Official Title of Study:

A Phase 2 Randomized, Double-Blind, Placebo-Controlled Study to Evaluate Efficacy and Safety of BMS-986165 in Subjects with Systemic Lupus Erythematosus

NCT Number: NCT03252587

Document Date (Date in which document was last revised): Dec 20, 2021

Statistical Analysis Plan

Sponsor:	Bristol-Myers Squibb
Protocol No:	IM011021
Protocol Date:	15-April-2020
Protocol Version No.:	Revised Protocol v4.0

Title:	A Phase 2 Randomized, Double-Blind, Placebo-Controlled Study to Evaluate Efficacy and Safety of BMS-986165 in Subjects with Systemic Lupus Erythematosus
[redacted]	
SAP Version No.	4.0

Approved v3.0 [redacted]

Table of Contents

Table of Contents 2
1.0 Purpose 6
2.0 Scope 6
3.0 Introduction 6
4.0 Study Objectives 6
4.1 Primary Efficacy 6
4.2 Secondary Efficacy 6
4.3 Safety 6
4.4 Pharmacokinetics 6
4.5 Pharmacodynamics 6
[redacted]
5.0 Study Design 7
5.1 Sample Size Considerations 8
5.2 Randomization 8
5.3 Unblinding Information 8
5.4 Summary of Changes to Version 2.0 of SAP 9
5.4.1 Changes due to protocol amendments after finalization of SAP V1.0 9
5.4.2 Additional Changes to SAP V2.0 9
5.5 Summary of Changes to Version 3.0 of SAP 10
5.6 Summary of Changes to Version 4.0 of SAP 11
6.0 Study Variables 11
6.1 Efficacy Assessments 11
6.2 Composite Measures 13
[redacted]
6.4 Safety 16
6.5 Predetermined Covariates and Prognostic Factors 16
[redacted]
7.0 Definitions 17
7.1 General 17
7.2 Analysis Visits Windows 18

8.0 Populations for Analyses 18
8.1 Enrolled 18
8.2 Randomized 18
8.3 All-treated 19
8.4 Biomarker 19
8.5 Pharmacokinetic 19
9.0 Analysis and Reporting 19
10.0 Data Review 19
10.1 Data Handling and Transfer 19
11.0 Statistical Methods 19
11.1 Subject Disposition 20
11.2 Demographic and Baseline Characteristics 20
11.3 Relevant Protocol Deviations 21
11.4 Exposure 21
11.4.1 Duration of Exposure 21
11.4.2 Summary of Dosing 21
11.4.3 Compliance 21
11.5 Prior and Concomitant Medications 22
11.5.1 Prior Corticosteroid Use 22
11.5.2 Concomitant Corticosteroid Use 22
11.6 Medical History 23
11.7 Efficacy Analyses 23
11.7.1 Primary Endpoint 24
11.7.2 Secondary Endpoints 26
11.7.2.1 SRI(4) Response at Week 48 26
11.7.2.2 BICLA Response at Week 48 27
11.7.2.3 CLASI-A Response (CLASI-50) at Week 48 28
11.7.2.4 LLDAS Response at Week 48 28
11.7.2.5 Tender and Swollen Joint Count at Week 48 29
[redacted]
11.7.4 Subgroup Analysis 31

[redacted]

11.8 Safety Analyses 36
11.8.1 Adverse Events 37
11.8.2 Adverse Events of Interest 38
11.8.3 Deaths and Serious Adverse Events....38
11.8.4 Laboratory Data....38
11.8.5 Vital Signs 39
11.8.6 ECG 39
11.8.7 Physical Examination 40
11.9 Statistical impacts due to Covid-19....40

11.9.1 Impact on efficacy endpoints....40
11.9.2 Impact on safety endpoints....40
12.0 Pharmacokinetic Analysis....41
12.1 Plasma Concentrations....41
12.2 Pharmacokinetic Parameters....41
13.0 Biomarker Analysis....42
14.0 Interim Analysis....42
15.0 References....42
Appendix 1 Glossary of Abbreviations....43
Appendix 2 Schedule of On-Treatment Activities....45
Appendix 3 Imputation Rules for Partial of Missing Dates....51
[redacted]
Appendix 5 Relevant Programmable PDs....57
[redacted]

1.0 **Purpose**

The statistical analysis plan (SAP) describes the statistical methods to be used during the reporting and analyses of data collected under Bristol-Myers Squibb IM011021 revised protocol version 4.0 dated 15 April 2020.

2.0 **Scope**

This plan is a document that will be created during the trial start-up and finalized prior to study unblinding. The SAP will be maintained throughout the lifecycle of the trial. If the SAP needs to be updated after finalization, then each version of the SAP will require sign off from the Biostatistics Asset Lead (BAL) or the BMS GBDS Lead.

The SAP outlines the following:

- Study objectives
- Study design
- Variables analyzed and analysis sets
- Applicable study definitions
- Statistical methods regarding relevant protocol deviations, study drug exposure, efficacy analysis, concomitant medications, adverse events handling, laboratory data and physical examinations
- Statistical methods for handling missing data and multiplicity

3.0 **Introduction**

The SAP describes the statistical methods to be used during the reporting and analyses of data collected under Bristol-Myers Squibb Protocol IM011021 version 4.0 (dated 15-Apr-2020). The SAP should be read in conjunction with the study protocol and case report form (CRF). This version of the plan has been developed using the protocol dated 15Apr2020 and CRF dated 14Jan2021.

4.0 **Study Objectives**

4.1 Primary Efficacy

- To assess the effect of BMS-986165 on Systemic Lupus Erythematosus (SLE) Responder Index (4) (SRI[4]) at Week 32 in subjects with SLE

4.2 Secondary Efficacy

- To assess the effect of BMS-986165 on measures of global and organ-specific SLE clinical response

4.3 Safety

- To assess the safety and tolerability of BMS-986165 in subjects with SLE

4.4 Pharmacokinetics

- To assess PK of BMS-986165 and its active metabolite (BMT-153261) in subjects with SLE

4.5 Pharmacodynamics

- To assess the effect of BMS-986165 on PD markers in subjects with SLE

5.0 **Study Design**

This is a randomized, placebo-controlled, double-blind, multicenter study to evaluate the efficacy and safety of BMS-986165 in subjects with SLE. Following a screening period of up to 4 weeks, approximately 360 subjects are to be randomized (1:1:1:1) via interactive response technology (IRT) to 1 of 4 treatments:

- BMS-986165 12 mg once daily (QD)
- BMS-986165 6 mg twice daily (BID)
- BMS-986165 3 mg BID
- placebo

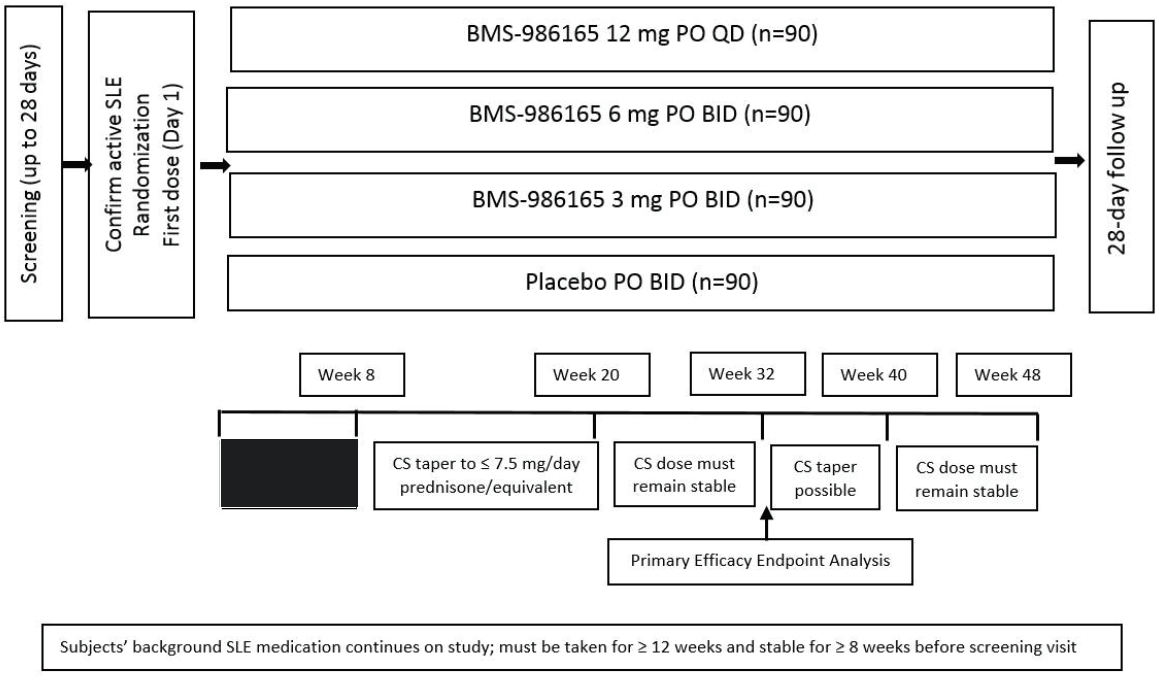
Randomization is planned to be stratified by incoming corticosteroid (CS) dose, SLE Disease Activity Index 2000 (SLEDAI-2K) score and region (see section 5.2 for details). The total duration of study participation for each subject is up to approximately 56 weeks (392 days), including up to a 4-week (28-day) screening period (SP), a 48-week (336-day) treatment period (TP), and a 4-week (28-day) follow-up period (FP) for subjects who do not continue into the long term extension (LTE) study (IM011074). The primary efficacy analysis will be based on Week 32 efficacy assessments. The study design is summarized in Figure 1.

Enrolled subjects will have received background therapy for at least 12 weeks including at a stable dose for at least 8 weeks before the screening visit, throughout the screening period and remain stable throughout the study. If on CS, the dose must be stable for at least 2 weeks before the screening visit, and throughout the screening period. A dose up to 30 mg/day prednisone or equivalent (as described in protocol Appendix 6) is permitted at the screening visit. Tapering of CS may begin at any time after the first dose on Day 1 and must start by Week 8. By Week 20 all subjects must have a CS dose of $\leqslant$ 7.5 mg prednisone or equivalent (if this is not possible, the subject is to remain on treatment, but will be a non-responder for analysis purposes). The CS dose must remain stable from Week 20 until Week 32 and tapering may resume at Week 32. The last day for CS taper is the Week 40 visit, and the CS dose must remain stable from Week 40 until Week 48. [redacted]

Details of physical examination, vital sign measurement, 12-lead ECG, and clinical laboratory evaluations are to be collected at all or at selected visits during the treatment and follow up period. Blood samples for PK and PD analysis are also to be collected for all subjects. [redacted]

A listing of measurements to be collected by visit are detailed in the schedule of events in Appendix 2.

Figure 1. Study Design Schematic



BID = twice daily; CS = corticosteroid; PO = per os (by mouth); SLE = systemic lupus erythematosus; QD = once daily

5.1 Sample Size Considerations

Approximately 360 subjects are to be randomized into 1 of 4 treatment groups in a 1:1:1:1 ratio, resulting in around 90 subjects per group. [redacted] With a sample size of 90 subjects per arm, there will be at least 84% power to detect a treatment difference of 20% in SRI(4) response rates between the BID BMS-986165 treatment arms and placeb [redacted].

5.2 Randomization

Approximately 360 subjects are planned to be randomized via IRT in a 1:1:1:1 ratio to 1 of 4 treatments: BMS-986165 3 mg BID, 6 mg BID, and 12 mg QD, or placebo. Randomization is stratified by screening CS dose (≥ 10 mg/day or < 10 mg/day), screening SLE Disease Activity Index 2000 (SLEDAI-2K) score (≥ 10 or < 10), and region (North America, Latin America, Japan [other randomization factors will not be applied in Japan], and rest of world).

5.3 Unblinding Information

Investigative site staff, Sponsor and designated personnel, and subjects and their families will remain blinded to treatment assignments. After the week 48 database lock, the Sponsor and designated personnel will be unblinded. Those subjects who are enrolled or will enroll in study IM011-074 will continue to remain blinded to their treatment assignment (including investigative site staff and subjects' families) in this study after the week 48 database lock. Those subjects who do not enroll or will not enroll in study IM011-074 may be unblinded to their treatment assignment (including investigative site

staff) after the week 48 database lock. However, if there is a medical emergency or pregnancy where knowledge of investigational product (IP) is critical to the subject's management the blind may be broken by the Investigator.

The Data Monitoring Committee (DMC) provides oversight of safety consideration throughout the study. A separate unblinded team, comprised of an unblinded Independent Reporting Statistician (IRS) and unblinded programmer(s), will produce output for the DMC using masked treatments. Treatment decodes may only be requested by DMC Chair and will be provided by the IRS. Data summaries and listings will be transmitted via a secure portal by IRS to only the DMC members. Additional details regarding the DMC process and unblinding are provided in the DMC charter.

Designated staff of BMS may be unblinded prior to database lock to facilitate the bioanalytical analysis of PK and biomarker samples and immunogenicity. [redacted]

5.4 Summary of Changes to Version 2.0 of SAP

No analyses were performed prior to the revision of this SAP.

5.4.1 Changes due to protocol amendments after finalization of SAP V1.0

[redacted]

2. Revised the testing strategy to be used for the primary and secondary endpoints (Section 9.3.1.4 of protocol)
3. Week 48 endpoints moved into the multiplicity-controlled hierarchy for better alignment with regulatory guidance documents
4. Non-responder imputation (NRI) section updated: updated to clarify that this NRI is applicable to the primary and secondary endpoints at Week 32 and Week 48.
5. Single randomization stratum combining Region (North America, Latin America, Rest of World, Japan [other stratification factors will not be applied in Japan]), CS dose at baseline (≥ 10 mg/day or < 10 mg/day) and SLEDAI-2K at baseline (≥ 10 or < 10) added for logistic regression, Cochran-Mantel-Haenszel and repeated measures analysis: Clarification of how to handle the lack of CS and SLEDAI-2K stratification within Japanese subjects
6. Revised the testing hierarchy for the QD and BID branches: Testing strategy revised to align with revision of the Week 48 endpoints to secondary endpoints [redacted]. The revised testing strategy ensures the familywise Type I error is protected for the primary and secondary endpoints.
7. The planned 32-week database lock has been removed: Since the key secondary endpoints will now be assessed at Week 48 no database lock will occur until all randomized subjects have completed 48 weeks of double-blind treatment (or have discontinued)

5.4.2 Additional Changes to SAP V2.0

1. Changes to the description of the composite endpoints for efficacy (aligned with source literature):
    a. SRI response:
        i. correction: no new BILAG A (severe disease activity) and (instead of "or" as indicated in the protocol) not more than 1 new BILAG B (moderate disease activity) organ domain grade since baseline
        ii. clarification: "new BILAG grade" should be interpreted as "worsened BILAG grade compared to baseline"

iii. correction: response only if PGA change from baseline <0.3 (instead of "<=0.3" as implied in the protocol description) (see Furie et al. Arthritis Rheum. 2009 61(9): 1143–1151; and Furie R, Petri MA, Strand V, et al. Clinical, laboratory and health-related quality of life correlates of Systemic Lupus Erythematosus Responder Index response: a post hoc analysis of the phase 3 belimumab trials. Lupus Science & Medicine 2014;1: e000031. doi:10.1136/lupus2014-000031)

b. BICLA response: "new BILAG grade" should interpreted as "worsened BILAG grade compared to baseline"
c. LLDAS response: "new BILAG grade" should interpreted as "worsened BILAG grade compared to previous assessment"

2. Added for clarification to the CTP: All hypothesis test will be based on Wald tests for odds ratio estimates for comparisons of the treatment groups and the placebo group obtained from the logistic regression models

[redacted]

4. Removal of per protocol analyses
5. Removal of all LOCF analyses
6. Modified sensitivity analyses, including addition of a treatment policy approach
7. Added a tipping point analysis for the primary endpoint

[redacted]

9. Various changes to format and structure to improve clarity

[redacted]

11. Refinement of NRI imputation method including the list of prohibited medications
12. Updated PK and Biomarker analyses

5.5 Summary of Changes to Version 3.0 of SAP

No analyses were performed prior to the revision of this SAP.

1. Added clarification for the first dose date in section 7.1
2. Added one additional condition in sections 11.7.1 and 11.7.2.2:
    a. Subjects takes any medications included in protocol Appendix 7 between first dose date and efficacy assessment visit
3. Deleted one condition in sections 11.7.1 and 11.7.2.2:
    a. Initiations and/or increases which occur for less than 4 days are permitted
4. Deleted the bar chart in sections 11.7.1, 11.7.2.3 and 11.7.2.4
5. In section 11.7.1, in case the logistic regression model fit is not stable the region stratification factor may be removed from the model. The stratification factors were updated to: CS dose ≥ 10 mg/day Yes or No, SLEDAI-2K score ≥ 10 Yes or No, and Japan.
6. Added clarification for analysis time points in section 11.7.2.5
7. Changed "interferon regulated genes" to "interferon" in section 11.7.4
8. Added "at baseline" for subgroup analyses referring to CS Use, anti-malarial use, and immunosuppressant use in section 11.7.4

[redacted]

10. Added clarification for analysis time points in section 11.7.5
11. Deleted the following condition in section 11.8.4:

a. No other immediately apparent possible causes of liver function test elevation and hyperbilirubinemia, including, but not limited to, viral hepatitis, preexisting chronic or acute liver disease, or the administration of other drug(s) known to be hepatotoxic.

12. Modified the QTcF calculation in section 11.8.6
13. Updated the pharmacokinetic analyses in section 12.0
14. Updated the biomarker analyses in section 13.0
15. Updated the biomarker listing in appendix 6

5.6 Summary of Changes to Version 4.0 of SAP

No analyses were performed prior to the revision of this SAP.

1. Deleted “between visits as collected from ‘Rescue Medication’ eCRF page” in section 11.5.2 [redacted]

6.0 **Study Variables**

6.1 Efficacy Assessments

The following efficacy procedures or tools will be used to assess subjects’ SLE disease activity during the study:

British Isles Lupus Assessment Group (BILAG)-2004

The BILAG-2004 grades are based on recording disease activity for 97 items in 9 different organ systems (constitutional, mucocutaneous, neuropsychiatric, musculoskeletal, cardiorespiratory, gastrointestinal, ophthalmic, renal, and hematological, see Appendix 10 of the study protocol). Each item is measured qualitatively by clinical observation (yes/no, not present/improving/same/worse/new) based on lupus activity over the past 4 weeks compared with the previous 4 weeks or quantitatively by measuring hematologic and renal lab values. For abnormal lab values, the item is noted as SLE related or not SLE related. Based on scores of the 97 items, an alphabetical grade is derived for each organ system representing 5 levels of disease with A being the highest level of disease activity. The grades are given a numerical score equivalent:

- A (very active) = 12
- B (moderate disease activity) = 8
- C (mild stable disease) = 1
- D (no current disease activity but the system had previously been affected) = 0
- E (no current or previous disease activity) = 0

A global BILAG-2004 score (range from 0 to 108) is derived by summing-up the numerical score equivalents for each organ system. Grades are assigned by central reviewers [redacted] based on BILAG-2004 Index Scoring along with grading conventions. The detailed grading conventions are documented in the Central Review Plan including the handling of missing data when assigning the grades. Grades assigned in ‘BILAG Grading’ eCRF page will be used for analysis purpose.

SLEDAI-2K

The SLEDAI-2K is a global index based on weighted scores for each of 24 clinical findings rated as present or absent at the time of the visit or in the last 30 days (see Appendix 11 of the Study Protocol). The total score for SLEDAI-2K ranges from 0 to 105 with higher scores representing higher disease activity. The total score collected in ‘SLEDAI-2K’ eCRF page will be used for analyses. The detailed grading conventions are documented in the Central Review Plan including the handling of missing data when assigning the grades.

Cutaneous Lupus Erythematosus Disease Area and Severity Index (CLASI)

CLASI assesses cutaneous disease activity and damage by body surface area (see Appendix 12 of the study protocol). Only the CLASI activity score derived in 'Cutaneous LE Disease Area and Severity Index (CLASI)' eCRF page will be used in this study.

For disease activity, points are given for presence of erythema, scale, hypertrophy, mucous membrane lesions, recent hair loss, and physician-observed alopecia and a total activity score is derived. The CLASI total activity score (CLASI-A) ranges from 0-70, where a higher score is associated with high disease activity.

For the damage assessments, points are given for dyspigmentation, scarring, and scarring alopecia. A total damage score is derived. The CLASI total damage score (CLASI-D) ranges from 0 – 56, where a higher score is associated with larger extent of damage.

40-joint Count

Each of 40 joints is evaluated for the presence/absence of tenderness, swelling, and simultaneous presence/absence of both tenderness and swelling. The following joints will be evaluated: the bilateral wrists, elbows, ankles, knees, interphalangeal joints of the thumb, individual proximal interphalangeal joints of the hand, second through fifth metacarpophalangeal joints of the hand, and individual metatarsophalangeal joints of the feet (which make up the 36-joint count). Bilateral first metacarpophalangeal joints and shoulders are also included, bringing the joint count to 40. From the above assessment, joint counts will be calculated for:

- Tender joints (0-40)
- Swollen joints (0-40)
- Tender and swollen joints (0-40): joints that are both tender and swollen (= active) are counted.

A larger joint count indicates more severe disease. The total joint counts collected from the eCRF will be used for analysis.

Physician's Global Assessment (PGA) of Disease Activity

The Physician's Global Assessment of Disease Activity (PGA) will be assessed using a 3-point visual analog scale (VAS) as the subject's average disease activity over the past 30 days. A score of 0 represents no disease activity and a score of 3 represents severe disease activity. An approximate score at tenth intervals between 0 and 3 (inclusive) will be reported.

[redacted]

[redacted]

6.2 Composite Measures

The following composite measures will be determined based on assessments described above.

BILAG-based Composite Lupus Assessment (BICLA)

BICLA response is achieved if all of the following criteria are met:

- Improvement in all organ systems with activity graded as BILAG-2004 A or B at baseline
- No new organ system with activity graded as BILAG-2004 A and no more than 1 new organ system with activity graded as BILAG-2004 B since baseline
- No increase from baseline in SLEDAI-2K (change from baseline of $\leqslant$ 0 score)
- No increase from baseline $\geqslant$ 10% (or $\geqslant$ 0.3 points) in the Physician's Global Assessment of Disease Activity on a 3-point VAS
- No discontinuation of IP or use of restricted medications beyond the protocol allowed threshold before assessment (refer to section 11.7.2 with more details)

SLE Responder Index (SRI[X])

The SRI is a composite responder endpoint. SRI(X) response is achieved if all of the following criteria are met:

- Reduction from baseline of ≥X points in the SLEDAI-2K
- No new BILAG A (severe disease activity) and not more than 1 new BILAG B (moderate disease activity) organ domain grade since baseline
- No worsening from baseline in the Physician's Global Assessment of Disease Activity scale by > 0.3 points

Lupus Low Disease Activity State (LLDAS)

LLDAS response is achieved if all of the following criteria are met:

- SLEDAI-2K ≤ 4, with no activity in major organ systems (renal, central nervous system, cardiopulmonary, vasculitis, fever) and no hemolytic anemia or gastrointestinal activity (measured as maintaining a D or E score in the BILAG gastrointestinal body system)
- No new lupus disease activity compared with the previous assessment measured as no new or worsening individual BILAG parameters
- Physician's Global Assessment of Disease Activity ≤ 1
- Prednisone or equivalent dose ≤ 7.5 mg/day
- Well tolerated standard maintenance doses of immunosuppressive drugs and approved biologic agents (refer to section 11.7.2 with more details)

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

6.4 Safety

Safety data to be collected includes physical examination and measurements, vital signs,12-lead ECG, hematology, clinical chemistry, urinalysis, coagulation, fasting lipid and glucose, pregnancy test results, concomitant medication use, and incidence and severity of all reported AE and SAEs. The planned time points for all safety assessments are listed in the Appendix 2. The change from baseline at each scheduled time point during the treatment period will be calculated for relevant measures.

6.5 Predetermined Covariates and Prognostic Factors

The following stratification factors are applied in this study:

- CS dose (≥ 10 mg/day or < 10 mg/day)
- SLEDAI-2K score (≥ 10 or < 10)
- Region (North America, Latin America, Japan, rest of world)

However, as CS dose and SLEDAI-2K score are not applied in Japan, the levels for the factor that is used in the analysis to represent the randomization strata is redefined as follows:

1. CS dose = '≥ 10 mg/day', SLEDAI-2K score = '≥ 10', Region= ‘North America’;
2. CS dose = '≥ 10 mg/day', SLEDAI-2K score = '≥ 10', Region= ‘Latin America’;
3. CS dose = '≥ 10 mg/day', SLEDAI-2K score = '≥ 10', Region= ‘rest of world’;
4. CS dose = '≥ 10 mg/day', SLEDAI-2K score = '< 10', Region= ‘North America’;
5. CS dose = '≥ 10 mg/day', SLEDAI-2K score = '< 10', Region= ‘Latin America’;
6. CS dose = '≥ 10 mg/day', SLEDAI-2K score = '< 10', Region= ‘rest of world’;
7. CS dose = '< 10 mg/day', SLEDAI-2K score = '≥ 10', Region= ‘North America’;
8. CS dose = '< 10 mg/day', SLEDAI-2K score = '≥ 10', Region= ‘Latin America’;
9. CS dose = '< 10 mg/day', SLEDAI-2K score = '≥ 10', Region= ‘rest of world’;
10. CS dose = '< 10 mg/day', SLEDAI-2K score = '< 10', Region=‘North America’;
11. CS dose = '< 10 mg/day', SLEDAI-2K score = '< 10', Region=‘Latin America’;
12. CS dose = '< 10 mg/day', SLEDAI-2K score = '< 10', Region=‘rest of world’;
13. CS dose = Not Applicable, SLEDAI-2K score = Not Applicable, Region=‘Japan’;

The above redefined stratification factor collected from IVRS system will be included in statistical analysis models or be used to examine treatment effect. For subgroup analysis, the value of the above redefined stratification factor collected on the eCRF will be used.

6.6 Pharmacokinetic and Pharmacodynamic Variables

The pharmacokinetics of BMS-986165 and its metabolite BMT-153261 will be derived from plasma concentration versus time data. Biomarker data will be collected through serum, plasma, and urine samples to gain further understanding of the PD effects of BMS-986195. These samples will be analyzed for markers of tissue damage and regeneration including serum complement, anti-dsDNA and other autoantibodies, blood RNA for interferon-regulated gene (IRG) and steroid signature, genome-wide expression, and inflammatory markers.

Blood samples for pharmacokinetics and interferon-regulated gene (IRG) assessment will be collected predose and postdose at pre-specified study days and times. Actual sample times will be recorded and missed samples will be noted.

For a substudy of approximately 120 subjects, samples will be collected anytime on Days 2 and 3 for blood RNA and inflammatory markers.

Blood flow cytometry data will also be collected in a substudy defined by geography and accessibility to the analyzing lab. Study will occur in USA and Europe.

7.0 **Definitions**

7.1 General

The following data handling conventions will be used for general analysis:

Term	**Definition**
Study Day	The assessment date – date of first dose + 1. When the assessment date is before the date of first dose then study day is calculated as the assessment date - date of first dose.
Baseline	Baseline is defined as the last non-missing measurement prior to dosing on Day 1 (Week 0). If the measurement on Day 1 is missing or not available, then the last available measurement during the screening period will be used as a baseline.
Baseline CS Dose Categorized	For subjects taking CS at screening, the dose (prednisone or equivalent) must be stable for at least 2 weeks before the screening visit and must remain stable until randomization. The last CS dose during the screening period will be considered the baseline dose. For subjects not taking CS at screening, the baseline dose will be 0. Baseline CS Dose will be categorized as $\geqslant$ 10 mg/day or < 10 mg/day prednisone or equivalent for stratification. Baseline CS dose categorization will be from baseline CS dose eCRF form.
Change from Baseline	Change from baseline is defined as (value at post-baseline visit – value at baseline).
Concomitant and Prior Medication	Prior medications are defined as medications with a start or stop date prior to the first dose of study treatment. Concomitant medications are defined as any medication that occurred (continued, started, or stopped) on or after first dose of study medication and on or before last dose of study medication.
End of Study (EOS) Date	The EOS date is the date recorded on the eCRF that a randomized subject either discontinued or completed the study. If the subject is lost to follow-up, the EOS date will be the date of the last visit assessment obtained.
First Dose Date	The date a subject received their first dose on Day 1 as recorded in the 'Exposure - PK Sampling' eCRF form, given that the first dose should be entered even if no PK sampling is collected at Day 1. Or if the first dose date is not collected in the 'Exposure - PK Sampling' eCRF form, the earliest drug dispensation date should be used.
Last Dose Date	The Date of Last Dose of Study Treatment on 'End of Treatment' eCRF form for a randomized subject.
Percent Change from Baseline	Percent change from baseline is defined as ([value at post-baseline visit – value at baseline]/value at baseline) x 100. If the baseline value is 0 and the post-baseline value is also 0, then the percent change from baseline is set to 0.

7.2 Analysis Visits Windows

Subjects do not always adhere strictly to the visit schedule timing in the protocol. Therefore, the designation of visits during the study period will be based on the day of evaluation relative to the Day 1 of the trial (day of first dosing = study Day 1) rather than the nominal visit recorded in the case report form (CRF). Visit window based on the study day (regardless the visit labeling) will be applied for efficacy analysis and safety analysis. Each efficacy and safety measurement will be windowed. In case that two or more measurements exist in a certain visit window, the one closest to the target day will be selected. If still more than one measurements are selected, the later one will be selected for analysis.

Efficacy measurements up to last dose will be used for windowing.

Analysis Visit	Target Day	Study Day Range
Baseline	1	[-28, 1]
Week 2	15	[2, 22]
Week 4	29	[23, 43]
Week 8	57	[44, 71]
Week 12	85	[72, 99]
Week 16	113	[100, 127]
Week 20	141	[128, 155]
Week 24	169	[156, 183]
Week 28	197	[184, 211]
Week 32	225	[212, 239]
Week 36	253	[240, 267]
Week 40	281	[268, 295]
Week 44	309	[296, 323]
Week 48	337	[324, 351]
Safety Follow-up	EOT + 30	

8.0 **Populations for Analyses**

The following populations will be used in the summary and analysis of study data.

8.1 Enrolled

The Enrolled population is defined as all subjects who signed informed consent.

8.2 Randomized

The Randomized population is defined as all subjects who are randomized to a treatment. Subjects will be analyzed as per randomized treatment according to the IRT.

8.3 All-treated

The All-treated population consists of all randomized subjects who receive at least one dose of any study treatment. A subject will be considered to have received their randomized treatment unless they receive the incorrect treatment for the entire duration of the double-blind treatment period.

8.4 Biomarker

The Biomarker population is defined as all randomized subjects who receive at least one dose of study treatment and have at least one post-treatment biomarker measurement. Subjects will be analyzed according to the treatment received.

8.5 Pharmacokinetic

The Pharmacokinetic (PK) population is defined as all randomized subjects who receive at least one dose of BMS-986165 and have any available concentration data. The evaluable PK population is a subset of the PK Population that includes all subjects who have adequate PK profiles or have at least one evaluable PK parameter.

The bioanalytical lab will receive the true randomization file and will only analyze the plasma samples from the subjects who received BMS-986165 drug. Subjects will be analyzed according to the treatment received.

9.0 **Analysis and Reporting**

The final (and only) database lock will occur once all randomized subjects have completed 48 weeks of double-blind treatment (or have discontinued earlier) and the 4-week safety follow-up for those who have not enrolled into the LTE study IM011074.

10.0 **Data Review**

10.1 Data Handling and Transfer

Data will be entered and exported as SAS® version 9.4 or higher datasets. Converted datasets will be created using SAS® and following standard Clinical Data Interchange Standards Consortium Standard Data Tabulation Model (CDISC SDTM, version 1.4 or higher, Implementation Guide version v3.2 or higher) conventions. Analysis datasets will be created using SAS® and following CDISC Analysis Data Model (ADaM, version 2.1 or higher, Implementation Guide 1.1 standards or higher).

Additional details can be found in the [redacted] Data Management Plan for this study.

11.0 **Statistical Methods**

Unless otherwise noted, variables will be reported using summary statistics according to the below methods by treatment group and analysis visit defined in section 7.2.

Categorical variables will be summarized using frequency distributions (counts and percentages). Percentages will be rounded to one decimal place, and percentages will not be displayed for zero counts.

For continuous variables summary statistics will include the number of observations (n), mean, standard deviation (SD), median, minimum, maximum, 25th percentile (Q1), and 75th percentile (Q3). The median, minimum, maximum, Q1, and Q3 values will be displayed to the same level of precision as the raw data, the mean to a further decimal place and the SD to two additional decimal places.

P-values will be displayed to 3 decimal places, with values less than 0.001 presented as <0.001.

All data will be listed by subject, analysis visit and treatment group.

All statistical analysis and reporting will be done using SAS® Version 9.4 or higher (SAS Institute, Inc.).

11.1 Subject Disposition

The number of subjects enrolled/screened, reasons of screen failure and the number and percentage of subjects randomized, treated, and in each analysis population will be presented.

Additionally, the following summaries will be provided by treatment group and overall:

- Number and percentage of subjects enrolled, randomized and treated by geographic region (North America, Latin America, Japan, rest of world), country, and center
- Number and percentage of subjects who completed the study, discontinued the study, reason for study discontinuation, completed the treatment phase, discontinued treatments, and reason for treatment discontinuation
- Number and percentage of subjects on treatment at each analysis visit

Subject disposition data will be taken from the End Of Treatment and End Of Study CRF forms.

11.2 Demographic and Baseline Characteristics

Demographic characteristics will be reported by treatment group for both the randomized, and all-treated populations using summary statistics. Demographic characteristics include the following:

- Geographic Region (North America, Latin America, Japan, rest of world)
- Sex
- Race
- Ethnicity
- Age (in years, at time of signing informed consent) and age category (<65 vs ⩾65)
- Baseline CS Use (yes, no)
- Baseline Anti-malarial Use (yes, no)
- Baseline Immunosuppressant Use (yes, no)

[redacted]

- Oral CS Dose (observed dose as categories in mg/day, prednisone or equivalent)
- Oral CS Dose (⩾ 10 mg/day vs < 10 mg/day; prednisone or equivalent)
- Weight (in kg)
- Height (in m)
- Body mass index (BMI in kg/m^2)
- Number and percentage of subjects within each geographic region categorized by baseline CS dose (⩾ 10 mg/day or < 10 mg/day) and baseline SLEDAI-2K score (⩾ 10 or < 10) as collected on eCRF and on IRT

The following baseline disease characteristics will be summarized by treatment group for the randomized population:

- BILAG-2004 grade by organ system
- BILAG-2004 global score
- Categorization of BILAG-2004 A/B grades (at least one A, no A and at least 2 Bs, No A and < 2 Bs)
- SLEDAI-2K score
- SLEDAI-2K (⩾ 10 or < 10)
- CLASI-A score
- swollen joint count
- tender joint count
- count of joints that are both tender and swollen (= active joints)

[redacted]

- Physician's Global Assessment of disease activity

11.3 Relevant Protocol Deviations

Relevant protocol deviations considered to have a potential impact on the primary or key secondary efficacy endpoint assessments are:

- Failure to meet any one or more inclusion criteria that define the patient population
- Failure to adhere to concomitant medication restrictions affecting the primary endpoint
- Poor compliance to study medication, i.e. overall compliance <75% for treatment period up to Week 32
- Mis-dosed (treated dose not equal randomized dose)

The detailed relevant protocol deviation criteria will be listed in Appendix 5.

11.4 Exposure

Summaries of extent of study drug exposure will be provided for the all-treated population.

11.4.1 Duration of Exposure

Duration of treatment will be calculated using the date of first dose of study drug and the date of last dose of study drug as recorded on the eCRF. Duration of treatment, in days, will be calculated according to the below formula:

Date of last dose of study drug - date of first dose of study drug + 1

Duration of treatment will be reported by treatment group using summary statistics.

11.4.2 Summary of Dosing

A single dose is defined as the randomized treatment dose. Subjects are to take a total of 8 capsules per day to equal one dose.

The total number of doses taken for each subject will be determined as the following:

Doses taken = (number capsules dispensed – number of capsules returned)/8

The total number of doses taken will be reported using summary statistics by treatment group.

11.4.3 Compliance

Treatment compliance will be determined from data captured in the eCRF.

The number and percentage of subjects who have missed at least one dose will be provided by treatment group. Additionally, summary statistics for the number of missed doses within each analysis visit period and overall will be provided by treatment group. The number of missed doses for each subject will be calculated as the following:

((8 x number of days in the analysis visit period) – (Number of capsules given – number of capsules returned)) /8

Overall compliance will be calculated based on eCRF data as specified below.

1. Number of doses taken = (number capsules dispensed – number of capsules returned)/8
2. Number of expected doses taken = last dose date – first dose date + 1
3. Compliance = (Number of doses taken / Number of expected doses taken) * 100

If a subject does not return the container, then this dispensation event will be excluded in the calculation for the total number of doses taken, total number of expected doses, and total compliance. Overall compliance for the treatment period up to the Week 32 and up to Week 48 will be reported using summary statistics by treatment group and for the combined treatment groups. The number and percentage of subjects with <75%, 75% to 100%, and >100% compliance will be reported.

11.5 Prior and Concomitant Medications

Prior and concomitant medications will be coded according to World Health Organization-Drug Dictionary WHO-DD (June 2017 DDE HD B3 or an updated version at the time of database lock (DBL)) and will be summarized by Anatomic Therapeutic Classification (ATC) level 1, level 2 and preferred term (PT) by treatment group for the randomized population. The number and percentage of subjects using each medication will be displayed by treatment group. Subjects taking more than one medication within the same ATC level and/or PT will be counted once. Details of prescribing information will be provided in a listing.

Medication dates will be imputed according to algorithms detailed in Appendix 3. The imputed dates will be used to assess whether medications should be included in the summaries as prior or concomitant, however the original, partial dates will be included in data listings.

Only medications collected in the study database from first dose date of IP up to last dose day of IP or up to first dose day of study IM011-074 for roll over subjects will be summarized for concomitant medications.

11.5.1 Prior Corticosteroid Use

Prior systemic CS medications (prednisone or equivalent) will be summarized as described above for the number and percentage of subjects using each CS medication. A summary of other prior medications used for SLE may also be presented.

The list of used CS medications to be selected for analyses is available in Appendix 6 of the protocol.

Prior corticosteroid use data will be taken from the Prior and Concomitant Medications CRF form.

11.5.2 Concomitant Corticosteroid Use

Concomitant oral CS medications (prednisone or equivalent) will be summarized as described above for the number and percentage of subjects using each CS medication as well as continuous summaries of CS dose. The frequency and percentage of successful CS tapering and maintenance (defined as CS dose $\leqslant$ 7.5 mg/day prednisone or equivalent by Week 20 continuing through Week 32 and no CS rescue therapy) will be presented by treatment group as well a summary of the reasons CS was not reduced; due to a new or worse BILAG grade, a new or worse SLEDAI-2K score, new or worse hemolytic anemia, a CLASI-A score of at least 10, a tender and swollen joint count of at least 8, or others as collected on the Corticosteroid Dose Log CRF and Steroid Taper CRF forms. The need for CS rescue will be summarized with percentages of bursts <=1 or >1. A listing will provide details of CS therapy use throughout the study. A blinded review of the corticosteroid use will be conducted to assess whether a subject received a CS burst. The logic for the determination of a CS burst and a spreadsheet with the CS bursts identified will be saved in the eTMF. This spreadsheet will be used for programming of CS bursts.

The list of used CS medications to be selected for analyses is available in Appendix 6 of the protocol. Concomitant corticosteroid use data will be taken from the Corticosteroid Dose Log CRF forms.

11.6 Medical History

General medical history and medical history related to SLE will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) (an updated version at the time of DBL). General medical history data will be summarized for each SOC and PT by treatment group and overall for the randomized population. Separate tables will be provided for general medical history and medical history related to SLE. Details of medical history will be provided in a listing. Medical history data will be taken from the General Medical History CRF.

Additionally, BILAG SLE history (captured at screening) will be summarized for each body organ system by treatment group and overall for the randomized population.

11.7 Efficacy Analyses

All efficacy analyses will be performed using the randomized population.

The final statistical analysis will be performed at the end of study after all subjects have completed their end of study visit (or discontinued early from the study).

Separate methods will be used for binary, continuous, time-to-event and categorical efficacy endpoints.

Subjects listings will report all collected efficacy endpoints over time until study end or early study discontinuation, regardless of treatment discontinuation before the study end or discontinuation.

Binary endpoints for analysis are:

- SRI(4)[redacted]
- BICLA response
- LLDAS response
- CLASI-A response

[redacted]

Continuous endpoints for analysis are:

- Change from baseline in the 40-joint count for tender joints, swollen joints, and joints that are both tender and swollen (= active joints)
- Global BILAG-2004 score

[redacted]

Categorical endpoints:

[redacted]

11.7.1 Primary Endpoint

The primary efficacy endpoint is the proportion of subjects who have a SRI(4) response at Week 32 as per non-responder imputation. All subjects in the randomized population will be categorized as a responder or non-responder.

Definition of Endpoint

A responder is defined as meeting all of the following criteria. If one criterion is not met, the subject will be considered a non-responder.

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

  The total SLEDAI-2K score is the sum of the scores for all items. The total score will be taken from the eCRF. A reduction in at least 4 points is achieved if SLEDAI-2K at Week 32 – SLEDAI-2K at baseline $\leq$ -4.

- No new BILAG A (severe disease activity) and not more than 1 new BILAG B (moderate disease activity) organ domain grade

  An A grade is new if the baseline grade for the same organ system was B, C, D or E. A B grade is new if the baseline grade for the same organ system was C, D, or E.

  This criterion is met if both of the following conditions are met:
  - All organ domains with an A grade at Week 32 were graded A at baseline
  - The number of organ domains with a B grade at Week 32 that were graded C, D, or E at baseline is $\leq$ 1

- No worsening from baseline in the Physician's Global Assessment of Disease Activity scale by at least 0.3 points
  - Physician's Global Assessment of Disease Activity at Week 32 – Baseline Physician's Global Assessment of Disease Activity < 0.3

Non-Responder Imputation (NRI) Method

Subjects will be considered non-responders at a given time point per imputation (i.e. regardless of observed SRI-response as defined above) if any of the following apply:

- Subject takes any of the following prohibited medications:
  - Corticosteroid dosages above 7.5mg/day prednisone or equivalent after week 20 (Week 20+1 day)
  - Use of IM or IV CS within 8 weeks before the W32 assessment timepoint
  - Use of IA or intrabursal CS within 8 weeks before the W32 assessment timepoint

- Use of modified-release oral CS formulations within 8 weeks before the W32 assessment timepoint
- Addition (compared to baseline) of topical CS within 2 weeks before the W32 assessment timepoint
  - Ocular CS are allowed for non-SLE indications
  - Inhaled CS are allowed
- CS bursts/rescue above the protocol-defined maximum dosage (40 mg/day prednisone equivalent)
- More than 1 CS burst or any CS burst for treatment of SLE after week 8
  - Oral steroid burst for non-SLE indication is permitted up to week 24 or between week 32 and week 40
- Initiation of or increases in oral NSAIDs (compared to baseline) occurring within the 7 days before the W32 assessment timepoint
- Initiation of or increase in immunosuppressant dosage/regimen and/or antimalarial dosage/regimen (compared to baseline)
  - Dose decrease or discontinuation is permitted
  - Discontinuation with subsequent restart at the same or lower doseis permitted
  - Immunosuppressants are defined as (combinations of these are not permitted): Azathioprine, 6-MP, methotrexate, luflunomide, MMF
  - Antimlalarials are defined as: chloroquine, hydroxychloroquine, or quinacrine

- Subjects takes any medications included in protocol Appendix 7 between first dose date and efficacy assessment visit
- Permanent treatment discontinuation of IP before the considered time point
- Missing value for the endpoint at the considered time point
- Subjects who are lost to follow-up or discontinue the study early are considered non-responders for all time points at and after loss to follow up or study discontinuation

Primary Analysis

The primary analysis will be conducted on the randomized population.

The analysis of NRI-imputed SRI(4) response at Week 32 will be conducted using logistic regression modeling, with treatment group (four levels) and redefined stratification factor (see section 6.5, thirteen levels) as fixed factors. The odds ratios of response will be estimated and tested between the BMS-986165 treatment groups and the placebo group [redacted]. In case the logistic regression model fit is not stable the region stratification factor may be removed from the model (resulting in 5 stratification factors: non-Japan subjects with baseline CS dose ≥ 10 mg/day and SLEDAI-2K score ≥ 10, non-Japan subjects with baseline CS dose < 10 mg/day and SLEDAI-2K score ≥ 10, non-Japan subjects with CS dose ≥ 10 mg/day and SLEDAI-2K score < 10, non-Japan subjects with CS dose < 10 mg/day and SLEDAI-2K score < 10 and Japan subjects). Estimated odds ratios of response with corresponding 2-sided 95% confidence interval (CI) and p-value will be reported.

The number and percentage of subjects achieving NRI-imputed SRI(4) response at Week 32 and those who do not will be reported by treatment group with the estimated response rate and asymptotic 95% confidence intervals (CIs). Differences in response rates between each active treatment group and placebo with corresponding 2-sided 95% CIs based on the continuity-corrected Newcombe method will also be provided.

Tipping Point Analysis

Tipping point analysis will further be used to assess the robustness of the primary study results. Different number of events between treatment groups will be assessed until the study conclusion is

changed. Each imputed value is initially imputed as a responder. The imputed values in the placebo group will remain as responders while the BMS-986165 imputed values are replaced as a non-responders one at a time, therefore changing the number of events between groups. Once all imputed values in the BMS-986165 group have been replaced with non-responders values, the data will reset to where the BMS-986165 group imputed values are all responders and one by one the imputed values in the placebo group are replaced with a non-responder until all placebo are non-responders and all BMS-986165 are responders. Furthermore, every pair of imputations between the placebo and BMS-986165 groups will be assessed similarly, creating a matrix of possible patterns.

At each iteration, the statistical analysis is assessed, and the direction of the analysis is recorded. A graph will be provided to identify at what point to which the difference in events cause a direction shift in study results. The tipping point analyses will be based on a two-sample chi-square test approach rather than a stratified CMH test approach for simplicity of the analysis.

Sensitivity analysis #1 (Treatment Policy Strategy)

The primary analysis for the primary endpoint will be repeated using a treatment policy estimand approach (Intention-To-Treat (ITT) principle). For this approach, subjects will only be imputed as non-responders if there is a missing value for the endpoint at the considered timepoint.

Sensitivity analysis #2

The primary analysis for the primary endpoint will be repeated, however the NRI imputation method will be modified. The prohibition medications restriction for the NRI imputation method will not apply. That is, subjects who take any of the prohibited medications listed above will not be imputed as non-responders. Subjects will continue to be imputed as non-responders if any of the following apply:

- Permanent treatment discontinuation of IP before the considered time point
- Missing value for the endpoint at the considered time point
- Subjects who are lost to follow-up or discontinue the study early are considered non-responders for all time points at and after loss to follow up or study discontinuation

Supportive Analysis #1 – Assessment of Covid-19 related impact

For this supportive analysis, the primary analysis for the primary endpoint will be repeated, but observations that were missed due to the impact of COVID-19 will be imputed with the last value observed before the visit that was missed due to COVID-19 restrictions.

11.7.2 Secondary Endpoints

11.7.2.1 SRI(4) Response at Week 48

SRI(4) response at week 48 is defined in an analogous way as SRI(4) response at week 32 (see details in 12.7.1).

Similar analyses will be conducted as for the primary endpoint, including the primary analysis and sensitivity analysis.

11.7.2.2 BICLA Response at Week 48

All subjects in the randomized population will be categorized as a responder or non-responder. A subject will be considered a BICLA Week 48 responder if all of the following criteria are met. If one criterion is not met, the subject will be considered a BICLA Week 48 non-responder.

- Improvement in all organ systems with activity graded as BILAG-2004 A or B at baseline
    - This criterion is met if both of the following conditions are met:All organ domains with an A grade at baseline are graded B, C, D, or E at Week 48
    - All organ domains with an B grade at baseline are graded C, D, or E at Week 48

- No new organ system with activity graded as BILAG A and no more than 1 new organ system with activity graded as BILAG B
  An A grade is new if the baseline grade for the same organ system was B, C, D or E. A B grade is new if the baseline grade for the same organ system was C, D, or E.

  This criterion is met if both of the following conditions are met:
    - There are no organ domains with an A grade at Week 48
    - The number of organ domains with a B grade at Week 48 that were graded C, D, or E at baseline is $\leqslant$ 1

- No increase from baseline in SLEDAI-2K ($\leqslant$ 0 points for change from baseline score)

  SLEDAI-2K score at Week 48 - SLEDAI-2K score at baseline $\leqslant$ 0

- No increase $\geqslant$ 10% (or $\geqslant$ 0.3 points) in the Physician's Global Assessment of Disease Activity on a 3-point VAS:

  Physician's Global Assessment of Disease Activity at Week 48 – Baseline Physician's Global Assessment of Disease Activity < 0.3

- No discontinuation of IP or use of prohibited/restricted medications beyond the protocol allowed threshold before assessment.
    - No discontinuation of IP is defined as remaining on IP within 14 ($\leqslant$ 14) days prior to the Week 48 visit. Treatment discontinuation date is captured in the eCRF.
    - Restricted/prohibited medications are defined as in section 11.7.1 and are:
        - Corticosteroid dosages above 7.5mg/day prednisone or equivalent after week 20 (Week 20+1 day)
        - Use of IM or IV CS within 8 weeks before the W32 and W48 assessment timepoints
        - Use of IA or intrabursal CS within 8 weeks before the W32 and W48 assessment timepoints
        - Use of modified-release oral CS formulations within 8 weeks before the W32 and W48 assessment timepoint
        - Addition of topical CS within 2 weeks before the W32 and W48 assessment timepoint
            - Ocular CS are allowed for non-SLE indications
            - Inhaled CS are allowed

        [redacted]

        - More than 1 CS burst or any CS burst for treatment of SLE after week 8

- Oral steroid burst for non-SLE indication is permitted up to week 24 or between week 32 and week 40

- Initiation of or increases in oral NSAIDs occurring within the 7 days before the W32 and W48 assessment timepoints
- Initiation of or increase in immunosuppressant regimen and/or antimalarial dosage/regimen
  - Dose decrease or discontinuation is permitted
  - Discontinuation with restart at the same or lower dose is permitted
  - Immunosuppressants are defined as (combinations of these are not permitted): Azathioprine, 6-MP, methotrexate, luflunomide, MMF
  - Antimlalarials are defined as: chloroquine, hydroxychloroquine, or quinacrine

- Any medications included in protocol Appendix 7 between first dose date and efficacy assessment visit

Similar analyses will be conducted as for the primary endpoint, including the primary analysis and sensitivity analysis.

11.7.2.3 CLASI-A Response (CLASI-50) at Week 48

Subjects who have a baseline CLASI activity score $\geqslant$10 will be included in this analysis. Subjects meeting this criterion will be categorized as CLASI-A responders or non-responders. Responders are defined as subjects who have at least a 50% decrease in their CLASI activity score at Week 48.

$$\frac{\text{CLASI activity score at Week 48} - \text{CLASI activity score at baseline}}{\text{CLASI activity score at baseline}} \times 100 \leq -50$$

Subjects not meeting this criterion will be considered non-responders.

Similar analyses will be conducted as for the primary endpoint, including the primary analysis and sensitivity analysis.

11.7.2.4 LLDAS Response at Week 48

A subject is considered an LLDAS responder at Week 48 if all of the following criteria are met:

- SLEDAI-2K $\leqslant$ 4, with no activity in major organ systems (renal, central nervous system, cardiopulmonary, vasculitis, fever) and no hemolytic anemia or gastrointestinal activity (measured as maintaining a D or E score in the BILAG gastrointestinal body system)

  This definition will be implemented as follows:
  - SLEDAI-2K $\leqslant$ 4
  - The following items on the SLEDAI-2K are scored as 'no': seizure, psychosis, organic brain syndrome, visual disturbance, cranial nerve disorder, lupus headache, CVA (cerebrovascular accident), vasculitis, urinary casts, hematuria, proteinuria, pyuria, pleurisy, pericarditis, and fever
  - No hemolytic anemia on BILAG (item 96 on BILAG)
  - All BILAG gastrointestinal items scored as 0 (all items under gastrointestinal organ system graded D or E)

- No new lupus disease activity compared with the previous assessment measured as no new or worsening individual BILAG parameter

  This definition will be implemented as follows:
  - None of the 97 BILAG items are scored 4 (=worse) or 5 (=new)
  - None of the 97 BILAG items are newly marked abnormal SLE related

If an item is rated Abnormal SLE Related at a visit then it must have been rated Abnormal SLE related at the previous visit.

- Physician's Global Assessment of Disease Activity $\leqslant$ 1
- Prednisone or equivalent dose $\leqslant$ 7.5 mg/day
- Well tolerated standard maintenance doses of immunosuppressive drugs and approved biologic agents

[redacted]

Similar analyses will be conducted as for the primary endpoint, including the primary analysis and sensitivity analysis.

11.7.2.5 Tender and Swollen Joint Count at Week 48

The change from baseline in counts for tender, swollen, and tender + swollen (active) joints will be calculated on all subjects in the randomized population.

- Tender 40-joint count at Week 48 – tender 40-joint count at baseline
- Swollen 40-joint count at Week 48 – swollen 40-joint count at baseline
- Tender + swollen (active) 40-joint count at Week 48 - tender + swollen 40-joint count at baseline

Analyses of each of these three endpoints will be conducted using a mixed model repeated measures (MMRM) model. For analysis visits where the NRI-imputation criteria are met (see section 11.7.1), no imputation is applied but the observed values are set missing. The MMRM model will include the fixed categorical effects of treatment group, analysis visit (Week 24, 32 and 48), treatment-by-analysis visit interaction, randomization stratum defined in section 6.5 and the continuous fixed covariate of baseline endpoint. An unstructured covariance matrix will be used to model the correlation of the repeated measures within each subject. The parameter estimations will be based on the assumption of data being missing at random (MAR) and using the method of restricted maximum likelihood (REML). The denominator degrees of freedom will be calculated using the Kenward-Rogers method. Adjusted mean changes (LSMEANS) with corresponding SE and 95% CI per treatment arm will be reported. The difference in adjusted means (LSMEANS) between BMS-986165 and placebo at the predefined time point, and corresponding 95% CI and p-value based on this longitudinal repeated measures model will be provided. In case of non-convergence of the preferred longitudinal model, memory space issues, or other computational issues the following covariance structures will be used in order of preference (decreasing number of parameters to be estimated): heterogeneous toeplitz, heterogeneous compound symmetry, and compound symmetry.

A longitudinal plot of mean change from baseline in tender joint count, swollen joint count, and tender +swollen joint count with standard errors by treatment group will be provided.

A sensitivity analysis will be conducted similar to the primary endpoint (treatment policy strategy).

[redacted]

[redacted]

11.7.4 Subgroup Analysis

Subgroup analyses will be performed for the primary endpoint, SRI(4) response at Week 32. Due to the reduced sample size within subgroup, no model-based adjusted odds ratios will be calculated. These analyses will include point estimates for response rate with asymptotic confidence intervals and the estimated unadjusted difference between study drug and placebo groups with Newcombe 95% confidence limits. P-values will not be reported for subgroup analyses. The following subgroups will be considered:

- Baseline CS dose (< 10 mg/day, ⩾ 10 mg/day)
- Baseline SLEDAI-2K Score (< 10, ⩾ 10)
- Geographic region (North America, South America, Japan, rest of world)
- Country
- Gender (male, female)
- Age group (18 - <45, 45 - <65, ⩾ 65 years)
- Weight (< median weight, ⩾ median weight)
- BMI (⩽ 30 kg/m^2, > 30 kg/m^2)
- Race
- Ethnicity (Hispanic or Latino, not Hispanic or Latino)
- CS Use at baseline (yes, no)
- Anti-malarial Use at baseline (yes, no)
- Immunosuppressant Use at baseline (yes, no)
- Interferon (IFN) status (high, low)
- Years since initial diagnosis (<3 yrs, 3-6 yrs, >6 years)

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

11.8 Safety Analyses

Safety summaries will be presented for the All-treated population, unless stated otherwise.

11.8.1 Adverse Events

The collection of nonserious AE information begins at initiation of study treatment and continues until the last follow-up visit. All SAEs that occur from the time the informed consent form (ICF) is signed through 30 days after the final dose of the study treatment must be reported. After that time, only SAEs deemed by the investigator to be related to the study treatment or a study procedure would be reported.

An adverse event (AE) including SAEs is defined as any new untoward medical occurrence or worsening of a preexisting medical condition in subjects that do not necessarily have causal relationship with treatment. Treatment-emergent adverse events (TEAEs) are those which first occur after the first dose of study drug or increase in severity after the first dose of study drug. The collected AEs (including SAEs) on CRF with an onset date between first dose date (including) and last dose day +30 days post last day or up to first dose day of study IM011-074 for roll over subjects will be summarized. All adverse events which change in severity or relationship to study drug are assigned a new start date and captured as a new record.

AEs will be coded according to MedDRA (version 21.0 or an updated version at the time of DBL). All adverse event tables will be summarized by treatment group and combined BMS-986165 groups.

A summary of the number and percentage of subjects reporting the following along with the number of events reported will be presented by treatment group and combined BMS-986165 groups:

- Subjects with at least one treatment-emergent adverse event
- Subjects with at least one treatment-related TEAE
- Subjects with at least one serious adverse event
- Subjects with at with one related serious adverse event
- Subjects discontinuing study drug due to an adverse event
- Subjects discontinuing study due to an adverse event
- Subjects who died

A breakdown of the number and percentage of subjects reporting each TEAE, categorized by coded body system and preferred term will be presented. Note that counting will be by subject not event and subjects are only counted once within each body system or preferred term; however, the number of total events including multiple events per subject will also be presented. A summary of treatment-related TEAEs will be provided. Furthermore, TEAEs and treatment-related TEAEs will be summarized for each preferred term in descending order.

A summary of TEAEs categorized by severity (mild, moderate, or severe) will also be provided. Events reported under multiple categories will be considered only at their maximum severity. Subjects with multiple events within a particular body system or preferred term will be counted under the category of their most severe event within that body system or preferred term.

A summary of TEAEs leading to discontinuation of study drug will be provided, grouped by body system and preferred term for all TEAEs and treatment-related TEAEs.

All adverse events (including non-treatment-emergent events) recorded on the CRF will be listed.

AE (including deaths) dates will be imputed according to algorithms detailed in Appendix 3. The imputed dates will be used to assess whether AEs should be considered as treatment-emergent and included in the safety summaries. The original, partial dates will be included in data listings.

No imputation will be performed on missing AE severity or relationship, they will be left as missing.

The exposure-adjusted incidence rate (EAIR) per 100 person-years of exposure (IR/100 P-Y) is calculated as follows:

$$100\ (person) * \frac{\text{number of subjects with the specific event}}{\text{total exposure-time (in years) among the subjects of that particular treatment}}$$

where total exposure time (years) among the subjects of that particular treatment

$$= \sum(exposure\ time\ (days)\ for\ all\ subjects\ of\ that\ particular\ treatment) * \frac{1\ year}{365.25\ days}$$

Subject Exposure Time (days) is calculated as follows:

- If a subject has at least 1 event while on that particular treatment, the exposure time for that subject is:

  First AE onset date – Treatment start date (of that particular treatment) + 1 day

- If a subject does not have an event, the exposure time for that subject is:

  Treatment stop date (of that particular treatment) – Treatment start date (of that particular treatment) + 1 day

11.8.2 Adverse Events of Interest

Adverse events of interest (AEI) include the following:

- Skin-related AEs
- Influenza
- Herpes Zoster
- Tuberculosis
- Opportunistic/Other Infections
- Malignancy events

These events will be summarized by body system and preferred term. cSMQ defined at program level will be applied to select the AEIs.

Creatine kinase (CK) elevation for CK elevation > 2.5 upper limit of normal will be summarized as CTCAE grade 2 or higher in the clinical laboratory summaries.

11.8.3 Deaths and Serious Adverse Events

The number and percentage of subjects reporting each SAE, categorized by body system and preferred term coded will be presented. Note that counting will be by subject not event and subjects are only counted once within each body system or preferred term; however, the number of total events including multiple events per subject will also be presented. A further tabulation of these events if related to the study drug, as reported on the CRF, will be provided.

All SAEs will be provided in a listing.

Adverse events with an outcome of death will be summarized.

All AEs with an outcome of death will be listed for the enrolled population.

11.8.4 Laboratory Data

Laboratory parameters will be summarized using both the International System (SI) of Units and conventional system of Units, unless otherwise specified. Summaries will be provided separately by different panels including hematology, general chemistry, chemistry - hepatobiliary, chemistry - renal function tests, chemistry - coagulation, and urinalysis. Only lab assessments with collection date up to last dose day +30 days post last dose day or up to first dose day of study IM011-074 for roll over subjects will be summarized. Data will be summarized by analysis visit, as applicable. The following summaries will be provided for each parameter:

- Absolute and change from baseline values
- Laboratory abnormalities (as determined by Controlled Terminology Criteria for Adverse Events [CTCAE v5.0] grading) presented as the worst postbaseline toxicity group

- Shifts from baseline to maximum postbaseline value
- Potential drug induced liver injury (DILI) is defined as a subject who meets the following criteria on the same assessment date:
  - 1) ALT or AST elevation >3 times ULN

  AND
  - 2) Total bilirubin >2 times ULN

  AND
  - Normal serum alkaline phosphatase (No initial findings of cholestasis)
- ALT or AST elevation >5 times ULN

- All abnormalities in laboratory data will be provided in listings.

11.8.5 Vital Signs

Vital signs including body temperature, respiratory rate, systolic blood pressure (SBP) and diastolic blood pressure (DBP), and heart rate will be listed. Absolute lab values, change from baseline values will be reported with summary statistics by treatment group and analysis visit.

The summaries of marked abnormalities at any timepoint postbaseline will also be tabulated for definitions provided as below.

- Absolute and change from baseline values
- Marked abnormalities defined by the below categories:
  - Heart rate:
    - Value > 100 and change from baseline > 30
    - Value < 55 and change from baseline < -15
  - Systolic blood pressure:
    - Value > 140 and change from baseline > 20
    - Value < 90 and change from baseline < -20
  - Diastolic blood pressure:
    - Value > 90 and change from baseline > 10
    - Value < 55 and change from baseline < -10

11.8.6 ECG

Summary statistics will be presented for ECG observed values, change from baseline values for all ECG parameters by treatment group and analysis visit.

The summaries of marked abnormalities at any timepoint post-baseline will also be tabulated for definitions provided as below.

- Absolute and change from baseline values
- Marked abnormalities defined by the below categories:
  - QT interval corrected using Fridericia's formula (QTcF):
    - 450 -< 480 msec
    - 480 -< 500 msec

- ⩾ 500 msec
- 30 < change from baseline ⩽ 60 msec
- Change from baseline > 60 msec
- Males: < 450 msec, ⩾ 450 msec
- Females: < 470 msec, ⩾ 470 msec

▪ PR interval ⩾ 200 msec

▪ QRS interval ⩾ 200 msec

In cases where QTcF is missing, then QTcF will be derived using the following definition: QTcF = QT/ (RR^(1/3)), where RR (seconds) = 60 / HR. In case where both QTcF and QT are missing, then QTcF will be derived using the following definition: QTcF = QTcB/ (RR^(1/6)), where RR (seconds) = 60 / HR.

Additionally, all ECG data will be provided in a listing.

11.8.7 Physical Examination

Complete physical examinations results will be summarized using frequency distributions for each body system by treatment group and analysis visit.

All physical examination results will be provided in a listing.

11.9 Statistical impacts due to Covid-19

11.9.1 Impact on efficacy endpoints

Some subjects in this study did not complete all planned visits due to the impact of COVID-19, which can potentially impact the primary and secondary efficacy endpoints. It is plausible to assume that the underlying missingness mechanism due to the COVID-19 is 'missing at random (MAR)'. Based on that assumption, the following approaches will be applied:

- For statistical analyses that are based on methods that are in agreement with the MAR assumption (e.g. MMRM model for continuous endpoints), the point estimates of treatment-specific means and differences can be assumed not be affected. Therefore, no modifications will be done to these analyses and no additional analyses will be conducted to address it.
- For statistical analyses that implement a missing-not-at-random imputation strategy (e.g. NRI for binary endpoints), estimated treatment-specific means and differences could be affected if the frequency of COVID-19 related missing observations differs between treatment groups. Therefore, for binary endpoints, a sensitivity analysis will be conducted where observations that were missed due to COVID-19 will be imputed with the last value observed before the visit that was missed due to COVID-19 restrictions

[redacted]

11.9.2 Impact on safety endpoints

No modifications will be made for the analysis of safety data due to COVID-19-related issues. There will not be any additional safety data handling considerations for missing data or treatment interruptions due to COVID-19.

12.0 **Pharmacokinetic Analysis**

The Pharmacokinetic (PK) population will be used for all listings. The Evaluable PK population will be used for summaries and statistical analyses.

12.1 Plasma Concentrations

PK concentrations of BMS-986165 and BMT-153261 will be listed by subject and treatment at each scheduled time point.

The following groups will be summarized:

- Subjects who received 12 mg BMS-986165 QD
- Subjects who received 6 mg BMS-986165 BID
- Subjects who received 3 mg BMS-986165 BID

For the summaries of concentration-time data, concentrations that are less than the lower limit of quantification (LLOQ) should be displayed as "< LLOQ" in the listings and be treated as missing.

Summaries of plasma concentrations of BMS-986165 and BMT-153261 will be provided for the PK population by day and scheduled time point as outlined above using summary statistics (n, mean, SD, geometric mean, %CV, median, min, max).

12.2 Pharmacokinetic Parameters

The following PK parameters may be estimated for BMS-986165 and the active metabolite BMT-153261 using noncompartmental methods by a validated PK analysis program, if data permits.

Parameter	Description
Plasma PK Parameters	
Cmax	Maximum observed concentration, obtained directly from the experimental data without interpolation, expressed in concentration units determined on Day 85.
Tmax	Time of maximum observed plasma concentration associated with Cmax on Day 85
Ctrough	Observed predose (trough) concentration determined on Days 15, 29, 57, 85, 169, 225, and 337.

All individual PK parameters will be listed for each analyte. Summaries of PK parameters will be tabulated for each analyte by treatment using descriptive statistics (n, mean, SD, median, min, max, geometric mean, and %CV).

The plasma PK parameters will be estimated from the concentration-time profiles. For the purpose of calculating PK parameters except Ctrough, predose concentrations below the lower limit of quantitation (<LLOQ) and concentrations prior to the first quantifiable concentration that were <LLOQ were set to "zero," all other <LLOQ concentrations were set to "missing." Actual sampling times, rather than scheduled sampling times, will be used in all computations involving sampling times.

Only true trough concentrations will be summarized. True trough concentrations are defined as PK concentrations measured before the next dose within the window of 24 hrs +/- 4 hrs for QD dosing and 12 hrs +/- 2 hrs for BID dosing. Samples taken accidentally after the dose occurred as well as samples taken earlier than specified windows will not be included in summaries.

Summary statistics for Ctrough concentrations will be calculated by imputing values less than LLOQ as ½ × LLOQ. This imputation is performed for Ctrough concentrations and not conducted for Day 1 pre-dose concentrations. Individual Ctrough listings will display these concentrations as "< LLOQ."

The comparison of trough concentrations of parent drug and metabolite(s) for all 3 groups will be presented graphically in the box/scatter plots.

13.0 **Biomarker Analysis**

Summaries will be provided for the biomarker analysis population. Summaries include mean, SD, median, IQR, range, change from baseline and percent change from baseline. Summaries for the following biomarkers will be provided for treatment arms as well as for baseline IFN status (IFN-high and IFN-low) under each arm at each analysis visit. IFN-high vs. IFN-low will be designated based on IFN evaluation of samples collected during the screening period. Demographic information of subjects for each IFN status will be provided.

[redacted]

- aggregated genes (IRGs [redacted])

[redacted]

- clinical markers: C3, C4, Coomb's test, anti-dsDN [redacted]

[redacted]

% Change of individual IRGs will be derived as follows:

- The log2 Fold Change of each gene of each subject = [normalized log2 value at time point]-[normalized log2 value at Day 1 predose]
- % Change from Baseline of each gene at a time point = [2^(log2 Fold Change)-1] x 100

14.0 **Interim Analysis**

No interim analysis will be performed during the course of the study.

15.0 **References**

Livak KJ, Schmittgen TD (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ Method. Methods. 25(4):402-8

Greenland and Mansournia (2015). Penalization, Bias reduction, and default priors in logistic and related categorical and survival regressions. Statistics in Medicine 34, 3133-3143

Furie R, Petri MA, Strand V, et al. Clinical, laboratory and health-related quality of life correlates of Systemic Lupus Erythematosus Responder Index response: a post hoc analysis of the phase 3 belimumab trials. Lupus Sci Med 2014;1(1):e000031.

Appendix 1 Glossary of Abbreviations

Glossary of Abbreviations:	
ACR	American College of Rheumatology
AE	Adverse Event
AEI	Adverse Events of Interest
ATC	Anatomic Therapeutic Classification
BICLA	British Isles Lupus Assessment Group-based Composite Lupus Assessment
BILAG	British Isles Lupus Assessment Group
CI	Confidence Interval
CLASI	Cutaneous Lupus Erythematosus Disease Area and Severity Index
CRF	Case Report Form
CRP	C-reactive Protein
[redacted]	
CS	Corticosteroid
CSR	Clinical Study Report
DBL	Database Lock
DBP	Diastolic Blood Pressure
DMC	Data Monitoring Committee
DSMB	Data Safety Monitoring Board
[redacted]	
ECG	Electrocardiogram
EOS	End-Of-Study
ESR	Erythrocyte Sedimentation Rate
FP	Follow up Period
[redacted]	
IFN	Interferon
IRG	Interferon-Regulated Gene
IRT	Interactive Response Technology
IVRS	Interactive Voice Response System
LLDAS	Lupus Low Disease Activity State
LS	Least Square
NSAID	Nonsteroidal Anti-Inflammatory Drug
MedDRA	Medical Dictionary for Regulatory Activities

[redacted]	
PCI	Potentially Clinically Important
PGA	Physician's Global Assessment
PK	Pharmacokinetic
[redacted]	
PPS	Per Protocol Set
[redacted]	
[redacted]	
SAP	Statistical Analysis Plan
SAE	Serious Adverse Event
SBP	Systolic Blood Pressure
[redacted]	
[redacted]	
SLE	Systemic Lupus Erythematosus
SLEDAI-2K	Systemic Lupus Erythematosus Disease Activity Index 2000
[redacted]	
SP	Study Period
SRI(X)	Systemic Lupus Erythematosus Responder Index of ≥ X
TB	Tuberculosis
Tyk2	Tyrosine Kinase 2
VAS	Visual Analog Scale
WHO	World Health Organization

Appendix 2 Schedule of On-Treatment Activities

Treatment Period Procedural Outline Up to Week 24

Procedure	Week 0 D1	Week 2 D15 (±3 d)	Week 4 D29 (±3 d)	Week 8 D57 (±3 d)	Week 12 D85 (±3 d)	Week 16 D113 (±3 d)	Week 20 D141 (±3 d)	Week 24 D169 (±3 d)
Eligibility/randomization criteria	X							
[redacted]								
Safety Assessments								
Complete PE	X							
Targeted PE		X	X	X	X	X	X	X
Body weight	X		X	X	X	X	X	X
Vital signs	X	X	X	X	X	X	X	X
Electrocardiogram	X		X	X	X			
Concomitant medication use	X	X	X	X	X	X	X	X
Monitor for AEs and SAEs	X	→	→	→	→	→	→	X
Laboratory Tests								
Hematology, chemistry, urinalysis, coagulation	X	X	X	X	X	X	X	X
Fasting lipid panel	X				X			
Fasting plasma glucose	X				X			
Pregnancy test (WOCBP only)	X	X	X	X	X	X	X	X

Procedure	Week 0 D1	Week 2 D15 (±3 d)	Week 4 D29 (±3 d)	Week 8 D57 (±3 d)	Week 12 D85 (±3 d)	Week 16 D113 (±3 d)	Week 20 D141 (±3 d)	Week 24 D169 (±3 d)
Spot urine for protein creatinine ratio	X		X	X	X	X	X	X
Coomb's Test (direct)	X		X	X	X	X	X	X
hsCRP	X				X			X
TBNK	X				X			X
Serum Ig	X		X		X			X
PK Assessments	X	X	X	X	X			X
Biomarker Assessments								
Serum Complement	X	X	X	X	X	X	X	X
Anti-dsDNA autoantibodies	X		X	X	X	X	X	X
[redacted]								
Blood RNA for IRG and steroid signature, and genome-wide expression	X	X	X	X	X	X	X	X
[redacted]								
Efficacy Assessments								

Procedure	Week 0 D1	Week 2 D15 (±3 d)	Week 4 D29 (±3 d)	Week 8 D57 (±3 d)	Week 12 D85 (±3 d)	Week 16 D113 (±3 d)	Week 20 D141 (±3 d)	Week 24 D169 (±3 d)
BILAG-2004 Index	X		X	X	X	X	X	X
SLEDAI-2K	X		X	X	X	X	X	X
[redacted]								
Physician's Global Assessment	X		X	X	X	X	X	X
40-joint count	X		X	X	X	X	X	X
CLASI	X		X	X	X	X	X	X
Photography	X		X	X	X	X	X	X
Assess suitability for CS taper	(X)	(X)	(X)	X	X	X	X	
[redacted]								
Patient's Global Assessment	X		X	X	X	X	X	X
Pain assessment	X		X	X	X	X	X	X
Clinical Drug Supplies								
Randomization	X							
Dispense Study Treatment	X	→	→	→	→	→	→	X
Study Treatment Administration	X	→	→	→	→	→	→	X
Study Treatment Compliance		X	X	X	X	X	X	X

Treatment Period Procedural Outline After Week 24

Procedure	Week 28 D197 (±3 d)	Week 32 D225 (±3 d)	Week 36 D253 (±3 d)	Week 40 D281 (±3 d)	Week 44 D309 (±3 d)	Week 48 D337/EOT (±3 d)	FP (28 days after EOT)[a]
Safety Assessments							
Complete PE		X				X	
Targeted PE	X		X	X	X		X
Body weight	X	X	X	X	X	X	X
Vital signs	X	X	X	X	X	X	X
Electrocardiogram		X				X	
Concomitant medication use	X	X	X	X	X	X	X
Monitor for AEs and SAEs	X	→	→	→	→	→	X
Laboratory Tests							
Hematology, chemistry, urinalysis, coagulation	X	X	X	X	X	X	X
Fasting lipid panel		X				X	
Fasting plasma glucose		X				X	
Pregnancy test (WOCBP only)	X	X	X	X	X	X	X
Spot urine for protein creatinine ratio	X	X	X	X	X	X	X
Coomb's Test (direct)	X	X	X	X	X	X	X
hsCRP		X		X		X	
TBNK		X				X	

Procedure	Week 28 D197 (±3 d)	Week 32 D225 (±3 d)	Week 36 D253 (±3 d)	Week 40 D281 (±3 d)	Week 44 D309 (±3 d)	Week 48 D337/EOT (±3 d)	FP (28 days after EOT)[a]
Serum Ig		X				X	X
PK Assessments		X				X	
Biomarker Assessments							
Serum Complement	X	X	X	X	X	X	X
Anti-dsDNA autoantibodies	X	X	X	X	X	X	X
[redacted]							
Blood RNA for IRG, steroid signature, and genome-wide expression	X	X	X	X	X	X	X
[redacted]							
Efficacy Assessments							
BILAG-2004 Index	X	X	X	X	X	X	X
SLEDAI-2K	X	X	X	X	X	X	X
[redacted]							
Physician's Global Assessment	X	X	X	X	X	X	X

Procedure	Week 28 D197 (±3 d)	Week 32 D225 (±3 d)	Week 36 D253 (±3 d)	Week 40 D281 (±3 d)	Week 44 D309 (±3 d)	Week 48 D337/EOT (±3 d)	FP (28 days after EOT)[a]
40-joint count	X	X	X	X	X	X	X
CLASI	X	X	X	X	X	X	X
Photography	X	X	X	X	X	X	X
Assess suitability for CS taper		X	X	X		X	
[redacted]							
Patient's Global Assessment	X	X	X	X	X	X	X
Pain assessment	X	X	X	X	X	X	X
Clinical Drug Supplies							
Dispense Study Treatment	X	X	X	X	X		
Study Treatment Administration	X	X	X	X	X	X	
Study Treatment Compliance	X	X	X	X	X	X	

Appendix 3 Imputation Rules for Partial of Missing Dates

Start Date		Stop Date						
		Complete: *yyyymmdd*		Partial: *yyyymm*		Partial: *yyyy*		Missing/ Ongoing
		<1st dose	≥1st dose	<1st dose *yyyymm*	≥1st dose *yyyymm*	<1st dose *yyyy*	≥1st dose *yyyy*	
Partial: *yyyymm*	= 1st dose *yyyymm*	2	1	n/a	1	n/a	1	1
	≠ 1st dose *yyyymm*		2	2	2	2	2	2
Partial: *yyyy*	= 1st dose *yyyy*	3	1	3	1	n/a	1	1
	≠ 1st dose *yyyy*		3		3	3	3	3
Missing		4	1	4	1	4	1	1

1 = Impute as the date of first dose

2 = Impute as the first of the month

3 = Impute as January 1 of the year

4 = Impute as January 1 of the stop year

Note: If the start date imputation leads to a start date that is after the stop date, then there is a data error and do not impute the start date.

Imputation rules for partial or missing stop dates:

1. Initial imputation
    a. For partial stop date “mmyyyy”, impute the last of the month.
    b. For partial stop date “yyyy”, impute December 31 of the year.
    c. For completely missing stop date, do not impute.
2. If the stop date imputation leads to a stop date that is after the death date, then impute the stop date as the death date.
3. If the stop date imputation leads to a stop date that is before the start date, then there is a data error and do not impute the stop date.

Imputation rules for partial or missing death dates:

1. If death year and month are available but day is missing:

    a. If “mmyyyy” for last contact date = “mmyyyy” for death date, set death date to the day after the last contact date.

    b. If “mmyyyy” for last contact date < “mmyyyy” for death date, set death date to the first day of the death month.

c. If “mmyyyy” for last contact date > “mmyyyy” for death date, data error and do not impute.

2. If both month and day are missing for death date or a death date is totally missing, set death date to the day after the last contact date.

[redacted]

[redacted]

[redacted]

[redacted]

Appendix 5 Relevant Programmable PDs

- Eligibility Deviations:
  - Subjects were not diagnosed ≥ 24 weeks before the screening visit
  - Subjects did not meet SLICC classification
  - Subjects did not meet at least one of the following serology criteria:

    elevated antinuclear antibodies (ANA) ≥ 1:80, or positive anti-dsDNA (positive includes indeterminate results), or positive anti-Smith (anti-Sm) as determined by the central laboratory
  - Subjects did not meet BILAG grading criteria
  - Subjects did not meet SLEDAI-2K criteria
  - Subjects did not meet background meds requirements for immunosuppressants, antimalarials (must be on one of these med type, no dose requirements) based on a list of WHO Drug coding terms
  - For subjects receiving prednisone or equivalent, the dose of CS was not stable for 2 weeks before the screening visit (signing of informed consent) or throughout the screening period until randomization
  - A more than 30 mg/day prednisone or equivalent dose was used during the screening period (or in the 2 weeks preceeding the screening visit)
  - Subjects have drug induced lupus, other autoimmune diseases, SLE overlap syndromes

  Note: All eligibility related relevant PDs will be determined based on data collected from Eligibility – Screening eCRF page and assessments on day 1
- Incorrect dosing:
  - Subjects took medication other than randomized treatment for the entire duration of the study.
  - Overall compliance <75% for the treatment period to Week 32.
- Restricted and Prohibited medications:
  - Subjects should not take any medications included in protocol Appendix 7 between first dose and efficacy assessment. A list of WHO Drug coding terms will be defined at protocol level.
  - Subjects who take any prohibited medications listed in section 11.7.1

Appendix 6 List of Biomarkers

Interferon regulated genes (IRGs)				
HERC5*	IRF7	MX1*	CXCL10	STAT2
DDX58	DHX58	EIF2AK2	IFI44*	IFI44L*
IFIT1*	ISG15*	LGALS3BP	PARP9	RSAD2*
TNFSF13B				

[redacted]

[redacted]