3.2 Unblinding of the Study Sites

Blinding of treatment assignment is critical to the integrity of this clinical study. However, in the event of a medical emergency or pregnancy in an individual subject in which knowledge of the investigational product is critical to the subject's management, the blind for that subject may be broken by the investigator. The subject's safety takes priority over any other considerations in determining if a treatment assignment should be unblinded.

Before breaking the blind of an individual subject's treatment, the investigator should determine that the unblinded information is necessary, i.e., that it will alter the subject's immediate management. In many cases, particularly when the emergency is clearly not related to the investigational product, the problem may be properly managed by assuming that the subject is receiving active product. It is highly desirable that the decision to unblind treatment assignment be discussed with the Medical Monitor, but the investigatory always has ultimate authority for the decision to unblind. The Principal Investigator should only call in for emergency unblinding AFTER the decision to discontinue the subject has been made.

For this study, the method of unblinding for emergency purposes is IVRS.

For information on how to unblind in an emergency, consult the IVRS manual.

In cases of accidental unblinding, contact the Medical Monitor and ensure every attempt is made to preserve the blind.

Any request to unblind a subject for non-emergency purposes should be discussed with the Medical Monitor.

In case of an emergency, the Investigator(s) has unrestricted access to randomization information via the Interactive Voice Response System (IVRS) and is capable of breaking the blind through the IVRS system without prior approval from sponsor. Following the unblinding the Investigator shall notify the Medical Monitor and/or Study Director.

2.3.3 Other Unblinding Information

For analysis of pharmacokinetic and anti-BMS-931699 samples, the Bioanalytical Sciences Group at BMS, and/or the designated vendor performing the analysis will be unblinded to the randomized treatment assignments in order to minimize unnecessary analysis of samples from control group subjects.

In collaboration with the DMC, an unblinded internal safety monitoring team, which is independent of the study team, will review the safety data throughout the study and will provide recommendations to the blinded study team. Therefore, the unblinded internal safety monitoring team can be unblinded. The unblinded internal safety monitoring team will have members from the Global Pharmacovigilance and Epidemiology group (GPV&E) within BMS.

The following BMS personnel will form the unblinded team (Independent of the blinded study team): one or more GPV&E physicians, one or more PK scientists, pharmacometrics analysts, one or more statisticians and programmers.

2.4 Protocol Amendments

Table 2.4-1: Protocol Amendments

Amendment No.	Amendment Date	Main Purpose of Amendment
1	20-AUG-2014	Pharmacogenetics amendment
2	24-SEP-2014	Brazil specific amendment
3	16-DEC-2014	Japan specific amendment
4	11-FEB-2015	Argentina specific amendment
5	16-MAR-2015	Japan specific amendment to add Exclusion Criteria regarding Laboratory Test of Hepatitis-B and Allergies
6	28-APR-2015	Revised Protocol no.1 (adding LTE, allowing MMF, edits, adding clarity)
7	29-JUN-2015	Canada specific amendment to limit LTE to 1yr (until further data are available)
8	01-Oct-15	Removing RNA testing (Site 152 (Japan) only)
9	22-Apr-16	To add intensive PK blood sample collections and also add sparse receptor occupancy samples in the PK substudy

3 OBJECTIVES

3.1 Primary Objective

To compare the proportion of patients who achieve BICLA response (BICLA response rate) at Day 169 following 24-week treatment with BMS-931699 or placebo administered on a stable background therapy.

3.2 Secondary Objectives

To assess:

- The safety and tolerability of treatment with BMS-931699 in patients with active SLE
- The proportion of patients who meet response criteria for the SLE Responder Index (4) [SRI (4)], SRI(5), and SRI(6)] at Day 169 following 24-week treatment with BMS-931699 or placebo administered on a stable background therapy.
- The proportion of patients with SLE Responder Index [SRI(4), SRI(5), and SRI(6)] at Day 85 following 12-week treatment with BMS-931699 or placebo administered on a stable background therapy.
- The proportion of patients with a BICLA response at Day 85 following 12-week treatment with BMS-931699 or placebo administered on a stable background therapy.
- The improvement in the extent of cutaneous and mucous membrane activity, as measured by the Cutaneous Lupus Erythematosus Disease Area and Severity Index (CLASI) at Day 85 and Day 169
- Joint pain, tenderness and swellings as measured by the by American College of Rheumatology (ACR)28 at Day 85 and Day 169
- The other indices of SLE activity measured at Days 29, 57, 85, 113, 141, and 169 including:
  - The change from baseline in overall BILAG-2004 score. A major response is defined as described in Yee et. al. 2010 (A=12; B=8; C=1; D/E=0).[2]
  - The change in overall Systemic Lupus Erythematosus Disease Activity Index 2K score (SLEDAI 2K)
  - The change from baseline in Physician's Global Assessment of disease activity (MDGA) score
- The systemic exposure of BMS-931699 in patients with SLE
- The cumulative corticosteroids use and immunosuppressant use over time
- The immunogenicity of BMS-931699
- Serum biomarkers C3, C4, anti-double-stranded deoxyribonucleic acid (anti-dsDNA), anti-nuclear antibody (ANA) and other autoantibodies at Day 85 and Day 169

- The pharmacodynamics (PD) of BMS-931699, including assessments potentially associated with target engagement (including receptor occupancy [RO]).

[redacted]

4 ENDPOINTS

4.1 Primary Efficacy Endpoint(s)

The primary efficacy endpoint is the proportion of subjects who achieve BICLA response at Day 169.

4.2 Secondary Efficacy Endpoint(s)

- Proportion of subjects who meet response criteria for the SLE Responder Index(4) [SRI(4)], SRI(5), SRI(6), and SRI(8) at Day 169
- Proportion of subjects who meet response criteria for the SLE Responder Index(4) [SRI(4)], SRI(5), SRI(6), and SRI(8) at Day 85
- Proportion of subjects with BICLA response at Day 85
- Change from baseline in CLASI score at Day 85 and Day 169

- Percentage of subjects with an improvement of at least 4 or an improvement of at least 50% from baseline in their CLASI score at Day 85 and Day 169
- Proportion of subjects with an improvement of at least 4 or an improvement of at least 20% from baseline in their CLASI score at Day 85 and Day 169
- Change from baseline in arthritis, as assessed by American College of Rheumatology (ACR) 28 joint counts of tender, swollen, and as well as active joints, i.e. those that are both tender and swollen at Day 85 and at Day 169
- The change from baseline in the following other indices SLE activity over time:
  - The overall BILAG-2004 score
  - The overall SLEDAI-2K score
  - The overall MDGA (Physician's Global Assessment of Disease Activity) score
- Change from baseline in SLICC/ACR Damage Index
- The cumulative corticosteroid and immunosuppressant use over time

4.3 Safety Endpoints

- Incidence and severity of all Adverse Events (AEs), Serious AEs (SAE), and Events of Special Interest (EOSI) pre-established in Section 6.7.1 of the protocol (including but not limited to Infections, Autoimmunity, Malignancies and Injection site reactions)
- Incidence and severity of clinically significant changes in vital signs
- Incidence and severity of clinically significant electrocardiogram (ECG) abnormalities
- Incidence and severity of clinically significant abnormalities in general laboratory tests

4.4 Pharmacokinetic Endpoints

Ctrough: Trough level serum concentration of BMS-931699 at time points specified in Section 5.5 of the protocol.

[redacted]

4.6 Immunogenicity

Proportion of subjects with BMS-931699 induced antibody response over the time period for which samples are collected (specified in Table 5.5.3 of the protocol).

5 SAMPLE SIZE AND POWER

The sample size calculation is based on power to compare the proportion of subjects who achieve BICLA response (the BICLA response rate) at Day 169 between each BMS-931699 dose arm and the placebo arm. With a one-sided continuity corrected chi-squared test at significance level 0.05, data from 70 treated subjects per arm will provide at least 89% power to detect a 25% increase in the BICLA response rate in the active treatment arm (i.e.. 50% response rate in BMS active arm) compared to placebo and provide 74% power to detect 20% increase in the BICLA response rate in the active treatment arm (i.e., 45% response rate in BMS-931699 arm) compared to placebo, assuming that the placebo arm has 25% response rate which is reasonable based on the EMBLEM phase 2 trial result[3]. No adjustment will be made for multiplicity.

The primary efficacy analysis will be conducted on all randomized patients who received at least one dose of the study drug. In order to get 70 treated subjects per arm, an adequate numbers of subjects need to meet the inclusion/exclusion criteria at screening.

If the interim analysis for futility and dose adaptation suggests an additional dose arm at a lower dose level needs to be added, then up to 40 additional subjects may be randomized.

6 STUDY PERIODS, TREATMENT REGIMENS AND POPULATIONS FOR ANALYSES

6.1 Study Periods

The following periods are defined in this SAP for the purpose of analyses and reporting.

Pre-treatment period: It covers the time period which starts from the day of enrollment and lasts until the initiation of the randomized double-blind treatment period.

Short-term period (ST): It starts at the time of the first dose of blinded treatment (BMS-931699 or placebo, regardless if it was received during Part 1 or Part 2 of study design) and continues up to the start of the first long-term extension dose (Day 169). For subjects who discontinue during the short-term period or do not enter the long-term extension period, it continues up to 42 days post last dose.

Long-term extension period (LTE): It starts at the first dose date of long-term extension (Day 169) and continues up to 42 days post the last study dosing date (LTE dose). Cutoffs may be applied for interim locks.

6.2 Treatment Regimens

All subjects will receive BMS-931699 or Placebo every week (double-dummy).

- BMS-931699 at 12.5mg weekly sub-cutaneous (SC) injection
- BMS-931699 at 12.5mg every other week (EOW) SC injection
- BMS-931699 at 5mg EOW SC injection
- BMS-931699 at 1.25mg EOW SC injection
- Placebo SC injection (every week)

Subjects randomized to weekly subcutaneous injections of either placebo or BMS-931699 will be dosed weekly as per schedule and subjects randomized to one of the every other week arm will be alternating between receiving a subcutaneous injection of BMS-931699 one week and one of placebo the following week.

There may be an instance of new treatment group as per study design. Also one or more treatment arms (other than Placebo) may be dropped based on interim analyses.

For more details, see study design in IM128-027 Protocol.

6.3 Populations for Analyses

- All Enrolled Subjects: All subjects who signed an informed consent.
- All Randomized Subjects: All subjects who are randomized to a treatment.

- Modified Intent-to-Treat (MITT) Analysis Population: All randomized subjects who have received at least one dose of the study medication. Subjects will be grouped according to the treatment to which they were randomized by IVRS at the start of the study. The MITT analysis population will also be referred to as "All Randomized and Treated Subjects".
- As-Treated Analysis Population: All Subjects who have received at least one dose of study medication. Subjects will be grouped according to the treatments to which they are randomized, except in cases where a subject receives a different treatment for the entire course of short term period. In this case, the subject will be analyzed and presented by the treatment they actually received.
- Efficacy Analysis Population: This is a subset of the MITT Analysis Population where the subjects from Part 1 who are randomized to a treatment arm and discontinued based on the Part 1 Safety/RO analysis are excluded. All efficacy results that involve statistical comparison or modeling will be based on this analysis population unless otherwise specified.
- Biomarker Analysis Population: All subjects who receive any study medication and have at least one post-treatment biomarker measurement.
- Pharmacokinetic Population: All subjects who received BMS-931699 and have any available concentration-time data.
- Immunogenicity Population: All subjects who receive any study medication and have at least one post treatment immunogenicity assessment.

7 STATISTICAL ANALYSES

7.1 General Methods

The table below provides an overview of the efficacy analyses to be performed.

Table 7.1-1: Planned Efficacy Analyses

Measure of Interest (Primary and Secondary Efficacy Endpoints):	Analysis Method
Primary endpoint (Binary): BICLA response rate at Day 169	Chi-Square test, one sided p-value will be provided. Point estimates and 90% CI of treatment difference. No multiplicity adjustments.
Binary endpoints : SRI(X) [X= 4, 5, 6, 8] at Day 85 and Day 169, BICLA response rate at Day 85, CLASI Response at Day 85 and Day 169	Point estimate and 90% CI of response rate, point estimates and 90% CI of treatment difference
Continuous endpoints: CLASI, BILAG-2004, SLEDAI-2K, MDGA, ACR 28 [tender, swollen, and active, (i.e. joints that are both tender and swollen) joint	Mixed model analysis, point estimates and 90% CI for mean change from baseline within treatment groups, point estimates and 90% CI for mean change from baseline between treatment groups (BMS vs. PLA)

Table 7.1-1: Planned Efficacy Analyses

Measure of Interest (Primary and Secondary Efficacy Endpoints):	Analysis Method
Counts]	

The Chi-square tests will be used to compare the BICLA response rates at Day 169 between each of the active treatment arms and the placebo arm. The Chi-square p-value will be provided for each comparison between the active treatment and the placebo arm. In addition, the difference of the response rates between each active treatment and the placebo arm will be estimated and their corresponding 90% confidence interval will be calculated. No adjustment will be made for multiplicity.

For each of the binary endpoints, the estimate and its corresponding 90% confidence interval will be calculated for the difference of the proportions between each active treatment arm and the placebo arm at the specified visit, similar to the primary analysis except no p-values will be provided.

The construction of confidence intervals for binary variables will be based on normal approximation if the number of the events in each individual treatment arm is at least 5. Otherwise, confidence interval using an exact method will be provided. The exact confidence intervals for binary variable within treatment group and between treatment groups will be constructed.

For each of the continuous endpoints, the mixed model will be fit with treatment, visit, and treatment-by-visit interaction as the fixed effects and measurements within each subject as the repeated measurements. The baseline value will be added into the model as a covariate if necessary. Based on the mixed model, the estimate and 90% confidence interval will be calculated for the difference between each active treatment and the placebo at each specified visit.

The results will be presented by the following grouping scheme in different periods, unless otherwise specified.

- **Pre-treatment period**: Only one group: All enrolled subjects.
- **Short-term period (ST)**: The results will be presented by treatment groups.

7.2 Study Conduct

Relevant protocol deviations, which could have an impact on the primary/key efficacy endpoint, will be identified for all subjects who are randomized and receive study medication during the double-blind period. Details of relevant protocol deviations are provided in Appendix 1.

All subjects with relevant protocol deviations that could affect the primary efficacy will be identified prior to database lock and unblinding of treatment assignment. All relevant protocol

deviations will be listed and summarized by treatment group. The subjects identified as relevant protocol deviation will not be excluded from any analysis population described in Section 6.3. If at least 10% of total subjects in all treatment arms in the efficacy analysis population demonstrate relevant protocol deviations, a per-protocol analysis may be performed.

7.3 Study Population

Frequency distributions or summary statistics of data pertaining to subject disposition, demographic characteristics, baseline disease characteristics, and medical history will be tabulated and displayed by treatment group. No statistical test will be carried out for comparison of any baseline measures among treatment groups.

These presentations will be provided by randomized treatment group for the MITT analysis population, unless specified otherwise.

7.3.1 Subject Disposition

The number of subjects enrolled into the study at screening, the number of subjects randomized, the number of subjects entering the short-term period (ST) and the number of subjects who complete the ST period will be summarized by treatment groups. The number of subjects discontinuing from the ST period (subject disposition) will also be summarized, together with the reasons for discontinuation, taken from the CRF status pages. The subject disposition for the ST period will be based on the MITT analysis population.

7.3.2 Demography and Baseline Characteristics

Demographic and baseline disease characteristics will be summarized by randomized treatment group and overall based on the MITT analysis populations. For continuous variables, they will be summarized using means, standard deviations and ranges. The distribution of categorical variables such as gender, race etc. will be summarized by treatment group using frequency and percentage.

Demographic characteristics include the following:

- Age in years
- Weight in kg
- Gender: Male, Female
- Race
- Ethnicity (Hispanic and Non-Hispanic)
- Region (North America, South America, Europe and Asia)
- Stratification [Japan vs. Non-Japan, MMF use (Y/N)]
-

Baseline clinical characteristics include the following:

- Time from First Diagnosis of SLE (years)
- Tender Joint Counts
- Swollen Joint Counts

- Active (Both Tender and Swollen) Joint Counts
- BILAG classifications and numerical scores
- SLEDAI-2K Score
- CLASI Score
- Physician's global assessment of disease activity (MDGA)
- Subject's global assessment of disease activity (PGA)
- SLICC/ACR damage index
- Total corticosteroid daily dose (prednisone equivalent) at baseline

7.3.3 Medical History and Prior Medication

Medical history and prior medication information before the first study medication will be summarized by treatment group using relative frequencies. Listings of medical history and prior medication will be presented. The results will be based on the MITT analysis population.

7.4 Extent of Exposure

7.4.1 Study Therapy

In general, extent of exposure during ST will be summarized for the As-treated analysis population in two ways:

- by the number of study drug injections: Frequency distribution of number of injections will be summarized
- by the number of days: the subject is known to be on study drug, ignoring any dosing interruptions

In general, exposure to study drug during the short-term period (ST) is calculated as follows:

1. The subjects who enter in the long-term extension period (LTE) within 7 days of the last study medication in ST,

   Exposure (in days) = (date of first study medication in LTE - date of first study medication in ST).

2. All other subjects (including the subjects who discontinue early, complete the ST but do not enter in the LTE or enter LTE after 7 days post the last dose of the ST):

   Exposure (in days) = date of last study medication in ST - date of first study medication in ST + 7

Note that, the offset of 7 days represents the dosing interval of 7 days.

The frequency will be presented according to the duration ranges (in days) : <=28, 29-56, 57-84 and so on.

7.4.2 Discontinuations from Study Therapy

Discontinuation from study therapy is defined as subject's termination of the study medication without resumption prior to study completion. Note that, subjects are to be discontinued if more than 3 consecutive doses of study medication are missed. Discontinuation from study therapy during the ST period will be summarized by reason for the premature termination, taken from the study status case report form (CRF) pages. These summaries will be provided based on the MITT analysis population by randomized treatment group and overall.

7.4.3 Treatment Compliance

The CRF for this study will capture information on injections of the study medication. All subjects who skip any study drug injection will be listed. The number of subjects with missed injections (excluding missed injections due to premature discontinuation from the study) by number of missed injections will be summarized based on the As-treated analysis population for the ST period by treatment groups.

[redacted]

7.4.5 Immunosuppressant Medication

For the purpose of this analysis plan, the term DMARDs will refer to non-biologic DMARDs. Immunosuppressant medications summaries will be provided to present the medication use of corticosteroids, NSAIDs, and DMARDs. Number and percent of subjects receiving corticosteroids, NSAIDs, and DMARDs will be summarized by treatment group for MITT analysis population during the ST period.

A comprehensive listing of all concomitant medications including immunosuppressant will be also provided for the MITT analysis population.

7.4.6 Corticosteroids

The average daily corticosteroid dose (prednisone equivalent) and its change from baseline will be summarized over time by treatment group for the ST period.

7.5 Efficacy

The efficacy analyses will be performed using the efficacy analysis population. The results will be grouped by "as randomized" treatment groups, unless otherwise specified. No formal statistical testing was pre-specified for other efficacy endpoints except the primary endpoint of BICLA response at Day 169. The analyses of binary endpoints and continuous endpoints will be provided over time including the milestone visits (e.g. Day 85, Day 169 etc.).

7.5.1 Primary Efficacy Endpoints in the Short-term Period

The primary efficacy endpoint is BICLA response rate at Day 169.

The Chi-square tests will be used to compare the BICLA response rates at Day 169 between each of the active treatment arms and the placebo arm. The Chi-square p-value will be provided for each comparison between the active treatment and the placebo arm. No adjustment will be made for multiplicity. In addition, the difference of the response rates between each active treatment and the placebo will be estimated and their corresponding 90% confidence interval will be calculated for Day 169 visit and other applicable visits.

The point estimate for BICLA response rate along with 90% confidence interval will be calculated for each treatment arm for each applicable visit.

All subjects who prematurely discontinue the trial after receiving study medication will have missing data imputed as non-responder at all scheduled protocol visits subsequent to the point of discontinuation. The analysis will be performed based on Efficacy Analysis Population. More details of imputation rule for classification of binary response status are provided in Sections 8.3, 8.1.1, and 8.1.2.

For BICLA, the number of responders, together with the rescue status (with or without one-time rescue allowed per protocol) will be summarized by treatment group. At the same time, the number of nonresponders will be summarized by treatment group, together with the reasons for nonresponse.

7.5.2 Secondary Efficacy Endpoints in the Short-term Period

For each of the binary endpoints (SRI(X) etc.), the estimate and its corresponding 90% confidence interval will be calculated for the difference of the proportions between each active treatment arm and the placebo arm for each applicable visit, similar to the primary analysis except no p-values will be provided.

For each of the continuous endpoints (BILAG-2004, SLEDAI, MDGA, CLASI, etc.), the mixed model will be fit with treatment, visit, and treatment-by-visit interaction as the fixed effects and measurements within each subject as the repeated measurements. The baseline value will be added into the model as a covariate if necessary. Based on the mixed model, the estimate and 90% confidence interval will be calculated for the difference between each active treatment and the placebo at each specified visit. No p-values will be provided. In addition, Descriptive summary statistics of mean change from baseline and 90% confidence interval for each treatment arm will be provided. The missing data in continuous measures will be handled via utilization of mixed model procedure in SAS.

7.5.3 Subgroup Analyses

The subgroup analyses of the primary and selected secondary endpoints ((SRI (X): X = 4, 6.)) will be performed within the following predefined subgroup (Non-Japan subjects). If the value of

the grouping variable cannot be determined for a subject, the subject will be excluded from the corresponding subgroup analysis.

In addition, other subgroup analyses may be performed on primary and selected secondary endpoints as post-hoc analysis.

7.5.4 Sensitivity Analyses

Sensitivity analysis for the primary endpoint: Adjustment for the randomization strata using Cochran-Mantel-Haenszel (CMH) methods will be applied on the primary endpoint.

If at least 15-20% of total subjects through all treatment arms in the efficacy analysis population discontinued due to any reasons other than Lack of Efficacy (LOE), AE or Unknown reason, a sensitivity analysis may be performed. The subjects who have missing response due to premature discontinuation with reasons other than LOE, AE or Unknown will be excluded from the analyses, hence defining the sensitivity analysis population. The sensitivity analysis will be carried out for the primary and selected secondary endpoints (SRI (X): X = 4, 6.) based on the sensitivity analysis population.

7.5.5 Per-Protocol (PP) Analyses

If at least 10% of total subjects in all treatment arms in the efficacy analysis population demonstrate relevant protocol deviations, a per-protocol analysis may be performed. A per-protocol analysis will be carried out for the primary and selected secondary endpoints (SRI (X): X = 4, 6.) based on the per-protocol (PP) analysis population. A Per-protocol (PP) analysis population will exclude all subjects with at least one relevant protocol deviation (defined in Appendix 1) from the efficacy analysis population based on the following rules.

- The subjects satisfying eligibility criteria deviation in Appendix 1 will be excluded from the PP analysis population completely.
- The subjects satisfying incorrect dosing criteria deviation in Appendix 1 will be excluded from the PP analysis population partially (only the data after deviation occurs will be excluded from the per-protocol (PP) analysis.)

7.6 Safety

Analysis of all safety data will follow the BMS standard safety data conventions[4] [5] and supplements to the standard conventions as defined in this document.

The evaluation of drug safety is based primarily on clinical AE, laboratory abnormalities and vital signs reported during the study.

The summaries of AEs and laboratory marked abnormality during the ST period will be provided by treatment groups. AEs and laboratory marked abnormalities will be included in the summary tables based on ST period if the onset date is on or after the first dose date and event occurs

within the ST period for continuing subjects or within 42 days post last dose date for discontinued subjects.

Frequency distribution and individual listings of all AEs will be generated. Laboratory marked abnormality using pre-defined abnormality criteria will also be descriptively summarized. There will be no statistical testing of group difference with respect to frequencies of AEs or laboratory marked abnormalities.

All recorded adverse events will be listed and tabulated by system organ class, preferred term, treatment and overall. Any physical examination findings will be listed. ECG, vital signs and clinical laboratory test results and corresponding change from baseline values will be listed and summarized by treatment. Values for ECG, vital signs and clinical laboratory test results outside the pre-specified criteria will also be listed and summarized. ECG readings will be evaluated by the Investigator and abnormalities, if present, will be listed.

Unless otherwise specified, baseline is defined as the last non-missing result with a collection date-time less than the date-time of the first dose of study medication, for laboratory results, vital signs and ECG.

Presentations for the ST period will be provided by "as treated" treatment group for the as-treated analysis population, unless otherwise specified.

7.6.1 Adverse Events

The investigator will determine the intensity of each AE as mild, moderate, severe, or very severe. In addition, the investigator will determine the relationship of the AE to the administration of the study drug.

All AEs are coded and grouped into Preferred Terms (PT) by System Organ Class (SOC), using the latest approved version of the Medical Dictionary for Regulatory Activities (MedDRA) at the time of database lock. Listings and summaries will be based on the resulting SOCs and PTs.

AEs will be included in the frequency tabulations if they occur while a subject is on study medication up to and including 42 days after the last dose of study medication. These AE summaries will be based on proportions, which represent the number of subjects experiencing the AEs divided by the number of subjects received with at least 1 dose of study medication in the current period.

All AE listings will indicate the unique subject identifier, age, gender, current treatment, the date of onset, the date of resolution, day of onset relative to the start of treatment, action taken, investigator's assessment of severity and relationship to study drug. Additional listings will be provided for adverse events leading to discontinuation and adverse events without recorded resolution. Summaries of adverse events will include adverse events, adverse events by intensity and adverse events by relationship.

Where a subject has the same adverse event, based on preferred terminology, reported multiple times in a single analysis period, the subject will only be counted once at the preferred terminology level in adverse event frequency tables.

Where a subject has multiple adverse events within the same system organ class in a single analysis period, the subject will only be counted once at the system organ class level in adverse event frequency tables.

When a subject has the same adverse event, based on preferred terminology, reported multiple times in a single analysis period, the following criteria, in order of precedence, will be used to select the event to be included in summary tables:

- Relationship to study medication
- Intensity of event
- Onset date and time

When reporting adverse events by intensity, in addition to providing a summary table based on the event selection criteria detailed above, summary tables will also be provided based on the most intense event during the analysis period - independent of relationship to study medication. For these tables, the following criteria, in order of precedence, will be used to select the event to be included in summary tables:

- Intensity of event
- Onset date and time

Subjects will only be counted once in the 'Total' at their maximum intensity, regardless of SOC or PT.

Any reported AEs that occur more than 42 days after the last dose of study drug will be excluded from safety summaries. Any reported AEs that occur more than 42 days post the last dose of study drug will appear in the listing.

Summary information (the number and percent of subjects by treatment) will be tabulated for the following AEs:

- All AEs including clinical and laboratory AEs
- Most common AEs (reported in 5% of subjects or more in any treatment group)
- All related events
- Serious AEs (SAEs)
- Discontinuations due to AEs

Laboratory AEs are laboratory changes identified by the investigator as AEs and thus reported on the AE pages of the eCRF.

7.6.2 Adverse Events of Interest

The following adverse events of special interest will be listed and summarized by treatment groups during the ST period.

- Infections: All reported infections and infestations within the SOC: *Infections and infestations.* The severity of serious infections will be summarized. In addition, a listing of opportunistic infections will be provided.
- Malignancy: All reported events defined in MedDRA Maintenance and Support Services Organization (MSSO) Structured MedDRA Query (SMQ) list.
- Autoimmunity: All reported events defined in pre-specified MedDRA code of autoimmune disorders.
- Injection Reactions (Local and Systemic): All reported events defined in pre-specified MedDRA code of injection reactions (Local and Systemic).

7.6.3 Deaths

All AEs with the outcome of death reported during the study will be listed. All reported deaths after a subject is enrolled will be listed separately by subject.

7.6.4 Other Serious Adverse Events

SAEs are captured on an eCRF page. All SAEs and related SAEs will be summarized by treatment group and listed.

7.6.5 Adverse Events Leading to Discontinuation of Study Therapy

AEs leading to discontinuation of study drug are identified on the eCRF will be listed and summarize.

7.6.6 Multiple Occurrences of Adverse Events

Several descriptive summaries of adverse events that takes into account the number of occurrences that an AE was reported by individual patients will be provided. In order to prepare these summaries, the CRF data will be processed according to standard BMS algorithms to categorize each line of patient data as a new occurrence or a continuation of an existing event. This determination will be based upon onset and resolution dates. Each line of patient data will represent the maximum severity observed as well as the last known assessed relationship to study medication by the investigator.

The number and proportion of subjects who experienced an AE once or multiple times (0 Events, 1 Events, 2-3 Events, >=4 Events) will be summarized by treatment groups for most common AE (at least 5% subject in any treatment group) and event of special interest, infections.

Listing displaying the unique instances of all AEs will be provided.

Note: The unique instances of all AEs will be generated after duplicates have been eliminated and overlapping and contiguous occurrences of the same event have been collapsed.

7.6.7 Summaries of Laboratory Data

The results of all protocol-specified clinical laboratory tests, as well as laboratory results outside of the normal range, will be listed. Scheduled laboratory measurements and corresponding change from baseline values will be summarized by treatment and nominal visit for each laboratory test for the MITT analysis population.

The frequency of subjects with any marked laboratory abnormality (based on sponsor defined criteria) will be presented by laboratory test during the ST period. The results are based on the As-treated analysis population. The criteria used for classifying laboratory test results as markedly abnormal will be listed.

7.6.8 Other Safety Considerations

7.6.8.1 ECG

Summary statistics (n, mean, standard deviation, median, minimum, and maximum) will be presented for each ECG parameters and the corresponding changes from baseline by treatment and visit. The baseline value is defined as the last measurement before the first dose. ECG readings will be evaluated by the Investigator and abnormalities, if present, will be listed.

The frequency distribution of subjects' maximum recorded post-dose QTcF, PR, QRS and ΔQTcF will be tabulated by treatment and summarized within the CSR text for the following ranges:

For QTcF: $QTcF \leq 450$ msec, $450 \text{ msec} < QTcF \leq 480$ msec, $480 \text{ msec} < QTcF \leq 500$ msec, $QTcF > 500$ msec

For PR: $PR \leq 200$ msec, $PR > 200$ msec

For QRS: $QRS \leq 120$ msec, $QRS > 120$ msec

For ΔQTcF: $\Delta QTcF \leq 30$ msec, $30 \text{ msec} < \Delta QTcF \leq 60$ msec, $\Delta QTcF > 60$ msec

Individual QTcF, PR, QRS or ΔQTcF values meeting these criteria will be flagged in the data listing.

7.6.8.2 Vital Sign

Vital signs parameters (systolic blood pressure, diastolic blood pressure, heart rate, respiratory rate and body temperature) will be listed. Summaries of vital sign parameters will be provided for each vital sign parameter at corresponding visits by treatment and respective changes from baseline. Subjects with vital signs outside of a pre-specified range will also be listed.

The following criteria will be used to determine vital sign results that are outside of a pre-specified range, where changes from baseline are based on matched postural positions and are calculated as parameter value - baseline parameter value:

Table 7.6.8.2-1: Vital signs criteria

Vital sign	Criteria
Heart Rate(bpm)	Value > 100 and change from baseline > 30, or Value < 55 and change from baseline < -15
Systolic BP(mmHg)	Value > 140 and change from baseline > 20, or Value < 90 and change from baseline < -20
Diastolic BP(mmHg)	Value > 90 and change from baseline > 10, or Value < 55 and change from baseline < -10
Respiration(breaths/min)	Value > 16 or change from baseline > 10
Temperature (°C)	Value > 38.3°C or change from baseline > 1.6°C

7.6.8.3 Conventions of ECG Data

All available non-missing values of ECG parameters should be used in the listings, summarizations, and analyses. However, if QTcF is missing and RR in seconds is available, then QTcF will be calculated as

$$QTcF = \frac{QT}{RR^{1/3}}.$$

If both QTcF and RR in seconds are missing, then QTcF will be calculated as

$$QTcF = \frac{QT}{\left(60/HEART\ RATE\right)^{1/3}}.$$

7.6.8.4 Physical Examination Findings

All physical examination abnormal findings will be listed per subject by visit.

7.7 Pharmacokinetic Analysis

Summary statistics of Ctrough (mean, SD, geometric mean, % CV, median, min, and max) for the PK analysis population will be provided by treatment and study day.

Similar statistics will be provided for post-dose concentration on Day 46 and Day 48.

The listing of Concentration of BMS-931699 by treatment and study day will also be provided.

[redacted]

[redacted]

7.9 Immunogenicity Analysis

Serum for anti-BMS-931699 antibody titers will be drawn at Day 1, and during the ST period. Subjects who discontinue study medication will have a serum samples collected at the early termination visit and 42 days after the last dose of study medication (42 days follow up). Lack of immunogenicity is defined as the absence of a positive response.

For all summaries and listings unless specified otherwise, an anti-drug antibody (ADA) positive relative to baseline immunogenicity response is defined as:

1) a missing baseline immunogenicity measurement and a positive laboratory reported immunogenicity response post-baseline
2) a negative laboratory reported baseline immunogenicity response and a positive laboratory reported immunogenicity response post-baseline
3) A positive laboratory reported baseline immunogenicity response and a positive laboratory reported immunogenicity response post-baseline that has a titer value of 4 fold or greater than the baseline titer value.

All other immunogenicity measurements will be classified as negative immunogenicity response.

In addition, persistent anti-drug antibody is defined as anti-drug antibody positive occurrence relative baseline detected at 2 or more consecutive time points at least 3 months apart. Immunogenicity Frequencies

7.9.1 Immunogenicity Frequency

The number (and percentage) of anti-drug antibody positive subjects relative to baseline will be provided. Similar analyses will be performed for persistent anti-drug antibody positive subjects (relative to baseline).

The number (and percentage) of subjects whose last sample is the only anti-drug antibody positive sample relative to baseline will be presented also.

A listing of titer values for all ADA positive subjects (laboratory reported) will be provided for antibodies against the domain antibody portion of the drug.

7.10 Interim Analyses

7.10.1 Interim Safety/RO Analysis at Day 29

When at least 6 patients per treatment arm have either reached Study Day 29 or discontinued, an interim analysis (IA) for safety and RO will be performed. The safety analysis will focus on incidence and severity of all adverse events (AEs), serious AEs and pre-established Events of Special Interest such as infection AEs and any other safety analysis requested by DMC. A separate document, DMC charter, describes the scope of these analyses. For RO: the median RO at Day 29 for each treatment arm will be calculated. Based on the results of this interim analysis, a decision will be taken regarding the doses that will be carried forward into Part 2 of the study. Specifically, dosing regimens originally included in Part 1 may be discontinued and/or new dosing regimens may be added according to the following criteria:

Safety:

The DMC in conjunction with an unblinded internal monitoring team (The details of the members are provided in Section 2.3.2) may require one or more doses to be discontinued if stopping criteria are met or other safety signals arise that the Medical Monitor and/or DMC consider of sufficient concern.

Receptor Occupancy:

- If median RO of any dose is <20%, the sponsor may consider dropping that dose
- If the median RO for all doses is >90% the sponsor may consider adding or replacing a dose in Part 2 of the study to ensure an adequate pharmacodynamic range (dose not to exceed 12.5 mg weekly)

Dose decrease and/or reduction of frequency of administration may also be considered if RO results fall outside the parameters indicated above. This adjustment may occur for safety reasons or in case unforeseen RO profiles observed in SLE patients. The decision to adjust dose and/or frequency will be taken after review of the data by the clinical team. Any decision related to safety reasons will be taken in conjunction with DMC recommendations.

RO analysis in interim analyses 1 will be done by a Pharmacokinetic scientist or a Pharmacometrics analyst. If dose/doses are adjusted based on this RO analysis, then the interim analyses results may be included in CSR as an appendix.

7.10.2 Interim Analysis for Futility and Dose Adaptation at Day 85

After 30 subjects per treatment arm (including the subjects from Part 1) have completed at least 84 days of the treatment period or discontinued the treatment, an interim analysis will be conducted to assess the futility of the BMS-931699 dose arms and the overall safety. The results of the interim analysis will be reviewed by an unblinded sponsor team (Independent of the study team) who will provide recommendations to the blinded study team. The details of the members are provided in Section 2.3.2. The blinded study team will make the decision. The unblinded BMS team may recommend to:

- maintain the current design;
- drop a certain dose arm(s) and equally randomize the remaining unallocated subjects to the remaining arms;
- stop the study if all dose arms need to be dropped for safety or futility reasons.

Any decision related to safety reasons will be taken in conjunction with DMC recommendations.

The interim analysis for futility and dose adaption will be performed on the BICLA response at Day 85 of those 30 subjects per arm and BICLA response at Day 169 of those subjects who have completed 168 days of treatment or discontinued the treatment using a Bayesian predictive approach.[6] This analysis assumes that the BICLA at Day 85 and at Day 169 in the subjects who have not yet been observed will be similar to what were observed for the subjects included at the interim analysis at Day 85 and Day 169, respectively. As such the unobserved data will be simulated from the predictive distribution conditional on the interim data and the prior distribution of the treatment difference (using a non-informative prior). Under the Bayesian framework, the posterior distribution of the treatment difference (BMS-931699 active dose – Placebo) will be constructed to determine the predictive probability of a successful outcome (i.e., reach statistical significance) at the planned end of the trial for each dose arm. Based on computer simulations of the operating characteristics of the study design, the futility threshold in this interim analysis is set to 0.2. For the comparison of interest, if the predictive probability of a successful outcome at the planned end of the trial falls below 0.2, that BMS-931699 dose arm is deemed futile in terms of BICLA response. However, the final futility will also incorporate the results from the efficacy endpoints such as but not limited to SRI(4), ACR28 and some biomarkers such as but not limited to C3, C4, ANA and anti-dsDNA. The final decisions will be made based on the discussions using the pre-specified rules in the *Interim Analysis SAP* and/or the evidence from the totality of the data.

The proposed stopping rules at the interim analysis, based on the futility assessment and the overall safety assessment, are as follows:

- Rule 1: Drop the futile dose arm(s) and the dose arm cannot be dropped until the lower dose arm has been dropped.
- Rule 2: Drop the dose arm(s) with safety issues identified;

- Rule 3: If safety issues are identified in all dose arms or all dose arms are futile, stop the study.

In case dosing arms are dropped following the interim analysis, the remaining unallocated subjects will be randomized into the remaining arms.

If the predictive probability of success for each active arm versus placebo is greater than or equal to 0.9, then the BMS team may also recommend adding an additional dose arm at lower dose level than current active doses to explore the suboptimal dose.

In that case, the remaining un-allocated subjects with up to 40 additional subjects will be randomized into the remaining arms and the new dose arm based on the new randomizations schedule.

A separate document, *Interim Analysis SAP*, has been developed in order to provide a set of rules with clarifying details as a guidance to make any decisions about the dose adaptation.

7.11.2 Time to event Analysis

Time to treatment failure and time to first flare as assessed by new BILAG 2004 index flare definition will be summarized. For those time to event endpoints, if necessary, the point estimate and its 90% confidence interval for the hazard ratio of each active treatment arm versus placebo will be calculated based on the Cox proportional hazards models. In addition, the distribution of time to the occurrence of event will be estimated using Kaplan-Meier method. The analysis will be based on the Efficacy Analysis Population.

Treatment failure is defined as non-protocol treatment i.e., new or increased immunosuppressive treatment or antimalarials; or parenteral corticosteroids; or protocol prohibited dose of oral corticosteroid; or premature discontinuation from study treatment.

[redacted]

7.11.4 Immunogenicity and PK/Efficacy/Safety

A listing of drug concentration with immunogenicity assessment at each pre-specified time point will be provided for each individual subject who has positive immunogenicity response (laboratory reported ADA positive).

Efficacy listings for all anti-drug antibody positive subjects (laboratory reported ADA positive) may be provided.

Listings of AEs, SAEs, and local injection site events for all anti-drug antibody positive (laboratory reported ADA positive) subjects will be provided.

A list of neutralizing positive subjects (subjects with at least one ADA positive sample relative to baseline with neutralizing antibodies detected) will be provided also.

Depending on the incidence of ADA positive relative to baseline observed in SLE patients within this trial, individual plots of Ctrough with ADA status may be generated for all subjects with ADA positive relative to baseline.

7.11.5 Exposure-Response Analysis

The PK data obtained from this study may be pooled with data from other studies to perform an integrated population PK (PPK) analysis (including assessment of covariate effects on PK), as well as exposure-response analysis for selected safety and efficacy endpoints. These analyses will be described in a separate report(s).

7.11.6 Analysis for the PK Sub-Study

As specified in the amendment 9 of the protocol, participation for PK sub-study is voluntary. While there is no minimum required number of subjects, every effort will be made to enroll approximately 4 subjects/arm in the ST study (ie approximately 20 subjects total) and approximately 6 subjects/arm in the LTE (ie approximately 30 subjects total). This sample size is expected to provide sufficient data to adequately estimate PK parameters for lulizumab in this patient population. Subjects will be asked to read, understand, and sign an informed consent form designed for the purpose of collecting the additional PK/RO samples.

Pharmacokinetics Analysis

All PK summaries and analysis will be done separately for first dose exposure and steady state exposure:

Summary statistics will be provided for all PK parameters by treatment groups for first dose exposure and steady state exposure. Geometric means and coefficients of variation (CV) will be

provided for Cmax, AUC(TAU), and C(Day15). Medians and ranges will be provided for Tmax. Mean and standard deviation may be provided for T-HALF and CLT/F. Additionally, scatter plots of AUC(TAU), Cmax and C(Day 15) versus dose for the three every 2 weeks treatment groups (1.25 mg, 5 mg and 12.5 mg every 2 weeks) to assess the dose-dependency. All plots will be provided separately for first dose exposure and steady state exposure.

A statistical analysis using a power model will be performed to assess dose proportionality in AUC(TAU), Cmax and C(Day 15) for only the three every 2 weeks treatment groups (1.25 mg, 5 mg and 12.5 mg every 2 weeks). The analyses will be carried out separately for first dose exposure and steady state exposure.

In addition, a listing will be provided for all PK parameters and lulizumab concentration.

Receptor Occupancy Analysis:

Medians and ranges will be presented for RO by Day for all treatment groups in each subgroup. Box plots of RO by Day will be provided for all treatment groups in ST and LTE.

Immunogenicity Analysis:

The number and percentage of subjects will be summarized for each treatment, and the corresponding antibody titer values will be listed. Exploratory plots or analysis might be conducted to evaluate the impact of immunogenicity on PK of lulizumab.

All other safety, efficacy and biomarker observations that are being planned in the main study protocol will be reported in the clinical study report of the main study.

8 CONVENTIONS

8.1 Calculations of Key Measures

8.1.1 Evaluation of BICLA response

BICLA response is defined based on the following components:

- British Isle Lupus Assessment Group improvement, defined as BILAG ‘A’s at Baseline improved to ‘B’/’C’/’D’, and BILAG ‘B’s at baseline improved to ‘C’/’D’, and no BILAG worsening in other BILAG organ systems such that there are no new BILAG ‘A’s or greater than 1 new BILAG ‘B’; and,
- No worsening in the SLEDAI-2K total score compared to Baseline (defined as no increase in SLEDAI total score); and,
- No worsening in the physician’s global assessment (MDGA) of disease activity (“no worsening” is defined as less than 10% worsening, equivalent to a 10mm increase on a 100mm visual analog scale [VAS]) compared to Baseline; and,
- No changes in concomitant medications according to the following criteria:
  - No increase of or addition of a new immunosuppressant agent (azathioprine, methotrexate, anti-malarial) over baseline levels

– No increase in corticosteroid dose above baseline level outside of those allowed per protocol.

If a subject failed any of the above criteria above, the BICLA response status of the subjects will be set to non-responder.

For the determination of missing BICLA response status, all subjects who prematurely discontinue from the study medication will be imputed as non-responder at all scheduled efficacy visits subsequent to the point of discontinuation.

If any subject has at most one missing component (e.g. MDGA for BICLA) at a given efficacy visit not due to premature discontinuation, the component-wise response status will be imputed based on the following rules.

- If the given visit is not the last scheduled efficacy visit (Day 169), imputation will depend on the observed responses from the immediate previous efficacy visit and the immediate next efficacy visit within 28 days of the current visit date.
- If the given visit is the last scheduled efficacy visit (Day 169), the imputation will depend on the observed responses at the previous 2 consecutive efficacy visits (within 56 days).
- If a subject has response (at component level) at both efficacy visits, the missing response status at current visit will be imputed as responder. Otherwise, response status will be set to non-responder.

The BICLA response status (composite) for a subject who has more than one component (e.g. MDGA and SLEDAI-2K are missing for BICLA) missing at a given efficacy visit not due to premature discontinuation will be imputed as non-responder.

After any imputation above or none, if for some reason the BICLA response status still cannot be determined, including the case where baseline data is missing, then its value will be set to non-responder.

8.1.2 Evaluation of SRI(X)

In addition to BICLA, a novel endpoint, the SLE Responder Index [SRI] with a modification used as a primary endpoint in the Anifrolumab study, the results of which were presented and published as an abstract at the American College of Rheumatology (ACR) National Meeting in November, 2015[7]. The basic difference between the SRI defined in our protocol and the SRI in the Anifrolumab study is the addition of a fourth component of the endpoint. The original definition of SRI in the protocol does not include the fourth component based on the use of restricted medications.

The use of restricted medications is an important confounding factor when we analyze the endpoints such as BICLA, and SRI(X). Hence an additional component based on the use of restricted medications (same as the 4th component in the definition of BICLA) is deemed necessary to incorporate into the original definition of SRI (X) in the protocol. The modification to the existing definition of SRI is described below.

An SRI(X) [X= 4, 5, 6, and 8] response @ Day T (e.g. T = 85, 169) is defined as:

- a reduction in Day 1 SLEDAI-2K disease activity score of ≥ X points at Day T; and,
- no worsening of disease (defined as an increase of ≥ 10mm on a 100mm VAS from Day 1 as measured by the MDGA up to time T; and,
- no new BILAG-2004 Index 'A' organ system score and no more than one new or worsening BILAG-2004 Index 'B' organ system scores up to time T ; and,
- *No changes in concomitant medications according to the following criteria:*
  - No increase of or addition of a new immunosuppressant agent (azathioprine, methotrexate, anti-malarial) over baseline levels
  - No increase in corticosteroid dose above baseline level outside of those allowed per protocol.

If a subject failed any of the above criteria, the SRI(X) response status of the subjects will be set to non-responder.

For the determination of missing SRI(X) response status, all subjects who prematurely discontinue from the study medication will be imputed as a non-responder at all scheduled efficacy visits subsequent to the point of discontinuation.

If any subject has at most one missing component (e.g. MDGA for SRI(X)) at a given efficacy visit not due to premature discontinuation, the component-wise response status will be imputed based on the following rules.

- If the given visit is not the last scheduled efficacy visit (Day 169), imputation will depend on the observed responses from the immediate previous efficacy visit and the immediate next efficacy visit within 28 days of the current visit date.
- If the given visit is the last scheduled efficacy visit (Day 169), the imputation will depend on the observed responses at the previous 2 consecutive efficacy visits (within 56 days).
- If a subject has response (at component level) at both efficacy visits, the missing response status at current visit will be imputed as responder. Otherwise, response status will be set to non-responder.

The SRI(X) response status (composite) for a subject who has more than one component (e.g. MDGA and SLEDAI-2K are missing for SRI(X)) missing at a given (efficacy) visit not due to premature discontinuation will be imputed as non-responder.

After any imputation above or none, if for some reason the SRI(X) response status still cannot be determined, including the case where baseline data is missing, then its value will be set to non-responder.

Note that, the corrected threshold is 10mm in the second component of SRI(X) instead of 30mm (as in the original protocol) on a 100mm VAS of MDGA. The correct threshold is used in the definition of SRI(X) in this section.

8.1.3 Evaluation of Major Clinical Response and Partial Clinical Response

Major and Partial Clinical responses are defined as:

- **Major Clinical Response**: a patient with BILAG C scores or better at 6 months with no new BILAG A or BILAG B scores AND maintenance of response with no new BILAG A or B scores between 6 and 12 months.
- **Partial Clinical Response**: a patient with BILAG C scores or better at 6 months with no new BILAG A or BILAG B scores and maintenance of response without a flare for 4 months.

If a subject is not meeting any of the above criteria, the response status will be set to non-responder.

If a subject has any missing data at a particular visit, the subject will be excluded from the analysis at that particular visit. The details will be provided in the LTE SAP.

8.1.4 Evaluation of Flare and Severity of Flare

Flare and severity of flare based on new BILAG 2004 is defined as follows:

- Severe flare is any BILAG 2004 "A" score in any system due to items that are new or worse.
- Moderate flare is two or more "B" scores that are new or worse.
- Mild flare would be any patient with a single "B" score due to items that are new or worse or in those with three or more "C scores due to items that are new or worse.
- Anyone without these criteria would be categorized as no flare
- If a subject has any missing data at a particular visit, the subject will be excluded from the analysis at that particular visit.

8.1.5 Evaluation of SF-36 Summary Functions and Components

The SF-36 is composed of 36 items measuring 8 health concepts: Physical Functioning, Role-physical, Bodily Pain, General Health, Vitality, Social Functioning, Role-emotional, and Mental Health. These 8 scales form 2 summary scores: (1) Physical Component Summary (PCS) and (2) Mental Component Summary (MCS). Scoring conventions for the SF-36 scales and components are documented in "SF Health Outcomes ™ Scoring Software User's Guide"[8].

8.2 Baseline Measures

The baseline value is the last assessment taken prior to the first dose of study medication. In general, the baseline value is the assessment taken on study Day 1 before the administration of study medication unless otherwise specified.

If a measurement on Day 1 (baseline) is missing, the last assessment taken at the screening period prior to the first dose of study medication will be used.

8.3 Missing Measurements

For listings of binary response efficacy variables, missing values will be presented as missing. For determination of binary response variables, handling of missing values for BICLA and SRI(X) responses is addressed in Sections 8.1.1 and 8.1.2 unless otherwise specified.

If the response cannot be assessed due to missing data or a subject's early discontinuation, the following conventions will be implemented in a sequence:

- Any subject who prematurely discontinues the trial during the Short-term period after receiving study medication will have data imputed as non-responder at all scheduled protocol visits subsequent to the point of discontinuation up to the end of the Short-term period.
- After any imputation above or none, if for some reason the binary variable still cannot be determined, including the case where baseline data is missing, then its value will be set to non-responder.

For the analysis of continuous efficacy variables, mixed model will be used based on assumption of data being missing at random. Subjects with only baseline observations will be excluded from any changes from baseline analyses.

For safety, PK and immunogenicity measures and biomarkers, missing values will not be imputed.

8.4 Missing, Unknown or Partial Dates

The BMS safety guidelines for conventions relating to the handling of missing or partial dates and the determination of appropriate default values in such cases (in particular, for concomitant medication dose start-dates and end-dates and AE onset dates) will be utilized.

8.5 Day Ranges for Analysis of Time Points

Subjects who do not always adhere strictly to the visit schedule timing in the protocol. Therefore the designation of visits during the Short-term period of the study will be based on the day of evaluation relative to the trial (day of first study medication = study Day 1) rather than the nominal visit recorded in the eCRF. Mutually exclusive relative day windows are defined below to provide derived visits that correspond to the post-baseline time points. If a visit falls outside of the pre-specified visit windows, then the data collected at that visit will not be assigned a derived visit but will remain in the derived data sets. Determination of baseline values is addressed in Section 8.2.

If a subject has more than 1 visit where a measurement is recorded within a window, the measurement closest to the target day will be used. In case of 2 visits being equidistant from the target, the later measurement will be used in analyses. Exceptions to these rules apply to immunogenicity where the least favorable value (toward a positive response) in the window will be used.

For subjects who discontinue from study therapy prematurely, assessments performed after the last dose of the study drug will be included in the efficacy datasets provided that these assessments are made within 28 days of the last dose.

Designation of visits for efficacy assessments, immunogenicity samples and receptor occupancy samples during the Short-term period is tabulated below.

Table 8.5-1: Days Ranges for Assessments at Every Scheduled Visits During the Short-term Period

Visit	Target Day	BILAG-2004, SLEDAI 2K, MDGA, PGA, ACR28, CLASI, SF-36, Coomb's test-direct	FACIT, C3, C4	SLICC	Anti-dsDNA	CRP, IgG, IgA, IgM	Immuno-genicity (Anti-BMS-931699), Cytokines	ANA, Auto-antibodies
Baseline /Day 1	1	1	1	1	1	1	1	1
D15	15		8-21				8-21	
D29	29	15-43	22-43		15-43		22-57	15-43
D57	57	44-71	44-71		44-71			
D85	85	72-99	72-99		72-99	72-99	58-127	72-99
D113	113	100-127	100-127		100-127			
D141	141	128-155	128-155		128-155			
D169	169	156-190*	156-190*	156-190*	156-190*	156-190*	128-190*	156-190*

* For subjects who enroll in long-term extension (LTE) period, the upper inclusive limit of this window is the start of LTE dose.

9 CONTENT OF REPORTS

The results of the study conducted by Protocol IM128-027 will be presented in a standard BMS CSR appendix. Key results and any unanticipated findings that are unusual for this study will be identified. A meeting for the initial dissemination of study results will be held after database locked. Attendees at this meeting will review all efficacy and safety summaries and listings and will identify key results that should be highlighted in the CSR.

10 APPENDIX 1 RELEVANT PROTOCOL DEVIATIONS

Eligibility Deviations:

- Subjects who do not have SLE as defined by meeting 4 of the 11 classification criteria of the American College of Rheumatology for the classification of SLE, either sequentially or coincident at screening

- Subjects who do not have one 'A' in Musculoskeletal or Mucocutaneous body systems or two 'B's in BILAG score (at least one in Musculoskeletal or Mucocutaneous body systems and one in any other body system) at screening
- Elevated ANA<1:80 and anti-dsDNA is not positive and anti-Sm is not positive at screening
- Subjects who have SLEDAI < 6 at screening
- Subjects who receive more than 10mg/day of prednisone (or prednisone equivalent) at randomization day (Day 1)

Incorrect Dosing:

- At least 4 consecutive doses are missed prior to the Day 169 and not discontinued as per protocol

11 APPENDIX 2 DOCUMENT REVISION HISTORY

Version	Statistician	Date	Notes/Revisions
1.0	Pranab K. Mitra	01/16/2015	Original version
2.0	Xiaoni Liu Pranab K Mitra	04/28/2016	Incorporated the changes needed for the Protocol amendments. Revised the relevant deviation criteria. Modified the definition of SRI(X) by adding one component based on restricted medication into the composite response. Fixed several typos in the text. Further clarified the analysis and the convention sections.