BOS161721-02: A Randomized Double-Blind Phase 1b/2 Combined Staggered Multiple Dose Escalation Study of BOS161721 In Systemic Lupus Erythematosus (SLE) Patients on a Background of Limited Standard of Care

ClinicalTrials.gov Identifier: NCT03371251

Statistical Analysis Plan, Version 4.0, dated 23-April-2021
Statistical Analysis Plan, Version 3.0, dated 13-November-2020
Statistical Analysis Plan, Version 2.0, dated 29-November-2018
Statistical Analysis Plan, Version 1.0, dated 05-September-2018

BOS161721-02: A RANDOMIZED DOUBLE-BLIND PHASE 1b/2 COMBINED STAGGERED MULTIPLE DOSE ESCALATION STUDY OF BOS161721 IN SYSTEMIC LUPUS ERYTHEMATOSUS (SLE) PATIENTS ON A BACKGROUND OF LIMITED STANDARD OF CARE

Statistical Analysis Plan

VERSION 4.0
DATE OF PLAN:

23Apr2021

STUDY DRUG:
BOS161721

PREPARED FOR:

Boston Pharmaceuticals, Inc.

Approval Signature Page: CATO SMS

PPD

4/23/2021

Document Author:
PPD, MS
Director of Biostatistics

Date

PPD

4/23/2021

Document Reviewer:
PPD, MS
Vice President, Biostatistics &

Date

Note: Reviewed changes from V3.0

Approval Signature Page: Boston Pharmaceuticals, Inc.

PPD

PPD, MD
Vice President, Clinical Dev

4/23/2021
Date

PPD

PPD
Vice President, Biometrics

4/23/2021
Date

Contents

ABBREVIATIONS

ACR	American College of Rheumatology
ADA	Anti-drug antibody
AE	Adverse event
AESI	Adverse event of special interest
ANA	Antinuclear antibody
ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
APL	Antiphospholipid
ATC	Anatomic therapeutic class
AUC	Area under the curve
BICLA	BILAG-based Composite Lupus Assessment
BILAG	British Isles Lupus Assessment Group
BP	Blood pressure
C3	Complement C3
C4	Complement C4
C_L	Systematic clearance
CLASI	Cutaneous Lupus Area and Severity Index
C_{max}	Maximum plasma concentration
CS	Corticosteroid
CSR	Clinical Study Report
C-SSRS	Columbia Suicide Severity Rating Scale
CTCAE	Common Terminology Criteria for Adverse Events
CV	Coefficient of variation
DLT	Dose-limiting toxicity
DMC	Data Monitoring Committee
dsDNA	Double-stranded deoxyribonucleic acid
ECG	Electrocardiogram
eCRF	Electronic case report form
EWB	Emotional well being
CCI	[redacted]
FAS	Full analysis set
FS	Fatigue subscale
FWB	Functional well being
HR	Hazard ratio
IA	Interim Analysis
IEC	Independent Ethics Committee
IL-21	interleukin 21
IM	Intramuscular
IRB	Institutional Review Board
ISR	Injection site reaction
LLN	Lower limit of normal
LLOQ	Lower limit of quantification
LOCF	Last observation carried forward

LS	Least squares
MAD	Multiple Ascending Dose
MCS	Mental component summary
MedDRA	Medical Dictionary for Medical Affairs
NAb	Neutralizing antibody
NCI	National Cancer Institute
PCS	Physical component summary
PD	Pharmacodynamics
PGA	Physician's Global Assessment
PK	Pharmacokinetics
POC	Proof of concept
PP	Per protocol set
pSTAT3	Phosphorylated signal transducer and activator of transcription 3
PT	Preferred Term
PWB	Physical well being
Q1	25th percentile/first quartile
Q3	75th percentile/third quartile
SAD	Single ascending dose
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SAS	Statistical analysis system
SC	Subcutaneously
SD	Standard deviation
CCI	[redacted]
SI	International System of Units
SLE	Systemic Lupus Erythematosus
SLEDAI-2K	SLE Disease Activity Index 2000
SLICC	Systemic Lupus International Collaborating Clinics
Sm	Smith
SOC	System Organ Class
SRI-4	SLE Responder Index 4
SRI-5	SLE Responder Index 5
SRI-6	SLE Responder Index 6
SS	Safety analysis set
SSA	Sjögren syndrome-A (Ro)
SSB	Sjögren syndrome-B (La)
SWB	Social/Family well being
$t_{1/2}$	Terminal elimination half-life
TEAE	Treatment emergent adverse event
T_{max}	First time to maximum concentration
TTMF	Time to medication failure
ULN	Upper limit of normal
V_d	Volume of distribution
WHO	World Health Organization

1. INTRODUCTION

The Statistical Analysis Plan (SAP) details the planned analyses that will be included in the Clinical Study Report (CSR) of study number BOS161721-02: A Randomized Double-Blind Phase 1b/2 Combined Staggered Multiple Dose Escalation Study of BOS161721 in Systemic Lupus Erythematosus (SLE) Patients on a Background of Limited Standard of Care.

SAPs describing analyses (pharmacokinetics [PK], PK/pharmacodynamics[PD], etc.) not covered within this SAP will be included within Appendix 13.13 of this document. The author, finalization date, signatories, and page numbering of this SAP are independent of those SAPs and vice versa. Only the last final versions will be included in the appendix just prior to study conclusion.

The content of this SAP is based on the protocol Version 7.0 (Amendment 6) dated 30April2020.

SAP Revision Chronology:	**Date:**	**Comments:**
V1.0	05SEP2018	Original
V2.0	29NOV2018	Modifications due to Protocol Version 4 (Amendment 3) • Modified evaluable window for efficacy assessments to be inclusive of 35 days since prior assessment. • Modified structure of table shells to combine analyses (i.e., CLASI total activity and total damage; ACR-28 sum of swelling and sum of tenderness; CCI [redacted]; biomarker and pharmacodynamic parameters) • Modified structure of listing shells CCI [redacted] • Clarified Last Observation Carried Forward (LOCF) versus observed analyses in the output titles within the shells
V3.0	13NOV2020	Modifications due to Protocol Version 7 (Amendment 6) including: • Sample size changed based on recent published literature • LOCF algorithm updated • Modified planned sensitivity analyses based on missing data • Added, amended, and removed secondary and exploratory endpoints per protocol amdement 6 • Defined efficacy subgroups • Added COVID-19 related data handling details • Modified evaluable window for efficacy assessments • Modified table, listing, and figure shells in accordance with text changes
V4.0	23APR2021	Modifications due to pre-lock comments from blinded review of draft tables, listing, and figures include:

		• Clarifications surrounding duration of longest SRI-4 response (LOCF methods and addition of 30 days in the case of missing dates) • Clarifications provided around LOCF methods for baseline disease characteristics used in both the disease characteristics summaries and subgroup analyses • Handling of biomarker/PD results <LLOQ modified to utilize 0.5*LLOQ instead of 0 in the analysis • BICLA definition clarifications provided • Medication failure identification and date determination explained as being provided by the sponsor based on the blinded review of the prior and concomitant medication eCRF form as opposed to utilizing the protocol deviation form. • Refinement of the per protocol set • Additional corrections of minor typos and clarifications resulting from pre-lock review

2. STUDY OBJECTIVES AND ENDPOINTS

2.1. Multiple Ascending Dose (MAD) Phase 1b Study

Objectives	Endpoints
Primary	
• To assess safety, tolerability, and immunogenicity of repeat doses of BOS161721 (20, 60, and 120 mg) administered subcutaneously (SC) in adult patients with moderately to severely active SLE on limited background standard of care treatment, in order to estimate the optimal dose.	Safety Endpoints • Incidence and severity of adverse events (AEs) and serious adverse events (SAEs), related AEs, AEs leading to study drug discontinuation, AEs by severity and relatedness • Injection site reactions • Columbia Suicide Severity Rating Scale (C-SSRS) • 12-lead electrocardiograms (ECGs) parameter results at each visit and change from baseline • Vital signs (blood pressure [BP], heart rate, and temperature) parameter results at each visit and change from baseline • Clinical laboratory results and change from baseline • Physical examinations changes from baseline • Anti-drug antibodies (ADAs) • Study drug exposure/compliance

Objectives	Endpoints
Secondary	
• To characterize the PK of BOS161721 and select the optimal dose of BOS161721 based on safety, PK, and PD effects in patients with moderate to severely active SLE.	Pharmacokinetic (PK) Endpoints • BOS161721 concentration by visit and time point • Maximum observed concentration (C_{max}), first time to maximum concentration (T_{max}), area under the concentration-time curve (AUC), terminal half-life ($t_{½}$), systematic clearance (C_L), volume of distribution (V_d) Pharmacodynamic (PD) Endpoints • Results and changes (or shifts) from baseline to each visit in phosphorylated signal transducer and activator of transcription 3 (pSTAT3), C3 and C4 levels, and leukocyte immunophenotype • Results and changes (or shifts) from baseline in anti-double-stranded DNA (dsDNA), antinuclear antibodies (ANA), anti-Sjögren syndrome A and B (SSA, SSB), Smith (Sm), and antiphospholipid (APL) autoantibodies at each visit • Results and changes (or shifts) from baseline in abrogation of IL-21 gene signature at each indicated visit
Exploratory	
• CCI [redacted]	CCI [redacted]

Objectives	Endpoints
	- CCI • CCI • CCI
• CCI	• CCI
• CCI	• CCI

[1] See Section 11

2.2. Proof of Concept (POC) Phase 2 Study

Objectives	Endpoints
Primary	
• To demonstrate a superior effect of BOS161721 at the chosen dose compared with placebo for response on the SLE SRI-4	Primary Efficacy Endpoint • The proportion of patients with a SRI-4 response at Day 210
Secondary	
• To demonstrate a superior effect of BOS161721 at the chosen dose compared with placebo for response on clinical indicators of SLE activity, in adult patients with moderately to severely active SLE on limited background standard of care treatment	Secondary Efficacy Endpoints • The proportion of patients with: - SRI-4 response at each visit - SRI-5 and SRI-6 response at each visit - a sustained reduction of oral corticosteroid (CS) (≤ 7.5 mg/day and < Day 0 dose) between Day 150 and Day 210 - new or recurrent BILAG flares (≥ 1 qualifying BILAG A or > 1 qualifying BILAG B) through Day 210 - PGA worsening - a BICLA response - a CLASI response - medication failures • Change and percentage change from baseline at each visit in: - CLASI - PGA - Total number of swollen joints, tender joints, and active joints (swelling and tenderness in the same joint) in the ACR-28 joint count - SLEDAI-2K - SLICC/ACR damage index • TTMF • Group mean percent reduction in corticosteroid administration from baseline Day 0 dose through Day 210 in patients receiving ≥ 7.5 mg/day prednisone equivalent at Day 0

Objectives	Endpoints
	• Duration of longest SRI-4 response • Time to first BILAG flare (≥ 1 new or recurrent BILAG A or > 1 new or recurrent BILAG B) relative to baseline through Day 210
Safety	
• To assess safety and tolerability of repeat doses of BOS161721 at the chosen dose administered SC in adult patients with moderately to severely active SLE on limited background standard of care treatment	Safety Endpoints • Incidence and severity of AEs and SAEs, related AEs, AEs leading to study drug discontinuation, AEs by severity and relatedness • Injection site reactions • C-SSRS • 12-lead ECGs parameter results at each visit and change from baseline • Vital signs (BP, heart rate, and temperature) parameter results at each visit and change from baseline • Clinical laboratory results and change from baseline • Physical examinations changes from baseline • ADAs • Study drug exposure/compliance

Objectives	Endpoints
Exploratory	
• CCI	• CCI
• CCI	CCI
• CCI	• CCI
• CCI	• CCI

Endpoints will be discussed in the SAP based on their order in the phase 2 POC portion of the study. The phase 1b MAD endpoints overlap with and will be discussed within the context of the phase 2 POC endpoints.

3. STUDY DESIGN

3.1. Study Design and Population

This is a Phase 1b/2 combined, randomized, multicenter, double-blind, placebo-controlled trial to study the clinical efficacy, safety, and PK of multiple subcutaneous doses of BOS161721 in adult

patients with moderately to severely active SLE. After successfully completing a screening phase, eligible patients will be randomized to a specified dose of BOS161721 or placebo. The dosing schedule will be monthly.

The trial will consist of 2 double-blinded portions: MAD Phase 1b and POC Phase 2. Patients may receive a total of 7 SC monthly doses of study drug on Days 0, 30, 60, 90, 120, 150, and 180, followed by safety follow-up visit at Days 210, 240, and 270.

Investigators will assess SLE activity in accordance with accepted evaluation tools (further details provided in the SLE Assessment Reference Guide).

SLE disease activity assessment data will be centrally reviewed by the CRO medical monitor(s), independent subject matter expert reviewers/adjudicators and the Sponsor to ensure the scores are clinically meaningful, compliant with specific definitions and compatible across various disease scoring tools. The scope of responsibility includes, but is not limited to, review and confirmation of BILAG organ system disease grades, BILAG A and B flares, and clinical components of SLEDAI-2Kscore at screening and during the study as well as cross-validation of the instruments used in this study to assess disease activity. Further details on the content, processes, procedures and methods of disease activity data reports will be outlined in the Medical Data Review Plan and the BILAG-2004 Review Guidelines.

3.1.1. Multiple Ascending Dose Phase 1b

The MAD portion will consist of 3 cohorts:

- Cohort 1 (20 mg SC) will include 6 patients
    - 5 patients will receive BOS161721 (active group) and 1 patient will receive placebo (placebo group)
- Cohorts 2 (60 mg SC) and 3 (120 mg SC) will include 12 patients each
    - 9 patients in the active group and 3 in the placebo group

Doses selected for each of the 3 cohorts is based on a 90-day safety, tolerability, PK and PD data review from the Phase 1 single ascending dose (SAD) study (BOS161721-01) in healthy patients. All doses selected for the MAD part of the study are projected not to exceed the mean exposure of that achieved at the highest dose in the SAD study.

The MAD portion of the study design is staggered, where after the 6 patients in Cohort 1 have received 2 doses and have completed 2 weeks of follow up- after the second dose, Cohort 2 begins dosing after the Data Monitoring Committee (DMC) evaluation of the safety and tolerability data from Cohort 1. Similarly, after 8 of the 12 patients in Cohort 2 have received 2 doses and have completed 2 weeks of follow up- after the second dose, Cohort 3 begins dosing after the DMC evaluation of the safety and tolerability data from Cohorts 1 and 2. Each cohort will continue at their assigned dose level through their respective 6-month treatment periods (See

Study Schematic in the study protocol). If patients discontinue the study in a cohort prior to adequate safety follow-up, he/she may be replaced.

Criteria for dose escalation are further described in the DMC Charter.

3.1.2. Proof of Concept Phase 2

The dose for the POC portion of the study is 120 mg administered SC monthly (a total of 7 doses). The rationale for the BOS161721-02 phase 2 POC dose selection was based on cumulative safety, tolerability, immunogenicity, PK, and PD data available from an interim analysis (IA) performed during the MAD phase 1b portion of the trial.

The data cut-off for this IA occurred on CCI and included all 6 patients and 7 doses from Cohort 1 (20 mg), 12 patients and 6 doses from Cohort 2 (60 mg), and 12 patients and 4 doses from Cohort 3 (120 mg).

The safety analysis focused on incidence and severity of all AEs, SAEs, and pre-determined adverse events of special interest (AESI). The DMC and designated unblinded Boston Pharmaceuticals team met on CCI and did not identify any untoward safety signals at any BOS161721 dose levels.

Because there were no safety, tolerability, or immunogenicity trends observed at the time of the IA, the phase 2 POC dose selection was made based on available PK and PD data. pSTAT3 levels were assessed as the primary PD biomarker of IL-21R signaling levels. This is because IL-21R signaling, upon IL-21 binding, initially involves phosphorylation of JAK1/JAK3 which dissociate from the receptor complex, and phosphorylate STAT3 which translocates to the nucleus and drives IL-21-regulated gene expression. CCI

The 120 mg dose was communicated to site Investigators participating in the POC phase 2 portion, and to Institutional Review Boards (IRBs) and the Independent Ethics Committee (IEC).

For the POC part of the study, approximately 110 additional patients will be randomized to active or placebo groups in a 2:1 ratio.

As in the MAD part of the study, each patient in the POC portion may receive a total of 7 SC monthly doses of study drug on Days 0, 30, 60, 90, 120, 150, and 180. Assessments will be completed according to the Schedule of Assessments detailed in the study protocol.

DMC safety reviews will be conducted periodically throughout the study as described in the DMC charter.

3.2. Randomization and Blinding

This is a randomized, double-blind study.

Patients meeting all inclusion and exclusion criteria will be centrally randomized to either placebo or BOS161721 using an interactive web response system according to a randomization list generated by an independent, non-study statistician. Randomization will be performed separately for each study portion and separately for each cohort in the Phase 1b and 2 portions as follows:

Phase/Cohort	Number of Patients	Randomization Ratio BOS161721:Placebo
Phase 1b/Cohort 1	6	5:1
Phase 1b/Cohort 2	12	3:1
Phase 1b/Cohort 3	12	3:1
Phase 2	*Approximately 110	2:1

*Additional patients may be enrolled to ensure sufficient numbers of patients are in the full analysis set (FAS).

Eligible patients will be assigned to the study portion which is active at time of enrollment. Similarly, patients in the Phase 1b MAD will be assigned to the cohort which is active. Each patient will be assigned a unique randomization number which will not be reused.

All patients, investigators, and study team participants will be blinded to treatment assignment. An independent biostatistician not otherwise involved on the study will be unblinded and prepare materials for the IA, ad hoc analyses as needed, and the DMC safety reviews. The DMC will review unblinded data during safety reviews and the IA. A limited team at Boston Pharmaceuticals will review unblinded results from the Phase 1b MAD portion during the IA to determine the dose that will be used for the Phase 2 POC portion of the study. Details regarding maintenance of the blinding and content of data reviews will be described in the DMC charter or related study documentation.

3.3. Sample Size Considerations

Sample size in the Phase 1b part of the study is based on operational consideration.

The sample size in the Phase 2 POC study is based on the primary endpoint, SRI-4 Response at Day 210. Based on efficacy data from a recently completed study with Ustekinumab[iii], where the primary endpoint was also SRI-4, the sample size in the phase 2 part of the BOS161721-02 study is set at 96 evaluable patients. In the Ustekinumab study of 102 patients, a statistically significant difference in SRI-4 response was shown in favor of Ustekinumab (62% of Ustekinumab treated patients achieved SRI-4 response vs. 33% in patients receiving placebo). The Ustekinumab phase 2 study is an appropriate comparator to BOS161721-02 because both studies share the same primary endpoint of SRI-4 response and BOS161721 and Ustekinumab have overlapping mechanisms of action (T follicular helper and Th17 cell biology).

Approximately 110 additional patients will be randomized in a 2:1 ratio to BOS161721 versus placebo to achieve approximately 96 evaluable patients in the FAS of the phase 2 portion. This assumes 14 patients will discontinue prior to having received treatment or having completed a post-baseline efficacy measure. Enrollment and randomization will be monitored in a blinded fashion and may be increased if needed to ensure at least 96 evaluable patients are randomized into the FAS (see Section 6.1).

A total of 96 evaluable patients randomized provides 80% power to detect a treatment difference of 29% in SRI-4 response rates at Day 210, based on a targeted 2-sided significance level of 10% and using a 2-sided Pearson's chi-squared test. This assumes a response rate of 62% for BOS161721 and 33% for placebo.

3.4. Data Monitoring Committee (DMC)

This study will use an external DMC. The DMC is an independent committee established to provide safety oversight and to advise the sponsor regarding actions necessary to protect study participants and the scientific integrity of the study. The recommendations made by the DMC (i.e., dose escalation, etc.) will be forwarded to the sponsor for final decision. The sponsor will forward such decisions, which may include summaries of safety data which are not endpoints, to regulatory authorities, as appropriate.

The sponsor will appoint a DMC for the periodic review of available study data. The members of the DMC serve in an individual capacity and provide their expertise and recommendations. The primary responsibilities of the DMC are to (1) periodically review and evaluate the accumulated study data for participant safety, study conduct, and progress, (2) make recommendations to the sponsor concerning the continuation, modification, or termination of the study and (3) suggest dose for POC portion of the study as described in the study protocol.

The DMC considers study-specific data as well as relevant background knowledge about the disease, test agent, or patient population under study. The DMC is responsible for defining its deliberative processes, including event triggers that would call for an unscheduled review, stopping guidelines, unmasking (unblinding), and voting procedures prior to initiating any data review. The DMC is also responsible for maintaining the confidentiality of its internal discussions and activities as well as the contents of reports provided to it.

The DMC will have access to unblinded treatment information during the clinical trial. Details regarding management and process of this committee are found in the DMC Charter.

The DMC may recommend termination of BOS161721 treatment arm or the entire BOS161721 MAD/POC trial for any safety concern that is felt to outweigh potential benefits. The recommendation must be supported by the sponsor as indicated in the DMC Charter.

3.5. Interim Analysis (IA)

One IA is planned for this study and will be conducted at the time of dose selection for POC. The POC dose will be selected based on cumulative available safety, tolerability, PK and PD data based on the MAD data. Since there is no IA during the POC part of the study, there is no impact on the type 1 error.

3.6. Timing of Analyses

The following analyses are planned:

- An IA will be performed during the last cohort of the MAD portion to determine dose selection for the POC portion.
- The final analysis will be performed when all patients have either completed the POC safety follow up or withdrawn from the study.
- Safety analyses will be performed for DMC reviews throughout the study to evaluate accumulating safety data and make dose escalation decisions. The frequency and details of the content and format of the safety review meetings are described in the DMC charter and SAP and applicable mock shells.

4. DATA ANALYSIS CONSIDERATIONS

All analyses will be conducted based on SAS 9.3 or higher.

All data in the database will be presented in by-patient data listings.

Unless otherwise stated, all listings will be sorted by study phase, treatment group, patient number, and assessment date (and time, if available). If assessment date is missing, chronological visit will be used for sorting.

Unless stated otherwise, continuous data will be summarized by treatment group based on number of patients (n), mean, median, standard deviation (SD), 25^{th} percentile/first quartile (Q1), 75^{th} percentile/third quartile (Q3), minimum value, and maximum value.

The geometric % coefficient of variation (CV) is calculated as 100*sqrt(exp(SDlog^2)-1) where SDlog is the SD of log-transformed values.

Unless stated otherwise, categorical data will be summarized by treatment group using n and percentage based on the number of non-missing values.

- The number of missing values will be presented as a separate category with no percentage, but only if one or more patients are missing data.
- Counts of zero will be presented without percentages.

Statistics will be presented in the summary tables based on the following:

- Minimum and maximum: same number of significant digits as the raw value
- Mean, median, Q1, and Q3: one additional decimal place to that reported for minimum and maximum
- SD: two additional decimal places than presented for the minimum and maximum
- Percentages: reported to one decimal place with the exception of 100% which will be displayed without decimals.
- P-values will be reported to four decimal places. If the value is below 0.0001 it will be noted as < 0.0001; if the value above 0.9999 it will be noted as > 0.9999.

Unless otherwise noted, statistical inference will be based on a 10% significance level (i.e. 90% confidence intervals will be produced).

Statistical testing will be performed on data from the Phase 2 study.

Numbering for data displays will be based on ICH E3.

4.1. Stratification and Covariates

The effects of noncompliance, background therapy use, treatment discontinuations, premature withdrawal from study and covariates will be assessed to determine the impact on the general applicability of results from this study. Exploratory analyses of the data may be conducted as deemed appropriate to evaluate factors for which analyses are not controlled.

Continuous efficacy endpoints will be assessed via analysis of variance (ANOVA) or analysis of covariance (ANCOVA) when applicable, adjusting for the baseline value.

Sensitivity analyses of the primary efficacy endpoint may be performed with covariates defined based on the subgroups in section 4.2.

4.2. Evaluation of Subgroups

CCI [redacted]

- CCI [redacted]

[redacted]

The patient will be excluded from subgroup analyses if they are missing the subgroup parameter at baseline. In case of missing data, LOCF methods consistent with the primary efficacy analysis described in Section 7.1 will be used for determination of baseline CLASI and SLEDAI.

4.3. Multiple Comparisons and Multiplicity

The Type 1 error rate will not be adjusted for multiplicity.

5. GENERAL DATA HANDLING CONVENTIONS

5.1. Reference Dates

- Screening date is defined as the electronic case report form (eCRF) provided date on which a patient was screened for trial entry.
- Informed consent date is the date the patient signed the informed consent form.
- Randomization date is defined as the date on which the patient is randomized to study treatment.
- Treatment start date is defined as the date of first dose of study drug.
- Treatment end date is defined as the date of last dose of study drug.

- Age will not be calculated and will come directly from the eCRF. The eCRF uses the informed consent date as its reference date for age calculation.
- Safety data, such as AEs and laboratory assessments will use the treatment start date as a reference date.
- Efficacy data will use the randomization date as a reference date.
- Study day will be based on treatment start date as a reference date.

5.2. Study Day and Duration Variables

Reference date calculations will be defined as follows, assuming non-missing dates:

- date of interest – reference date + 1 when the date of interest is on or after the reference date;
- otherwise, date of interest – reference date.

If either date is missing or incomplete, reference date calculations will not be performed. Date imputation will be performed as identified in Section 5.5.

Study day will either have a negative value if collected before dosing or a positive value if collected on or after the day of drug dosing; there will be no study day zero.

Duration of time is dependent on reference dates and will be calculated in a manner similar to that of the reference date calculation. Duration on study is defined as the end of study date – informed consent date + 1. Duration of treatment is defined as treatment end date – treatment start date + 1, where treatment end date is the date of last dose of study drug.

Time-to-event endpoints are followed until first event or censoring. As a result, event time will be calculated as: event or censoring date – reference date + 1. These are further described in Section 7.

When reporting time-to-event data (TTMF and time to BILAG A flare or >1 BILAG B flare) or duration outcomes, the results (in days) above will be converted to an appropriate unit. When reporting in months it will be divided by 30.4375; for reporting in weeks it will be divided by 7; and for reporting in years it will be divided by 365.25.

5.3. Study Time Periods

Where applicable, data reporting will be classified by the following study periods for analysis:

- Pre-treatment is defined as the period prior to a patient's treatment start date.
- On-treatment is defined as the period between a patient's treatment start date and treatment end date, inclusive.
- Post-treatment is defined as the period of time following the on-treatment period.

5.4. Baseline, Post-Baseline Changes, and LOCF

Unless stated otherwise, baseline and post-baseline change values will be based on the following:

- Baseline will be based on the last non-missing value collected prior to or on the date/time of first dose of study drug.
- Post-baseline values will be those collected after the date/time of first dose of study drug.
- Change from baseline is defined as: value – baseline value.
- Percentage change from baseline is defined as: [(value – baseline value) / baseline value]*100. Patients with baseline values of zero are excluded due to division by zero.
- Most extreme change: The maximum most extreme change will be the maximum post-baseline change value; the minimum most extreme change will be the smallest post-baseline change value. This calculation will consider all assessments collected after the first dose of study drug, scheduled or unscheduled.
- LOCF:
    - Imputation rules for assessments that include multiple components:
        - In case of missing laboratory components, lab values from the most recent unscheduled visit <=14 days after the nominal visit will be utilized. If lab values are not available from an unscheduled visit <=14 days after the visit, laboratory data will be carried forward from the most recent laboratory assessment <=30 (scheduled or unscheduled) days prior to the nominal visit.
        - In case of missing non-laboratory components, no imputation will be performed.
    - In case of missing the entire score for an assessment, imputation will be based on LOCF. LOCF will only be implemented with the most recent prior visit (including unscheduled assessments). If the most recent prior visit is missing, occurred more than 30 days before (current visit date – most recent prior visit date > 30), or is not evaluable, LOCF is not implemented. See Section 7 for further details. Additional details regarding LOCF for components for specific assessments are described in Section 7.
    - Observed values do not incorporate any imputations and are presented as collected.

5.5. Imputation of Partial Dates

Adverse Events

- If the AE start date is completely missing, or if the patient was not treated, no imputation will be conducted.
- If the AE start date is missing day and month, do the following:
    - If the treatment start date is missing or the AE start year does not fall in the same year as that of the treatment start date or if the AE record contains information to

indicate that the event ended before the treatment start date (e.g. the AE end date month and year are earlier than the treatment start date or the full AE end date is known and occurs earlier than the treatment start date), then set the AE start month and day to January 1st.
    - Otherwise, set the AE start date to the treatment start date.
- If only the AE start day is missing, do the following:
    - If the study treatment start date is missing or the AE start month and year does not fall in the same month and year as that of the treatment start date or if the AE contains information to indicate that the event ended before the treatment start date, then set the AE start day to the 1st day of the month of the AE start date.
    - Otherwise, set the AE start date to the treatment start date.
- AE end dates will not be imputed.

Prior, Concomitant, and Subsequent Medications (Excluding Corticosteroids)

- The imputation rules for AE start dates will be used for medication start dates.
- Medication stop dates will be imputed as follows:
    - If the stop date is only missing the day, then the stop day is the last day of the month
    - If the stop date is missing both the day and month and the year matches the last study date, then the stop month and day is the earlier of the last study date and December 31
    - If the stop date is missing both the day and month and the year is earlier than the last study date, then the stop month and day is December 31
    - If the stop date is completely missing, no imputation is performed, and the medication will be classified as a concomitant medication for patients who were treated.

Corticosteroids

Due to the criticality of the corticosteroid usage, handling of partial start/stop dates will be performed on a case-by-case basis as it will require medical review of the patient records. Such determinations will be made in a blinded fashion prior to database lock and approved by the Sponsor.

Diagnosis Date

- If the diagnosis date is completely missing, no imputation will be conducted.
- If the month and day are completely missing but year is present, do the following:
    - Set missing month and day to July 1st
    - If July 1st of the non-missing year is after informed consent date, set month and year to January 1st
- If only the day is missing, set to the 1st of the month

5.6. Multiple Assessments and Visit Windows

Nominal visits (e.g. those identified by the study eCRF) will be the basis of summarization and statistical analysis unless otherwise specified. Unscheduled data may be included in summaries of most extreme, baseline, LOCF, and endpoint values; summaries of specific abnormalities any time post-baseline; and patient data listings

5.7. Treatment Group Display

Treatment groups will be displayed with the following columns. See the mock shells for additional details and a visual representation of the treatment group display. All overall columns will be excluded from efficacy outputs.

- Phase 1b MAD: Cohort 1 BOS 20mg
- Phase 1b MAD: Cohort 2 BOS 60mg
- Phase 1b MAD: Cohort 3 BOS 120mg
- Phase 1b MAD: Placebo
- Phase 1b MAD: Total BOS (All patients treated with BOS161721, excludes placebo)
- Phase 2 POC: BOS 120mg
- Phase 2 POC: Placebo
- Overall: Total BOS 120mg (All patients treated with BOS161721 120mg in MAD and POC)
- Overall: Total BOS (Patients treated with BOS161721 in MAD and POC, excludes placebo)
- Overall: Total Placebo (Placebo patients in MAD and POC, excludes patients treated with BOS161721)

5.8. Missing Data

Missing data imputation for the efficacy endpoints are discussed in the corresponding efficacy sections. AE, concomitant medication, and corticosteroid date imputations are described in Section 5.5. Missing data and other analysis conventions for noncompartmental analysis of pharmacokinetics will follow standard imputation methods, as described in a separate Clinical Pharmacology SAP. Otherwise, missing data will not be imputed.

6. STUDY PATIENT DATA

6.1. Analysis Populations/Sets

FAS: Defined as all patients who receive at least one dose (partial or complete) of study treatment and have at least one evaluable post-baseline efficacy evaluation (BILAG, PGA,

SLEDAI-2K, CLASI, ACR-28, SLICC/ACR, CCI , or CCI). FAS analyses will be conducted on the basis of the assigned (randomized) treatment. The FAS will be used as the basis for the primary efficacy analysis.

Safety Analysis Set (SS): Defined as all patients who receive at least one dose (partial or complete) of study treatment and have at least one evaluable post-baseline safety evaluation (adverse events, concomitant medications, vital signs, physical examination, ECG, or laboratory assessment). Safety analyses will be conducted on the basis of actual treatment received as reported on the eCRF. The SS will be the basis of all safety reporting.

Per Protocol Set (PP): Defined as all patients from the FAS except those with major deviations to the protocol deemed to impact the analysis of the primary endpoint. These deviations will be identified based on blinded data review prior to study unblinding. Patients with more than one missed dose and patients with both Day 180 and Day 210 efficacy assessments (BILAG, SLEDAI-2K, or PGA) missing will be excluded from the PP. PP analyses will be conducted on the basis of the assigned (randomized) treatment. Additional information on inclusion into this population can be found in Section 6.3.

6.2. Patient Disposition

Disposition data will be summarized for all randomized patients. The number of patients in each analysis set population, the number of patients who discontinued study treatment including reasons, and the number of patients who discontinued the study including reasons will be summarized. The number of patients who completed the study, duration on study treatment in days and duration on study in days will also be summarized. Data will be presented overall, by study phase, and treatment group.

A by-patient listing of patient disposition data including reason for discontinuation, if applicable, will be presented for all randomized patients. A by-patient listing of randomization details will also be provided for all randomized patients.

A by-patient listing of inclusion/exclusion criteria and screen failures will be produced for all enrolled patients. Screen failures will otherwise not be included in any analyses.

6.3. Protocol Deviations

Protocol deviations will be identified and classified as major or minor before database lock and unblinding. Major protocol deviations may include but are not limited to:

- Violation of Inclusion/Exclusion Criteria
- Use of prohibited therapies
- Incorrect treatment

Major protocol deviations that necessitate exclusion from the efficacy analysis in the PP population will be identified prior to database lock and unblinding.

Protocol deviations will be summarized by any deviation, deviation category, and major/minor designation. Protocol deviations will be summarized overall, by study phase, and treatment group for all randomized patients.

A listing of protocol deviations, including protocol version under which the deviation occurred, will be provided for all randomized patients. The protocol deviation listing will also include a column to indicate if the deviation resulted in the subject being excluded from the PP set based upon blinded clinical review. A separate listing of COVID-19 related deviations also will be provided.

6.4. Demographic and Baseline Characteristics

Patient demographics and baseline characteristics will be summarized in the SS. These will include age (years), sex (Male / Female), ethnicity (Hispanic or Latino / Not Hispanic or Latino), race (American Indian or Alaska Native / Asian / Black or African American / Native Hawaiian or Pacific Islander / White / Other), geographic regions (North America / South or Central America / Eastern Europe), baseline height (cm), baseline weight (kg), and baseline BMI (kg/m^2). Age is reported as collected in the clinical database and will also be categorized as a categorical variable (≤45, >45-<65, ≥65) for reporting. Patient eligibility for central review will be summarized. Childbearing potential and birth control methods for female and male patients will be summarized for the DMC meetings and interim analysis only. Demographics and baseline characteristics will be summarized overall, by study phase, and treatment group.

Baseline disease characteristics will also be summarized in the FAS. The SLEDAI total score, PGA, CLASI total activity score, CLASI total damage score, ACR-28 sum of swelling, ACR-28 sum of tenderness, ACR-28 sum of active joints (swelling and tenderness in the same joint) and SLICC/ACR damage index total score will be summarized at baseline. Frequencies and percentages of patients with SLEDAI presence by component, ACR-28 loss of functional range interfering with life and ACR-28 impairment of basic activities from inflammation as well as BILAG grades by body organ system will be presented. Frequency and percentage of patients with each organ system disease on BILAG SLE history will be summarized. Frequency and percentage of patients with SLICC criteria for SLE will be summarized. Frequencies and percentages of patients with CLASI Activity Score (<8 vs. >=8), SLEDAI-2K (<10 vs. >=10) and baseline corticosteroid use (0, <7.5, >= 7.5, <10, >=10 mg/day), will be presented. In case of missing data, LOCF methods consistent with the primary efficacy analysis described in Section 7.1 will be used for baseline disease characteristics summaries.

Frequencies and percentages of patients will also be presented for the following characteristics:

- Positive anti-ds-DNA
- Low C3 and/or C4

- Positive anti-dsDNA and normal C3 and normal C4
- Negative anti-dsDNA and low C3 and/or C4
- Positive anti-dsDNA and low C3 and/or C4
- Negative anti-dsDNA and normal C3 and normal C4
- Other

In addition, frequencies and percentages of patients will be presented for the following characteristics:

- SLEDAI-2K ≥ 10 and negative** anti-dsDNA and normal C3 and normal C4
- SLEDAI-2K ≥ 10 and positive** anti-dsDNA and normal C3 and normal C4
- SLEDAI-2K ≥ 10 and negative** anti-dsDNA and low C3 and/or C4
- SLEDAI-2K < 10 and positive** anti-dsDNA and low C3 and/or C4
- SLEDAI-2K < 10 and positive** anti-dsDNA and normal C3 and normal C4
- SLEDAI-2K < 10 and negative** anti-dsDNA and low C3 and/or C4
- SLEDAI-2K < 10 and negative** anti-dsDNA and normal C3 and normal C4
- SLEDAI-2K ≥ 10 and positive** anti-dsDNA and low C3 and/or C4

** Positive anti-dsDNA is defined as ≥ 15 IU/mL. Negative anti-dsDNA is defined as < 15 IU/mL

Time from diagnosis to informed consent in years defined as (informed consent date - date of diagnosis as recorded on the screening BILAG assessment)/365.35 will be summarized. Partial dates of diagnosis will be handled as defined in section 5.5. Baseline disease characteristics will be summarized overall by study phase and treatment group.

Demographics and baseline characteristics will be listed for all randomized patients. A separate by-patient listing will be created for childbearing potential and birth control for all randomized patients.

The following conversions and equations will be used as applicable:

Height (in cm) = height (in inches) * 2.54

Weight (in kg) = weight (in lbs) * 0.4536

BMI (kg/m^2) = weight(kg)/[height$(m)^2$]

6.5. Medical History

The incidence of medical history will be summarized by system organ class (SOC) and preferred term (PT) overall, by study phase, and treatment group in the SS. Patients will be counted once per SOC (likewise for PT within SOC). Medical history terms will be coded based on the Medical Dictionary for Regulatory Affairs (MedDRA). Medical history will be reported in

descending order of overall incidence of SOC and then in descending order of overall incidence for PT within the SOC.

Medical history will be presented in data listings for all randomized patients. Medical history coded terms will be provided, including the SOC and PT.

The MedDRA version used is defined in the study data management plan.

6.6. Prior and Concomitant Medication and Procedures

The incidence of medication use will be summarized by World Health Organization (WHO) Drug Dictionary anatomic therapeutic class (ATC) Level 2 classification (i.e. therapeutic main group) and preferred name. A patient will be counted only once at each level of reporting. Prior medications are those which have been identified to have been discontinued prior to the first dose of study medication. Concomitant medications are those which have been identified to have been taken at any point during the period of time between the first dose of study drug and last dose of study drug + 90 days, including medications which started prior to first dose of study drug that are ongoing at first dose. Subsequent medications are those which have been started after the last dose date + 90 days. The data management plan specifies the version of WHO Drug used.

Partial dates will be imputed according to Section 5.5 for the determination of prior, concomitant, and subsequent medications.

Concomitant medication use will be summarized and presented overall, by study phase, and treatment group in the SS.

A separate summary will be prepared for corticosteroids. Corticosteroids will be presented overall, by study phase and treatment group in the SS. The average daily dose of CS (prednisone equivalent) at each visit including summary statistics as well as categorical summaries for <5 mg/day, ≥5 mg/day to ≤10 mg/day, and >10 mg/day and for dosages of interest relevant to Phase 2, i.e. 0 mg/day, <7.5 mg/day, ≥7.5 mg/day, and for dosages of interest to Phase 1b, i.e. <10 mg/day, and ≥10 mg/day will be presented. Average daily dose is calculated as described in Section 7.2.4. Prednisone equivalence is also provided in the same section.

All prior and concomitant medication data will be provided together in a by-patient listing including the verbatim and preferred drug name and WHO ATC Class (ATC Level 2) for all randomized patients. Past biological, immunosuppressant, antimalarial and corticosteroid SLE medication history data will be listed separately for all randomized patients.

Concomitant procedures will be presented in a data listing for all randomized patients.

6.7. Study Drug Exposure and Compliance

The total dose administered (mL) and the number of doses received will be summarized as continuous variables. The number and percentage of patients that received any dose and the number and percentage of patients who received the total planned dose at each nominal, protocol specified, visit (Day 0 / Day 30 / Day 60 / Day 90 / Day 120 / Day 150 / Day 180) will be summarized. The number and percentage of patients who received 1-7 doses (mutually exclusive) will also be summarized categorically. The denominator for the number of patients dosed and the number of doses received will be the number of patients in the SS. The denominator for the number of patients who received the total planned dose will be the number of patients dosed at the corresponding nominal visit.

The percent compliance will be summarized descriptively and is defined as 100 * (sum of doses received / 7). Percent compliance will also be calculated categorically as <70%, ≥70% to <85%, ≥85% to <100%, ≥100%.

Relative dose intensity will be summarized descriptively and is defined as 100 * (total actual dose summed across all visits)/total planned dose summed across all planned visits) per the below table.

Dose Level (mg)	Dose Level (mL)*	Total Planned Dose Across All 7 Visits (mL)	Relative Dose Intensity Calculation
20	0.4	2.8	(Total actual dose in mL summed across all visits / 2.8) X 100
60	1.2	8.4	(Total actual dose in mL summed across all visits / 8.4) X 100
120	2.4	16.8	(Total actual dose in mL summed across all visits / 16.8) X 100

* Dose Level in mL is based on 50mg/mL administration.

The duration of exposure in days will also be summarized descriptively and is defined as the last dose date – first dose date +1.

Study drug exposure and compliance will be presented overall, by study phase, and treatment group in the SS.

By-patient listings of BOS161721 and placebo dosing data will be produced for all patients in the SS.

7. EFFICACY

7.1. Primary Efficacy Endpoint and Analyses

Safety will be the primary endpoint for the Phase 1b MAD portion of the study and is discussed in Section 10. Efficacy will be evaluated for the Phase 1b MAD portion of the study as an exploratory endpoint and details will be included in relation to the Phase 2 POC portion of the study within this section. Efficacy analyses will be presented by study phase and treatment group.

7.1.1. The proportion of patients with a SRI-4 response at Day 210

The primary efficacy endpoint for the POC is the proportion of patients who achieve a SRI-4 response at Day 210 in the FAS.

The SRI-4 is a composite index of SLE disease improvement that consists of scores derived from the SLEDAI-2K, the BILAG 2004 Index, and the PGA. SRI-4 response is defined as:

1) ≥ 4-point reduction from baseline in SLEDAI-2K total score

AND

2) No new severe disease activity (BILAG A organ score) or more than 1 new moderate organ score (BILAG B) compared with baseline

AND

3) No deterioration from baseline in the PGA by ≥ 30mm.

The test for treatment group difference in proportion of patients who achieve a SRI-4 response at Day 210 will be based on Pearson's chi-square test. The difference in percentages of patients who achieve SRI-4 response (BOS161721 vs. Placebo) will be presented with its corresponding (Wald) 90% CI. The number and percentage of SRI-4 responders, along with an exact (Clopper-Pearson) 90% CI, will be presented by treatment group. If an expected count for the Pearson chi-square test falls below 5, a Fisher's exact test will be used to assess the treatment difference and the CI will be based on an exact method (Santner Snell). The analyses will be repeated for each of the three components contributing to the SRI-4 response. The primary analysis of the primary efficacy endpoint will be performed in the FAS and presented by treatment group. The analysis will be repeated in the PP analysis set. SRI-4 Response at Day 210 will also be repeated by all subgroups in Section 4.2 and treatment group for POC Phase 2 only.

The superiority of BOS161721 relative to placebo will be evaluated. If the proportion of SRI-4 responders for the selected POC dose in BOS161721 arm is higher than that of the POC placebo

arm, then BOS161721 will be considered superior to placebo if the 2-sided p-value is less than 0.10.

Evaluability/Non-Responders:

Medication Failures: Patients who received prohibited medications or undergo unallowable corticosteroid (CS) usage will be considered "medication failures" and will be treated as non-responders at time points on and following the first date of prohibited medication or unallowable CS usage for the primary efficacy analysis. Prohibited medications are detailed in Appendix 4 of the study protocol. Prohibited medications and unallowable CS usage will be recorded as protocol deviations in the clinical database.

A patient is considered to have unallowable CS usage if any of the following is not met:

- A maximum daily dose of 30 mg/day of oral CS will be acceptable for eligibility for the study. For patients whose only SLE treatment is steroids, their stable steroid dose must be at least 10 mg/day and no more than 30 mg/day for a minimum of 6 weeks at time of randomization on Day 0
- Topical steroids may be used, but the dose must be stable for at least 6 weeks prior to Day 0, and be maintained at a constant dose throughout the study duration until the rashes resolve. PRN topical steroids are not permitted
- Once the patient has received the first dose of study drug (Day 0), tapering of oral steroids will be highly encouraged and should be continually evaluated during the protocol-allowed tapering windows (Day 0 through Day 150) with the target of achieving a CS daily dose of < 7.5 mg and < Day 0 dose between Day 150 and Day 210
- Between Day 150 and Day 210 (i.e., within 60 days of primary endpoint assessments), oral CS doses must be held constant
- CS Burst for SLE-related Indications
    - After Day 0 (first dose of study drug), a maximum of 1 oral CS "burst" equivalent to ≤ 40 mg/day prednisone for increased SLE disease activity will be allowed between Day 0 and Day 120, which must be tapered down to the baseline (Day 0) CS dose or lower within 14 days of initiation of the "burst". Any "burst" continuing after Day 120 or occurring after Day 120 is considered a protocol deviation.
    - Alternatively, a single intramuscular (IM) dose of methylprednisolone (< 40 mg) is permitted during this period.
- CS Burst for Non-SLE-related Indications
    - A single treatment of oral prednisone equivalent of ≤ 40 mg/day for 14 days is permitted for a non-SLE indication, though it must be completed prior to Day 120. No long acting steroid injections are permitted.
    - Note: Treatment with inhaled CS are allowed for the treatment of non-SLE-related indications only (e.g., for asthma).

Any other increase from baseline of CS dose or any other systemic use of CS of any kind (including intra-articular and intravenous administration) are not permitted from Day 0 through Day 210 and will result in the patient being considered a medication failure and will be treated as non-responders for the primary efficacy analysis. There is no restriction of CS usage after Day 210.

The determination of medication failures will be reviewed using blinded data and finalized prior to unblinding. Once a patient is determined to be a non-responder, the patient will remain a non-responder for the remainder of the study.

Out of Window Assessments:

- If a patient has an efficacy assessment (i.e. BILAG, SLEDAI-2K, PGA, CLASI, etc.) more than 30 days since the previous assessment (current visit – prior visit > 30), the corresponding result will not contribute to LOCF analyses.
- Since the Day 180 nominal visit is assigned as "Day 180/End of Treatment" in the clinical database, Day 180 nominal assessments must be >=159 days (180 – 21 day sponsor allowed treatment window) since first dose to be considered for "Day 180" nominal assessments and also to be considered for LOCF to Day 210 in the case of missing Day 210 data for all efficacy analyses. Note, under protocol amendment 6, patients will be allowed to be treated outside of window (after consultation with the sponsor) due to the ongoing COVID-19 pandemic.

Early Treatment Withdrawals/Discontinuations:
For early treatment withdrawal/discontinuation prior to study day 159 (180 days – 21 day treatment window), patient will be considered a non-responder after the date of early treatment withdrawal/discontinuation assuming the patient was not considered a medication failure prior to or on the date of early withdrawal/discontinuation for the primary efficacy analysis.

For early treatment withdrawal/discontinuation on or after study day 159 (180 days – 21 days) such that patient has data for nominal visit Day 180 but not for nominal visit Day 210, Day 180 data will be carried forward to Day 210 for assessment of SRI-4 response.

For early treatment withdrawal/discontinuation on or after study day 159 (180 days – 21 days) such that patient does not have data for nominal visit Day 180 and Day 210, the patient will be considered a non-responder on Day 210.

For early treatment withdrawal/discontinuation on Day 210 such that patient has data for nominal visit Day 210, the Day 210 data will be used for assessment of SRI-4 response.

Missing Data:

- **Pre-Treatment:** If a patient is missing data during the pre-treatment period for SLEDAI-2K, PGA, or BILAG, the patient will be excluded from the primary efficacy analysis.

- **Overall SRI-4 Components:** Missing data at Day 210 for overall components for SLEDAI-2K, BILAG, and PGA will be addressed by employing a LOCF analysis within each overall component as described in Section 5.4. For example, if a patient has data for SLEDAI-2K and BILAG at Day 210 but is missing PGA at Day 210, the PGA component would employ LOCF as described in Section 5.4. If the overall component is missing from Day 180 through Day 210 and the patient is not a medication failure on or prior to Day 210 and did not withdraw/discontinue treatment prior to Day 180, the SRI-4 response at Day 210 will be considered not evaluable.

- **Individual Items Contributing to Overall SRI-4 Components:** Data collection and scoring will be performed in the clinical database for BILAG, SLEDAI-2K, and PGA. If missing individual items contributing to any of the overall components occurs for Day 210, the overall component will be recalculated per the Table 1.

Table 1

SRI-4 Component	Method for Carrying Forward Individual Items Contributing to SRI-4 Component Result/Score
BILAG Grade	BILAG Grading is evaluated and entered into the clinical database by the medical monitoring team. Grades are adjudicated by independent subject matter experts. The medical monitoring team and the adjudication team will utilize BILAG index results to determine BILAG Grades. Missing data will be handled by the medical monitoring team as follows. If lab values are missing in the BILAG Index at Day 210, lab values from the most recent unscheduled visit <=14 days after Day 210 will be utilized for determination of the BILAG grade in the applicable organ system and entered into the clinical database. If lab values are not available from an unscheduled visit <=14 days after Day 210, laboratory data will be carried forward from the most recent laboratory assessment (scheduled or unscheduled) <=30 days prior to the Day 210 nominal visit for determination of BILAG grade. If the lab components for BILAG are missing from Day 210 and no unscheduled laboratory assessments occur within 14 days after or 30 days prior to the Day 210 assessment, the lab components from Day 180 will be utilized if Day 180 occurred within 30 days of Day 210. Otherwise, no individual BILAG Index results will be carried forward. If the BILAG grade for a given organ system is missing at Day 210, the grade will be carried forward from the Day 180 assessment. If the Day 180 BILAG grade is also missing, occurred more than 30 days before Day 210 (Day 210 visit date – Day 180 visit date >

	30), or is not evaluable, the Day 210 grade will be considered not evaluable.
SLEDAI-2K	If lab values are missing at Day 210, lab values will be imputed based on the below hierarchy of determination: (1) If any unscheduled laboratory assessments occurred <=14 days after the Day 210 visit, the laboratory values at the most recent unscheduled laboratory assessment will be used for the determination of missing laboratory SLEDAI-2K criteria. For example, if platelet count is missing at Day 210 and an unscheduled laboratory assessment occurs 10 days after the Day 210 visit date, the platelet count from the unscheduled laboratory assessment would be used to determine if the patient had thrombocytopenia at Day 210 for SLEDAI-2K. (2) Else if #1 does not occur or lab values are not available at the unscheduled visit noted in #1, the laboratory data will be carried forward from the most recent laboratory assessment (scheduled or unscheduled) <=30 days prior to the Day 210 nominal visit to make the SLEDAI-2K determination. For example, if a patient is missing thrombocytopenia and leukopenia SLEDAI-2K components at Day 210, they have an unscheduled lab assessment 10 days earlier with platelet counts populated and WBC not populated, and they have a Day 180 assessment within 30 days of the Day 210 visit with both platelets and WBC present, the platelet value from the unscheduled visit would be carried forward to the Day 210 for thrombocytopenia component designation and the WBC value would be carried forward from the Day 180 assessment to Day 210 for the Leukopenia component designation. (3) If the lab components for SLEDAI-2K are missing from Day 210 and no unscheduled laboratory assessments occur within 14 days after or 30 days prior to the Day 210 assessment, the lab components from Day 180 will be carried forward, if available and if Day 180 occurred within 30 days of Day 210. Otherwise, no non-lab individual SLEDAI-2K components will be carried forward. If non-lab individual components are missing at Day 210, the entire SLEDAI-2K score from Day 180 will be carried forward. If Day 180 is missing, occurred more than 30 days before Day 210, or is not

	evaluable, LOCF is not implemented and the Day 210 score will be considered not evaluable. In the case of missing SLEDAI-2K lab components at Day 210, where any SLEDAI-2K lab component is carried forward (or back) per above, the SLEDAI-2K total score will be recalculated outside of the clinical database as the sum of: (1) All of the non-missing weights for each descriptor at Day 210 (2) All individual lab component weights carried forward from the Day 180 visit or unscheduled visit, as applicable. If any components are still missing at Day 210 after carrying forward available lab components, the Day 210 SLEDAI-2K score will not be recalculated and will be considered not evaluable. Note: Clinical input may be obtained prior to database lock and unblinding to make a final determination of SLEDAI-2K criteria in the case of carrying forward (or back) individual lab component scores.
PGA	Individual components are not applicable for PGA and will not be carried forward. Total PGA VAS measurement will be carried forward from the Day 180 visit in the case of missing PGA measurement at Day 210. If total PGA score is also missing from Day 180, occurred more than 30 days before Day 210, or is not evaluable, LOCF is not implemented and the Day 210 visit will be considered not evaluable.

Table 2 summarizes the primary analysis and details several sensitivity analyses which may be performed:

Table 2

Analysis	**Analysis Population**	**Criteria**
Primary Analysis LOCF for missing data, Medication Failures=Non-Responder	FAS	The primary analysis of the primary efficacy endpoint will be performed based on LOCF methods for missing data. (1) Medication failures are considered non-responders on and after the date of medication failure (2) Early withdrawals/discontinuations prior to study day 159 will be considered non-responders after the date of early withdrawal/discontinuation

		a. If patient is a medication failure prior to withdrawal/discontinuation, patient will be considered a non-responder on and after the date of medication failure. b. In case of early withdrawal/discontinuation after study day 159 but before Day 210 such that patient has Day 180 data but is lacking Day 210 data in a patient without medication failure, the Day 180 data will be carried forward to Day 210. (3) Patients with missing Day 180 and Day 210 data without prior medication failure and without early withdrawal/discontinuation prior to study day 159 are considered not evaluable and are excluded from analysis
Sensitivity Analysis 1: Observed (no LOCF for missing data), Medication Failures=Non-Responder	FAS	A sensitivity analysis of the primary efficacy endpoint may be performed based on purely observed data. No LOCF imputations will be performed for missing data. Additional criteria are as follows: (1) Medication failures are considered non-responders on and after the date of medication failure (2) Early withdrawals/discontinuations prior to Day 210 will be considered non-responders after the date of early withdrawal/discontinuation a. If patient is a medication failure prior to withdrawal/discontinuation, patient will be considered a non-responder on and after the date of medication failure. (3) Patients with missing Day 210 data without prior medication failure and without early withdrawal/discontinuation prior to Day 210 are considered not evaluable and are excluded from the sensitivity analysis If lab component scores are carried forward from Day 180 to Day 210 in order to assign a Day 210 grade in the primary analysis, the Day 210 assessment will be considered not evaluable for this sensitivity analysis.
Sensitivity Analysis 2: LOCF for missing data, Exclude Medication Failures	FAS	A sensitivity analysis of the primary efficacy endpoint may be performed based on LOCF methods for missing data detailed for the primary efficacy analysis. Additional/modified criteria are as follows: (1) Medication failures are considered not evaluable and are excluded from the sensitivity analysis on and after the date of medication failure

		(2) Early withdrawals/discontinuations prior to study day 159 will be considered not evaluable and will be excluded from analysis after the date of early withdrawal/discontinuation a. If patient is a medication failure prior to withdrawal/discontinuation, patient will be considered not evaluable and excluded on and after the date of medication failure. b. In case of early withdrawal/discontinuation after study day 159 but before Day 210 such that patient has Day 180 data but is lacking Day 210 data in a patient without medication failure, the Day 180 data will be carried forward to Day 210. (3) Patients with missing Day 180 and Day 210 data are considered not evaluable and are excluded from the sensitivity analysis
Sensitivity Analysis 3: Observed (no LOCF for missing data), Exclude Medication Failures	FAS	A sensitivity analysis of the primary efficacy endpoint may be performed based on purely observed data. No LOCF imputations will be performed for missing data. Additional criteria are as follows: (1) Medication failures are considered not evaluable and are excluded from the sensitivity analysis on and after the date of medication failure (2) Early withdrawals/discontinuations prior to Day 210 will be considered not evaluable and are excluded from the sensitivity analysis after the date of early withdrawal/discontinuation a. If patient is a medication failure prior to withdrawal/discontinuation, patient will be considered not evaluable and excluded on and after the date of medication failure. (3) Patients with missing Day 210 data are considered not evaluable and are excluded from the sensitivity analysis If lab component scores are carried forward from Day 180 to Day 210 in order to assign a Day 210 grade in the primary analysis, the Day 210 assessment will be considered not evaluable for this sensitivity analysis.
Sensitivity Analysis 4: LOCF for missing data, Medication Failures=Non-Responder, Logistic Regression	FAS	CCI [redacted]

Additional sensitivity analyses may be performed to consider the effect of missing data.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

A summary of disposition of SRI-4 response per the primary analysis at day 210 will be presented by study phase and treatment group in the FAS and PP. Sensitivity analysis 2 (LOCF, Exclude Medication Failures) will also be performed for disposition of SRI-4 response at day 210. These summaries will include the number of SRI-4 responders and the number of SRI-4 non-responders with a breakdown of reasons for non-response (early withdrawal, medication failure, <4-point reduction from baseline in SLEDAI-2K total score, new BILAG 1A/2B, and PGA worsening). Denominators for percentages will be the number of patients eligible for SRI-4 response analysis (patients not deemed not evaluable) defined as those with data for PGA, SLEDAI-2K and BILAG at Day 210 (whether or not due to LOCF, FAS analysis).

A listing of SRI Response data at Days 30, 60, 90, 120, 150, 180 and 210 will be presented in the FAS. A listing of SRI-4 Response at Day 210 (Primary and Sensitivity Analyses) will also be presented in the FAS. A listing of patients with medication failures, early withdrawals/discontinuations and those excluded from the primary efficacy analysis will also be provided for all randomized patients.

7.2. Secondary Efficacy Endpoints and Analyses

PK and PD measures are the secondary endpoints for the Phase 1b MAD part of the study and exploratory endpoints for the POC part of the study. Those analyses are described in Section 8.

7.2.1. Proportion of Patients with SRI-4 Response at Each Visit

The primary analysis methods described in Section 7.1.1 and Table 2 will be repeated for each nominal visit and will be presented by study phase and treatment group in the FAS. Other analyses described in Section 7.1.1 may also be performed. The 3 components will not be summarized separately by visit. In case of missing data, LOCF methods will be implemented as described in section 5.4 and Table 1. LOCF will only be implemented with the most recent prior nominal visit. If the most recent prior nominal visit is missing, occurred more than 30 days before, or is not evaluable, LOCF is not implemented. For example, if Day 30 SLEDAI assessment is complete but Day 60 and Day 90 SLEDAI assessments are missing, Day 60 will implement LOCF from Day 30 assuming Day 30 is within 30 days of Day 60 and Day 60 SLEDAI will be included in the analysis but Day 90 will not carry forward from Day 60 and thus Day 90 will be considered not evaluable and will be excluded from the analysis. Similarly, if Day 150 assessments are complete but Day 180 and Day 210 are missing, Day 180 will implement LOCF from Day 150 assuming Day 150 is within 30 days of Day 180 and will be included in the analysis but Day 210 will not carry forward from Day 180, will be considered not evaluable, and will be excluded from the analysis.

If a patient is a medication failure, the patient is considered to be a non-responder at timepoints on or after the first date of prohibited medication/unallowable CS usage.

If a patient discontinues/withdraws early, the patient will be considered a non-responder after the date of discontinuation/withdrawal (unless the patient had a prior medication failure in which case the patient is considered a non-responder on and after the date of medication failure).

The denominators for analyses will be the number of patients in the FAS with an evaluable post-baseline SRI-4 assessment defined as those with data for PGA, SLEDAI-2K and BILAG (whether observed or imputed via LOCF depending on the analysis) at a given visit.

A graphical display of SRI-4 response by study phase, treatment group, and visit in the FAS will be presented.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.2. Proportion of Patients with SRI-5 Response at Each Visit

The SRI-5 is a composite index of SLE disease improvement that consists of scores derived from the SLEDAI-2K, the BILAG 2004 Index and the PGA. Response based on the SRI-5 is defined by:

1) ≥ 5-point reduction from baseline in SLEDAI-2K total score

AND

2) No new severe disease activity (BILAG A organ score) or more than 1 new moderate organ score (BILAG B) compared with baseline

AND

3) No deterioration from baseline in the PGA by ≥ 30mm.

The primary analysis described in Section 7.1.1 and Table 2 will be repeated for each nominal visit as described in Section 7.2.1 based on the SRI-5 response definition and will be presented by study phase and treatment group in the FAS. The 3 components will not be summarized separately by visit for SRI-5 response. Other analyses described in Section 7.1.1 may also be performed.

The denominators for analyses will be the number of patients in the FAS with an evaluable post-baseline SRI-5 assessment defined as those with data for PGA, SLEDAI-2K and BILAG (whether observed or imputed via LOCF depending on the analysis) at a given visit.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.3. Proportion of Patients with SRI-6 Response at Each Visit

The SRI-6 is a composite index of SLE disease improvement that consists of scores derived from the SLEDAI-2K, the BILAG 2004 Index and the PGA. Response based on the SRI-6 is defined by:

1) ≥ 6-point reduction from baseline in SLEDAI-2K total score

AND

2) No new severe disease activity (BILAG A organ score) or more than 1 new moderate organ score (BILAG B) compared with baseline

AND

3) No deterioration from baseline in the PGA by ≥ 30mm.

The primary analysis described in Section 7.1.1 and Table 2 will be repeated for each nominal visit as described in Section 7.2.1 based on the SRI-6 response definition and will be presented by study phase and treatment group in the FAS. The 3 components will not be summarized separately by visit for SRI-6 response. Other analyses described in Section 7.1.1 may also be performed.

The denominators for analyses will be the number of patients in the FAS with an evaluable post-baseline SRI-6 assessment defined as those with data for PGA, SLEDAI-2K and BILAG (whether observed or imputed via LOCF depending on the analysis) at a given visit.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.4. Oral Corticosteroid (CS) Reduction

The average daily CS dose will be presented over time starting with Day 0 as baseline. Oral CS usage for all visits will be based on the average daily dose, defined as the sum of the corticosteroid doses taken each day from the date of the previous visit to the day before the date of the current visit (converted to prednisone equivalent dose) / the duration between the two visits. For example, the average daily dose at Day 0 (Baseline) is the sum of daily doses of CS from the screening visit to the day prior to Day 0 divided by the duration between the screening visit to the day prior to Day 0, inclusive. Similarly, the average daily dose at Day 30 is the sum of daily doses of CS from Day 0 to the day prior to Day 30 divided by the duration between Day 0 to the day prior to Day 30, inclusive. Partial dates for CS use will be imputed based on details in Section 5.5. Oral CS will be converted to prednisone dose equivalence for analysis of reduction of oral CS use. Common prednisone dose equivalents are provided in Table 3.

Table 3

Medication	Prednisone Dose Equivalence
Prednisone	1 mg
Cortisone	5 mg
Hydrocortisone	4 mg
Deflazacort	1.2 mg
Meprednisone	1 mg
Prednisolone	1 mg
Methylprednisolone	0.8 mg
Methylprednisolone Acetate	0.8 mg
Triamcinolone	0.8 mg
Triamcinolone Acetonide	0.8 mg
Budesonide	0.25 mg
Dexamethasone	0.16 mg
Bethamethasone	0.16 mg

7.2.4.1. Proportion of Patients with a sustained reduction of oral CS between Day 120 and Day 210 (MAD)

The proportion of patients with a sustained reduction of oral corticosteroid, defined as ≤ 10 mg/day prednisone equivalent and < Day 0 dose, between Day 120 and Day 210 determined based on actual daily CS usage on and in between the nominal visit dates will be summarized for MAD Phase 1b by treatment group in the FAS. To qualify as sustained reduction, the actual daily CS usage needs to meet the definition above every day between nominal visit dates for Day 120 and Day 210 inclusive. LOCF and imputation for missing values will not be performed. Otherwise, statistical methods described in Section 7.1.1 will be implemented. Patients who withdraw prior to Day 210 or are designated as a medication failure on or prior to Day 210 or miss the Day 120 and/or the day 210 visit will not be considered to have sustained reduction in CS. Denominators for percentages are the number of patients taking oral CS at baseline.

7.2.4.2. Proportion of Patients with a sustained reduction of CS between Day 150 and Day 210 (POC)

The proportion of patients with a sustained reduction of oral corticosteroid, defined as ≤ 7.5 mg/day prednisone equivalent and < Day 0 dose, between Day 150 and Day 210 determined based on actual daily CS usage on and in between the nominal visit dates will be summarized for POC Phase 2 by treatment group in the FAS. To qualify as sustained reduction, the actual daily CS usage needs to meet the definition above every day between nominal visit dates for Day 150 and Day 210 inclusive. LOCF and imputation for missing values will not be performed. Otherwise, statistical methods described in Section 7.1.1 will be implemented. Patients who withdraw prior to Day 210 or are designated as a medication failure on or prior to Day 210 or miss the Day 150 and/or Day 210 visit will not be considered to have sustained reduction in CS. Denominators for percentages are the number of patients taking oral CS at baseline.

7.2.4.3. Percent Reduction in CS Administration from Day 0 through Day 210 (POC)

The percent reduction in CS administration from Day 0 through Day 210 determined based on the average daily CS usage will be summarized for POC Phase 2 by treatment group and visit. Average daily dose will be calculated for nominal visits as defined in Section 7.2.4. This will be performed for patients in the FAS who received ≥ 7.5 mg/day prednisone equivalent at Day 0. Patients considered medication failures will be excluded from the percent reduction analysis at timepoints on and after the first date of medication failure.

7.2.5. Proportion of new or recurrent BILAG A flare or >1 BILAG B flare through Day 210

The proportion of patients with a new or recurrent BILAG A flare or >1 BILAG B flares through Day 210 will be summarized by study phase and treatment group in the FAS. Statistical methods described in Section 7.1.1 will be implemented. BILAG flares will be adjudicated, recorded in the clinical database, and summarized as recorded in the clinical database. Flare assessments will be performed on BILAG grades whether or not designation of a BILAG grade required carrying component data forward by the assessor. However, no LOCF methods will be implemented with respect to flare data itself. When discrepancies between the two-reviewers occurs, the adjudicated flare determination will be used for analyses. The proportion of new or recurrent BILAG A flare or >1 BILAG B flare will be summarized overall and by nominal visit. The denominator for percentages of patients with a new or recurrent flare will be patients in the FAS with a post-baseline BILAG assessment at the given visit.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

Results as recorded in the clinical database will be presented in separate data listings for BILAG SLE history, BILAG-2004 index, and BILAG grading for all randomized patients.

7.2.6. Physician's Global Assessment (PGA)

The PGA is used to assess investigator's general impression on the patient's overall status of SLE disease activity via visual analogue scale (100 mm) with 0 being "very good, asymptomatic and no limitation of normal activities" with 100 mm being "most severe possible disease ever seen in all SLE patients."

7.2.6.1. Proportion of Patients with PGA Worsening at Each Visit

The proportion of patients with PGA worsening, defined as an increase of ≥30mm from baseline, will be summarized by study phase and treatment group in the FAS overall (at any post-baseline nominal visit) and by nominal visit. Statistical methods described in Section 7.1.1 will be implemented. In case of missing data, LOCF methods will be implemented as described in section 5.4 and Table 1 and in Section 7.2.1 for the "at each visit" analyses.

The analyses described in Section 7.1.1 and Table 2 (Primary Analysis methods) will be repeated for each nominal visit as described in Section 7.2.1 based on the PGA worsening definition and will be presented by study phase and treatment group in the FAS.

The denominators for analyses will be the number of patients in the FAS with an evaluable post-baseline PGA assessment (overall analysis) or with evaluable PGA assessment (whether or not imputed via LOCF depending on the analysis) at a given visit (nominal visit analysis).

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.6.2. Results, Changes, and Percentage Change from Baseline in PGA

The analysis of PGA will be performed separately using an ANCOVA model with change or percentage change from baseline to each nominal visit as the dependent variable, treatment group as the factor, and the corresponding baseline PGA result as a covariate. The p-value for the difference between the BOS161721 and placebo groups at each nominal visit is based on the F-test.

Descriptive statistics for PGA will be summarized separately at baseline and each visit. Change and percentage change from baseline will be summarized at each post-baseline visit. The least squares (LS) mean estimate, its standard error and 90% CI will be tabulated by treatment group and for the difference between treatment groups (BOS161721 versus placebo) for change and percentage change at each visit. In case of missing data, LOCF methods will be implemented as described in Section 5.4 and Table 1 and in Section 7.2.1.

Two separate approaches will be taken based on handling for medication failures and treatment withdrawals:

Approach 1: Corresponding to Primary Analysis Methods (Medication Failure LOCF): If a patient is a medication failure, the most recent assessment prior to the medication failure date will be carried forward to all subsequent assessments. In the case of patients who withdraw/discontinue early without prior medication failure, the most recent data on or prior to withdrawal/discontinuation will be carried forward to all subsequent assessments. In case of early withdrawal/discontinuation after Day 180 but before Day 210 such that patient has Day 180 data but is lacking Day 210 data in a patient without medication failure, the Day 180 data will be carried forward to Day 210.

Approach 2: Corresponding to Sensitivity Analysis 2 Methods (Exclude Medication Failure): If a patient is a medication failure, all assessments on and after medication failure date will be considered not evaluable and will be excluded. In the case of patients who withdraw/discontinue early without prior medication failure, all assessments after the early withdrawal date will be considered not evaluable and will be excluded. In case of early withdrawal/discontinuation after Day 180 but before Day 210 such that patient has Day 180 data but is lacking Day 210 data in a patient without medication failure, the Day 180 data will be carried forward to Day 210.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

PGA observed results will be presented in a data listing for all randomized patients.

7.2.7. Proportion of Patients with BICLA Response

The BICLA is a responder index developed to measure response to therapy, and it includes scores from the BILAG, SLEDAI-2K, and PGA. BICLA response is defined as:

1) at least 1 gradation of improvement in baseline BILAG 2004 scores in all body systems with all A [severe disease] scores falling to B [moderate], C [mild] or D [no activity] OR all B [moderate disease] scores falling to C [mild] or D [no activity])

AND

2) no new BILAG A or more than 1 new BILAG B scores

AND

3) no worsening of total SLEDAI-2K score from baseline

AND

4) ≤ 10% deterioration in PGA score

AND

5) no medication failure.

Statistical methods described in Section 7.1.1 will be implemented. In case of missing data, LOCF methods will be implemented as described in section 5.4 and Table 1 and in Section 7.2.1 for the "at each visit" analysis. The primary analysis and sensitivity analysis 2 (LOCF for missing data, Exclude Medication Failures) methods described in Section 7.1.1 and Table 2 will be repeated overall and as described in Section 7.2.1 for Days 210, 240, and 270 based on the BICLA response definition and will be presented by study phase and treatment group in the FAS.

Other analyses described in Section 7.1.1 may also be performed.

The denominators for analyses will be the number of patients in the FAS with an evaluable post-baseline BICLA assessment defined as those with data for PGA, SLEDAI-2K and BILAG (whether observed or imputed via LOCF depending on the analysis) (overall analysis) or at a given visit (nominal visit analysis).

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.8. Cutaneous Lupus Erythematosus Area and Severity Index (CLASI)

The CLASI is a comprehensive tool for assessment of disease activity and damage in cutaneous lupus, shown to be valid, reliable, and sensitive to changes in disease activity.

The total activity score will be calculated by summing the scores on the left side of the form (erythema, scale/hypertrophy, mucous membrane, alopecia – recent hair loss and alopecia clinically not obviously scarred).

The total damage score will be calculated by:

(1) Sum the responses for "Dyspigmentation" in all anatomical locations

(2) If a response for "Report duration of dyspigmentation after active lesions have resolved" is "Dyspigmentation usually lasts at least 12 months", then multiply the sum in #1 by 2.

(3) If a response for "Report duration of dyspigmentation after active lesions have resolved" is "Dyspigmentation usually lasts less than 12 months", then the sum in #1 will not change.

(4) If the sum in #1 is 0 and a response for "Report duration of dyspigmentation after active lesions have resolved" is missing, then the sum in #1 will not change.

(5) Determine the Dyspigmentation Score based on the results from 2, 3, or 4.

(6) Sum the remaining scores on the right side of the form (scarring/atrophy/panniculitis, alopecia - scarring of the scalp judged clinically).

(7) Sum the scores from #5 and #6 to obtain the total damage score.

If any individual scores which contribute to the total activity score or the total damage score are missing, the total score in question will not be calculated. If total activity score or total damage

score are missing at a visit, LOCF methods will be implemented as described in Section 5.4 and Section 7.2.1 for the "at each visit" analyses.

The mean and SD of the change from baseline in CLASI total activity and total damage scores will be presented graphically over time by study phase and treatment group.

7.2.8.1. Proportion of Patients with CLASI Response

CLASI response is defined as 50% improvement from baseline in CLASI. A CLASI-A response applies to "A" (total activity) and a CLASI-B response applies to "B" (total damage) score.

Statistical methods described in Section 7.1.1 will be implemented. In case of missing data, LOCF methods will be implemented as described in Section 5.4 and 7.2.1 for the "at each visit" analysis. Individual components of the CLASI will not be carried forward; however, the total activity score or total damage score may be carried forward from the most recent prior visit. If the most recent prior visit is missing, occurred more than 30 days before (current visit date – most recent prior visit date > 30), or is not evaluable, LOCF is not implemented. The primary analysis methods (LOCF for missing data, medication failure=non-responder) as described in Section 7.1.1 and Table 2 will be performed at each visit based on the CLASI response definition and will be presented by study phase and treatment group in the FAS.

CCI [redacted]

The analysis will also be repeated in line with Sensitivity Analysis 2 (LOCF for missing data, exclude medication failures) in Section 7.1.1.

Other analyses described in Section 7.1.1 may also be performed.

The denominators for analyses will be the number of patients in the FAS with an evaluable post-baseline CLASI assessment (whether observed or imputed via LOCF depending on the analysis) at a given visit.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.8.2. Results, Changes, and Percentage Change from Baseline in CLASI

The analysis of CLASI total activity score and CLASI total damage score will be performed separately using an ANCOVA model with change or percentage change from baseline to each nominal visit as the dependent variable, treatment group as the factor, and the corresponding baseline CLASI score as a covariate. The p-value for the difference between the BOS161721 and placebo groups at each nominal visit is based on the F-test.

Descriptive statistics for the activity score and damage score will be summarized separately at baseline and each visit. Change and percentage change from baseline will be summarized at each post-baseline visit. The LS Mean estimate, its standard error and 90% CI will be tabulated

by treatment group and for the difference between treatment groups (BOS161721 versus placebo) for change and percentage change at each visit. Statistical methods described in Section 7.1.1 will be implemented. In case of missing data, LOCF methods will be implemented as described in Section 5.4 and Section 7.2.1 for the "at each visit" analyses. Individual components of the CLASI will not be carried forward; however, the total activity score or total damage score may be carried forward from the most recent prior visit. If the most recent prior visit is missing, occurred more than 30 days before (current visit date – most recent prior visit date > 30), or is not evaluable, LOCF is not implemented. The analyses will be performed in line with the two approaches outlined in Section 7.2.6.2. CCI [redacted]

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

CLASI observed results will be presented in data listings for all randomized patients.

7.2.9. Results, Changes, and Percentage Changes from Baseline in Swollen, Tender, and Active Joints ACR-28

The ACR-28 joint count evaluates the number of tender and swollen joints in the shoulder, elbow, wrist, hand, knee joints. Joints of the feet, hip and ankle are excluded. The sum of active joints is also evaluated and is defined as the total number of joints with swelling and tenderness in the same joint. The joint sums for tender and swollen joints are recorded in the database. The sum of active joints will be derived.

The analysis of ACR-28 Swollen and Tender Joints for the sum of tenderness (left and right), the sum of swelling (left and right), and sum of active joints (left and right) will be performed separately using an ANCOVA model with change and percentage change from baseline to each nominal visit as the dependent variable, treatment group as the factor, and the baseline ACR-28 score as a covariate. The p-value for the difference between the BOS161721 and placebo groups at each nominal visit is based on the F-test.

Descriptive statistics for the tenderness, swelling, and derived active joint sums will be summarized separately at baseline and each visit. Change and percentage change from baseline will be summarized at each post-baseline visit. The LS Mean estimate, its standard error and 90% CI will be tabulated by treatment group and for the difference between treatment groups (BOS161721 versus placebo) for change and percentage change at each visit.

Statistical methods described in Section 7.1.1 will be implemented. In case of missing data, LOCF methods will be implemented as described in Section 5.4 and in Section 7.2.1 for the "at each visit" analyses. Individual components of ACR-28 will not be carried forward; however, the total ACR-28 score may be carried forward from the most recent prior visit. If the most recent prior visit is missing, occurred more than 30 days before (current visit date – most recent prior visit date > 30), or is not evaluable, LOCF is not implemented. The analyses will be performed in line with the two approaches outlined in Section 7.2.6.1.

The mean and SD of the change from baseline in ACR-28 sum of tenderness, sum of swelling, and sum of active joints based on primary analysis methods will be presented graphically over time by study phase and treatment group.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

Swollen, tender, and active joints ACR-28 observed results will be presented in data listings for all randomized patients.

7.2.10. Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K)

The SLEDAI-2K is a validated instrument that measures disease activity in SLE patients at the time of the visit and in the previous 30 days. It is a global index and includes 24 clinical and laboratory variables that are weighted by the type of manifestation, but not by severity. The total score falls between 0 and 105, with higher scores representing increased disease activity. The SLEDAI-2K has been shown to be a valid and reliable disease activity measure in multiple patient groups. A SLEDAI 2K of 6 or more generally represents moderately to severely active disease.

7.2.10.1. Proportion of Patients with a SLEDAI-2K Response

The proportion of patients with a SLEDAI-2K response, defined as ≥ 4-point reduction from baseline in SLEDAI-2K total score, will be summarized by study phase and treatment group in the FAS overall and by nominal visit. Statistical methods described in Section 7.1.1 will be implemented. In case of missing data, LOCF methods will be implemented as described in section 5.4 and Table 1 and in Section 7.2.1 for the "at each visit" analyses.

The analyses described in Section 7.1.1 and Table 2 (Primary Analysis methods and Sensitivity Analysis 2 methods) will be repeated for each nominal visit as described in Section 7.2.1 based on the SLEDAI-2K response definition and will be presented by study phase and treatment group in the FAS.

The denominators for analyses will be the number of patients in the FAS with an evaluable post-baseline SLEDAI-2K assessment (overall analysis) or with evaluable SLEDAI-2K assessment (whether observed or imputed via LOCF depending on the analysis) at a given visit (nominal visit analysis).

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.10.2. Results, Changes, and Percentage Change from Baseline in SLEDAI-2K

The analysis of SLEDAI-2K will be performed using an ANCOVA model with change and percent change from baseline to each nominal visit as the dependent variable, treatment group as the factor, and the baseline SLEDAI-2K score as a covariate. The SLEDAI-2K total score is

calculated and provided in the clinical database. In case of missing data, SLEDAI-2K will be re-calculated outside the clinical database after application of imputation rules.

Descriptive statistics for the SLEDAI-2K total score will be summarized separately at baseline and each visit. Change and percentage change from baseline will be summarized at each post-baseline visit. The LS mean estimate, its standard error and 90% CI will be tabulated by treatment group and for the difference between treatment groups (BOS161721 versus placebo) for change and percentage change at each visit. The p-value for the difference between the BOS161721 and placebo groups at each nominal visit is based on the F-test.

Statistical methods described in Section 7.1.1 will be implemented,. In case of missing data, LOCF methods will be implemented as described in section 5.4 and Table 1 and in Section 7.2.1 for the "at each visit" analyses. The analyses will be performed in line with the two approaches outlined in Section 7.2.6.1.

Approach 1 corresponding to the primary analysis methods (LOCF for missing data, medication failure LOCF) will be repeated for subgroups CCI as defined below:

- CCI

7.2.11. Results, Changes, and Percentage Change from Baseline in SLICC/ACR Damage Index

The SLICC/ACR damage index is a validated instrument to assess damage, defined as irreversible impairment, continuously persistent for 6 months (ascertained by clinical assessment), occurring since the onset of lupus, and it is based on a weighted scoring system. This index records damage occurring in patients with SLE regardless of cause, with demonstrated content, face, criterion, and discriminant validity. It will be performed on Days 0 and 180.

The analysis of SLICC/ACR Damage Index total score will be performed using an ANCOVA model with change and percentage change from baseline to Day 180 as the dependent variable, treatment group as the factor, and the baseline SLICC/ACR Damage Index total score as a covariate. The SLICC/ACR Damage Index total score is calculated and provided in the clinical database. A missing total SLICC/ACR Damage Index score will not be imputed. Observed results will be presented for the change and percentage change from baseline analysis. The p-value for the difference between the BOS161721 and placebo groups at each nominal visit is based on the F-test.

Descriptive statistics for the SLICC/ACR Damage Index total score will be summarized separately at baseline and Day 180. Change and percentage change from baseline will be summarized at Day 180. The LS Mean estimate, its standard error and 90% CI will be tabulated by treatment group and for the difference between treatment groups (BOS161721 versus placebo) for change and percentage change at Day 180.

In case of medication failure or early withdrawal/discontinuation prior to Day 180, the Day 180 assessment will be considered not evaluable and will not contribute to the analysis at that visit.

These analyses will also be performed for the MAD Phase 1b with the exception of the statistical testing.

Observed SLICC/ACR Damage Index data will be presented in a by-patient listing for all randomized patients. SLICC Criteria for SLE at screening will also be presented in a by-patient listing for all randomized patients.

7.2.12. Medication Failure

7.2.12.1. Proportion of Patients with Medication Failures

Patients who received prohibited medications or had unallowable CS usage as described in Section 4.6 of the study protocol will be considered “medication failures” and will be treated as non-responders at time points on and following the first date of prohibited medication or unallowable CS usage for the primary efficacy analysis and selected sensitivity and secondary efficacy analyses. The determination of medication failures and medication failure dates will be reviewed using blinded data and finalized prior to unblinding.

The number and proportion of patients with medication failures will be summarized by study phase and treatment group in the FAS overall and by visit. A patient will be counted as a medication failure at each visit occurring on and after the date in which a patient is considered a medication failure. Statistical methods described in Section 7.1.1 will be implemented for observed data only. LOCF and missing data imputation will not be performed.

The denominator is the number of patients in FAS (overall analysis) and the number of patients with non-missing visit data (for the by visit analysis).

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.12.2. Time to Medication Failure (TTMF)

Medication failure is defined in Section 4.6 of the study protocol and discussed in Section 7.1.1. TTMF will be computed as the event date (as described below) – randomization date + 1.

Kaplan-Meier methods will be used to estimate TTMF for each treatment group. Estimates of median TTMF will be provided along with 90% confidence intervals. Q1 and Q3 will also be tabulated. Comparisons between treatment groups will use the two-sided log-rank test. In addition, hazard ratio (HR) estimates with 90% CIs will be calculated from a Cox proportional hazards regression analysis that includes treatment as a factor. Censoring rules are defined as in Table 4.

Table 4

Situation	**Date of Event or Censoring**	**Outcome**
Patient is designated a medication failure (received prohibited meds or unallowable CS burst)	Date of first dose of medication leading to medication failure designation. These dates will be identified by the sponsor based on the prior and concomitant medication eCRF form.	Event
Patient withdraws from study (no medication failure)	Censored at date of withdrawal	Censored
Patient completes the study (no medication failure)	Censored at last study visit	Censored
Patient dies during study (no medication failure)	Censored at the date of death	Censored

Kaplan-Meier curves of TTMF will be plotted over time.

These analyses will also be performed for the MAD Phase 1b with the exception of the statistical testing.

Time to medication failure data will be provided in a by-patient listing in the full analysis set.

7.2.13. Duration of Longest SRI-4 Response

Duration of longest SRI-4 response will be computed for patients who have been identified as a responder at least once based on criteria defined in Section 7.1.1 utilizing LOCF methods consistent with the primary efficacy analysis. The duration of longest response will be defined as the longest period a patient meets the SRI-4 responder criteria at consecutive visits. This will be computed as the date of last consecutive SRI-4 response – date of first consecutive response + 1. If the last consecutive response is based on LOCF methods and the last consecutive response does not have a date, duration of longest SRI-4 Response will be computed as the (date of last non-missing consecutive SRI-4 response + 30) – date of first consecutive response + 1.

The duration of longest SRI-4 response will be analyzed using an analysis of variance (ANOVA) model with treatment group as the effect. The p-value for the difference between the BOS161721 and placebo group is based on the F-test.

The mean, SD, median, minimum, maximum, Q3, Q4, LS Mean estimate, its standard error and 90% CI will be tabulated by treatment group and for the difference between treatment groups (BOS161721 versus placebo).

If a patient has missed a visit between two scheduled visits with a response at both, the patient will be counted as having a response at the missed visit unless otherwise specified.

Duration of longest SRI-4 response will be displayed in the SRI response listing.

These analyses will also be performed for the MAD Phase 1b with the exception of the statistical testing.

7.2.14. Time to First BILAG A Flare or >1 BILAG B Flare through Day 210

Time to first BILAG A Flare or >1 BILAG B Flare will be computed as the event date (as described below) – randomization date + 1. If a patient is designated as a medication failure prior to or on their day 210 visit, they will be considered as having a BILAG flare (event) as of the date of first dose of the medication leading to the medication failure designation due to prohibited medication or unallowable CS burst.

Time to first BILAG A Flare or >1 BILAG B Flare will be analyzed with Kaplan-Meier methods described for TTMF and censoring as described in Table 5. Kaplan-Meier curves will be plotted over time.

Table 5

Situation	Date of Event or Censoring	Outcome
Patient has a BILAG A Flare or >1 BILAG B Flare (Noted as BILAG 1A/2B Flare) prior to or on their Day 210 visit	Date of first BILAG 1A/2B Flare	Event
Patient is designated a medication failure (received prohibited meds or unallowable CS burst) prior to or on their Day 210 visit	Date of first dose of medication leading to medication failure designation. These dates will be obtained from date of the medication failure identified by the sponsor based on the prior and concomitant medication eCRF form.	Event
Patient withdraws from study (no BILAG 1A/2B Flare, no medication failure) prior to or on their Day 210 visit	Censored at date of last non-missing BILAG flare assessment (prior to or on withdrawal) through Day 210	Censored
Patient completes the study (no BILAG 1A/2B Flare, no medication failure)	Censored at date of last non-missing BILAG flare assessment through Day 210	Censored
Patient dies during study (no BILAG 1A/2B Flare, no medication failure) prior to or on their Day 210 visit	Censored at the day of death	Censored

An event will be counted once in a patient who is a medication failure and also experiences a BILAG flare. Date of event will be the date of the earlier of the two reasons.

These analyses will also be performed for the MAD Phase 1b with the exception of the statistical testing.

Time to first BILAG A Flare or >1 BILAG B Flare data will be provided in a by-patient listing in the full analysis set.

7.3. Exploratory Endpoints and Analyses

Efficacy will be evaluated for the Phase 1b MAD portion of the study as an exploratory endpoints and details are noted throughout sections 7.1 and 7.2. Exploratory endpoints and analyses may be included in relation to the Phase 2 POC part of the study as described within this section.

7.3.1. CCI

CCI

CCI

CCI

CCI

8. PHARMACOKINETICS/PHARMACODYNAMICS/ IMMUNOGENICITY

Pharmacokinetics, pharmacodynamics, and immunogenicity are secondary endpoints for the Phase 1b MAD study and exploratory endpoints for the Phase 2 POC study.

PD results reported as less than the lower limit of quantification (LLOQ) will be counted as 0.5*LLOQ to calculate summary linear statistics and will be reported as <LLOQ in listings. All observed Immunogenicity and PD data, as provided by an external vendor, will be presented in by-patient listings.

SAPs describing analyses (pharmacokinetics [PK], PK/pharmacodynamics[PD], etc.) not covered within this SAP will be included within an Appendix of this document. The author, finalization date, signatories, and page numbering of this SAP are independent of those SAPs and vice versa. Only the last final versions will be included in the appendix just prior to study conclusion.

8.1. Pharmacokinetics

BOS161721 PK concentration and parameter analyses will performed by a separate PK vendor and described under a separate SAP.

8.2. Pharmacodynamics

Descriptive statistics for the observed results of the following PD endpoints/results will be summarized separately at baseline and each visit. Change from baseline and percentage change from baseline will be summarized at each post-baseline visit. In cases where both numeric and qualitative results are provided, the summaries will be based on the numeric results. In cases where only qualitative results are provided and results given only as titers, shifts will be summarized. Graphical presentations of mean and SD of the change from baseline over time will also be presented for the parameters in table 6.

Table 6

PD Endpoint	**Applicable Study Portion(s)**	**Parameters to be Included in Summary Tables and Figures ('NA (all)' is displayed if all collected parameters are to be included)**
Clinical PD Biomarkers		
CCI [redacted]	MAD and POC	NA (all)
C-Reactive Protein	MAD and POC	NA (all)
CCI [redacted]	MAD and POC	NA (all)
CCI [redacted]	MAD and POC	NA (all)
CCI [redacted]	MAD and POC	NA (all)
CCI [redacted]	MAD and POC	NA (all)
CCI [redacted]	MAD and POC	NA (all)

PD Endpoint	Applicable Study Portion(s)	Parameters to be Included in Summary Tables and Figures ('NA (all)' is displayed if all collected parameters are to be included)
Other Biomarkers		
Phosphorylated signal transducer and activator of transcription 3 (pSTAT3)	MAD only	% pSTAT3+ Lymphocytes –Stimulated; Ratio MEF (Stimulated / Unstimulated)
Leukocyte immunophenotype (B and NK Cells)	MAD only	CD19+% of CD45+ Lymphocytes; IgD+C27-% of CD19+; IgD+CD27-CD38++CD24++% of CD19+; IgD-CD27-% of CD19+; IgD+CD27+% of CD19+; IgD-CD27+% of CD19+; IgD-CD27+CD38++CD138-% of CD19+; IgD-CD27+CD38++CD138+% of CD19+; CD56+% OF CD45+ Lymphocytes
Leukocyte immunophenotype (T-Cells)	MAD only	CD4+CD8-% of CD3+; CXCR5+% of CD4+CD8-; PD-1+% of CD4+CD8-; ICOS+% of CD4+CD8-; CD25+CD127-% of CD4+CD8-; CXCR5+% of CD25+CD127-; PD-1+% of CD25+CD127-; ICOS+% of CD25+CD127-; CD8+CD4-% of CD3+
CCI	MAD only	CCI
CCI	MAD and POC	CCI

PD Endpoint	Applicable Study Portion(s)	Parameters to be Included in Summary Tables and Figures ('NA (all)' is displayed if all collected parameters are to be included)
CCI [redacted]	MAD	NA (all) – Listing only
CCI [redacted]	MAD and POC	NA (all) – Listing only (including fold change)

CCI [redacted]

CCI [redacted]

CCI [redacted]

Additional PD results may be plotted over time by treatment group and visit.

8.3. Immunogenicity

Immunogenicity data as collected on the eCRF will be provided in by-patient data listings including neutralizing antibodies (NAb) and ADA for all relevant timepoints.

All serum ADA samples that are confirmed positive in the ADA assay will be analyzed for the presence of NAb.

The proportion of patients with positive or negative ADA will be summarized by study phase, treatment group, and overall, at baseline (pre-dose), and post-baseline in the FAS. The frequency and percentage of patients with transient positive and persistent positive ADA will also be included in the summary. If warranted based on data, treatment-emergent transient positive and treatment-emergent persistent positive ADA will be summarized. The frequency and percentage of patients with confirmed positive, treatment-emergent confirmed positive, and negative ADA at each post-baseline timepoint will also be summarized. ADA designations and denominators are defined in table 7.

Table 7

Timepoint	ADA Designations	Definition	Denominator/Evaluable Patients
Baseline (Day 0 Pre-Dose)	Positive	Confirmed positive ADA at baseline	Number of patients with a baseline and at least one

	Negative	Negative ADA at baseline	post-baseline ADA assessment
Post-Baseline	Positive	Confirmed positive ADA at any time after baseline	Number of patients with at least one post-baseline ADA assessment
	Negative	Negative ADA at all times after baseline	
	Transient Positive[iv]	(1) Last ADA sampling timepoint is negative AND 2a OR 2b: (2a) Only one post-baseline confirmed positive ADA (excluding the last sampling timepoint) exists and last ADA negative sample is at least 16 weeks apart from the previous confirmed positive sample. OR (2b) Two or more post-baseline confirmed positive ADA where the first and last ADA confirmed positive samples are less than 16 weeks apart (irrespective of any negative ADA in between) and last ADA negative sample is at least 16 weeks apart from the previous confirmed positive sample.	
	Persistent Positive[iv]	1 OR 2 OR 3: (1) Two or more post-baseline confirmed positive ADA where the first and last ADA confirmed positive samples are more than 16 weeks apart (irrespective of any negative ADA in between) OR	

		(2) Confirmed positive ADA at the last post-baseline assessment OR (3) Negative ADA at the last post-baseline assessment and a prior confirmed positive ADA sample occurred less than 16 weeks before the last post-baseline negative ADA sample	
Treatment-Emergent	Positive	A post-baseline confirmed positive result exists with a titer >=4 times greater than the baseline titer OR Any post-baseline confirmed positive result after a negative baseline result Note: The number of treatment-emergent transient positive and treatment-emergent persistent positive ADA will sum to the total positive treatment-emergent ADA Note: If baseline ADA data is missing, the baseline value is treated as negative for this analyses	Number of patients with at least one post-baseline ADA assessment
	Transient Positive[iv]	<u>Scenario 1:</u> (1) Negative ADA at baseline AND (2) Last ADA sampling timepoint is negative AND (3a OR 3b): (3a) Only one post-baseline confirmed positive ADA (excluding the last sampling timepoint) exists and last ADA negative sample is at	

		least 16 weeks apart from the previous confirmed positive sample. OR (3b) Two or more post-baseline confirmed positive ADA where the first and last ADA confirmed positive samples are less than 16 weeks apart (irrespective of any negative ADA in between) and last ADA negative sample is at least 16 weeks apart from the previous confirmed positive sample. <u>Scenario 2:</u> (1) Confirmed positive ADA at baseline AND (2) Last ADA sampling timepoint is negative or is the same as baseline ADA levels AND (3a OR 3b): (3a) Only one post-baseline confirmed positive ADA (excluding the last sampling timepoint) exists and the titer at the post-baseline timepoint is >=4 times greater than the baseline titer OR (3b) Two or more post-baseline confirmed positive ADA where the first and last ADA confirmed positive samples are less than 16 weeks apart (irrespective of any negative ADA in between) and the titers at each of the two or more post-baseline confirmed positive	

		timepoints are >=4 times greater than the baseline titer Note: If baseline ADA data is missing, the baseline value is treated as negative for this analyses	
	Persistent Positive[iv]	<u>Scenario 1:</u> (1) Negative ADA at baseline AND (2 OR 3 OR 4): (2) Two or more post-baseline confirmed positive ADA where the first and last ADA confirmed positive samples are more than 16 weeks apart (irrespective of any negative ADA in between) OR (3) Confirmed positive ADA at the last post-baseline assessment OR (4) Negative ADA at the last post-baseline assessment and a prior confirmed positive ADA sample occurred less than 16 weeks before the last post-baseline negative ADA sample <u>Scenario 2:</u> (1) Confirmed positive ADA at baseline AND (2 OR 3 OR 4): (2) Two or more post-baseline confirmed positive ADA where the first and last ADA confirmed positive samples are more than 16 weeks apart (irrespective of any negative ADA in between) and the	

		titers at each of the two or more post-baseline confirmed positive timepoints are >=4 times greater than the baseline titer OR (3) Confirmed positive ADA at the last post-baseline assessment and the titer at the last post-baseline timepoint is >=4 times greater than the baseline titer OR (4) Negative/similar to baseline ADA at the last post-baseline assessment and a prior confirmed positive ADA sample occurred less than 16 weeks before the last post-baseline negative/similar to baseline ADA sample and the titer of the confirmed positive ADA is >=4 times greater than the baseline titer Note: If baseline ADA data is missing, the baseline value is treated as negative for this analyses	
Post-baseline at each timepoint	Positive	Confirmed Positive ADA at the post-baseline timepoint	Number of patients with ADA assessment at baseline and at the post-baseline timepoint
	Treatment-Emergent	<u>Scenario 1:</u> (1) Negative ADA at baseline (2) Confirmed positive ADA at the post-baseline timepoint <u>Scenario 2:</u> (1) Confirmed positive ADA at baseline	

		(2) Confirmed positive ADA at the post-baseline timepoint and the titer at the of the confirmed positive ADA is >=4 times greater than the baseline titer	
	Negative	Negative ADA at the post-baseline timepoint	
Overall	Positive	A patient with positive baseline (pre-dose) ADA or positive post-baseline ADA	Number of patients with a baseline or post-baseline ADA assessment
	Negative	A patient with negative baseline (pre-dose) ADA and negative post-baseline ADA at all time points	

The frequency and percentages of positive and negative neutralizing antibodies will be summarized in the FAS overall and at each relevant post-baseline timepoint (i.e., Day 15, Day 30, Day 60, Day 90, Day 180, Day 270). The denominators will be the number of patients with non-missing NAb data overall and at each timepoint, respectively.

The adverse event overview table from section 10.1 will be repeated by post-baseline positive and negative ADA as defined in table 7 in the safety analysis set. Additional adverse event analyses may be performed by post-baseline transient positive and persistent positive.

9. QUALITY OF LIFE

9.1. CCI

CCI [redacted]

CCI [redacted]

CCI

CCI

CCI

CCI

9.2. CCI

CCI

CCI

CCI

[redacted]

CCI

[redacted]

CCI

[redacted]

CCI

[redacted]

CCI

[redacted]

10. SAFETY

All safety analysis reporting will be based on the Safety Analysis Set. Unless otherwise noted, all safety analyses will be presented overall, by study phase, and treatment group.

10.1. Adverse Events

AEs will be collected and recorded for each patient from the date of the first dose of study drug until the end of their participation in the study, including the safety follow up period. Severity of adverse events will be graded by the investigator according to National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) version 4.03. AEs will also be assessed for relationship to study drug, and seriousness. Adverse events will be marked as AESIs, dose-limiting toxicities (DLTs), or injection site reactions (ISRs), where applicable. AEs will be considered treatment-emergent adverse events (TEAEs) if they started any time after first dose or if the year, month and day of start date were missing. Any missing severity assessments will be assumed to be Grade 3, missing relationship assessments will be assumed to be related, and missing seriousness assessments will be assumed as serious.

An overview of TEAEs will be produced, including counts and percentages of patients with any incidences of TEAE, TEAEs related to study treatment, SAEs, SAEs related to study treatment, any CTCAE grade 2 or higher TEAEs, any CTCAE grade 2 or higher TEAEs related to study treatment, any CTCAE grade 3 TEAEs, any CTCAE grade 3 or higher TEAEs related to study treatment, any CTCAE grade 4 TEAEs, any CTCAE grade 4 or higher TEAEs related to study treatment, TEAEs leading to study treatment discontinuation, treatment related TEAEs leading to study treatment discontinuation, TEAEs of special interest, related TEAEs of special interest, DLTs, related DLTs, TEAEs resulting in death, related TEAEs resulting in death, and injection site reactions. The maximum CTCAE grade (Grades 1-5) will also be included on the summary table.

Adverse events will be coded based on the MedDRA. Patients will be counted once per SOC (and likewise for PT within SOC). TEAEs will be summarized by SOC and PT in descending order of overall incidence of SOC (and PT within SOC). The MedDRA version used is defined in the study data management plan.

Summaries of adverse events by SOC and PT will include the following types:

- TEAEs;
- TEAEs related to study treatment;
- SAEs;
- Related Serious TEAEs;
- CTCAE Grade 3 or higher TEAEs;
- CTCAE Grade 3 or higher TEAEs related to study treatment;
- TEAEs leading to treatment discontinuation; and
- Dose-Limiting Toxicities
- TEAEs of Special Interest

A summary of TEAEs by SOC, PT, and maximum CTCAE grade will also be prepared. For these summaries, TEAEs will be sorted for each patient by SOC/PT and CTCAE grade; patients will be counted once within a SOC/PT based on their maximum CTCAE grade. Missing CTCAE grade will be counted as grade 3 and will also be presented in a 'Missing' row for completeness.

A summary of TEAEs by PT and maximum CTCAE grade will also be prepared with sections for any TEAE and TEAEs related to study treatment. For these summaries, patients will be counted once within a PT based on their maximum CTCAE grade. Missing CTCAE grade will be counted as grade 3. An additional summary of TEAEs by PT and maximum CTCAE grade will summarized for all TEAEs and TEAEs related to treatment for PTs with at least one patient experiencing a CTCAE grade 3 or higher and PTs where 10% or more of patients experience a CTCAE grade 1 or 2 for that PT. The summary will be sorted by the total incidence of the PT in the any TEAE section.

A comprehensive listing of all AEs will be provided in a by-patient data listing for the safety analysis set. In addition, the following listings will be provided for the safety analysis set:

- TEAEs related to study treatment;
- SAEs;
- TEAEs leading to treatment discontinuation; and
- AESIs
- CTCAE Grade 3 or Higher TEAEs

10.2. Clinical Laboratory Evaluations

Laboratory tests will be performed at times defined in the protocol Schedule of Assessments. Clinical chemistry and hematology parameters will be reported based on the International System of Units (SI). The following laboratory evaluations will be reported in data summaries, with asterisks indicating those that will be graded using NCI CTCAE. Grading will be performed by the Central Lab for relevant laboratory parameters based on the details provided in Appendix 13.11 . If a grade is not present in the central lab database for a graded lab parameter, the grade will be analyzed as grade 0.

Hematology: Hemoglobin*, Hematocrit, RBC count, RDW, MCV, MCH, MCHC, Platelet count*, WBC count*, CD4+ count*, Total neutrophils (Abs)*, Eosinophils (Abs), Monocytes (Abs), Basophils (Abs), Lymphocytes (Abs)*

Clinical chemistry: Alanine Aminotransferase*, Albumin*, Alkaline Phosphatase*, Aspartate Aminotransferase*, Bicarbonate, Bilirubin*, Blood urea nitrogen, C Reactive Protein, Calcium, Chloride, Creatine Kinase*, Creatinine*, Glucose (fasting)*, Glucose (random)*, Potassium*, Protein, Sodium*, Uric acid*, eGFR (Cockcroft-Gault)*, eGFR (MDRD)*, Total cholesterol (fasting)*, LDL-C (fasting), HDL-C (fasting), Triglycerides (fasting)*, Gamma Glutamyl Transferase (GGT)*. Per protocol, GGT is collected in cases of potential drug-induced liver injury.

Coagulation: Prothrombin Intl. Normalized Ratio*, Prothrombin Time

Urinalysis: pH, Glucose (qual), Protein (qual)*, Blood (qual)*, Ketones (qual), Nitrites (qual), Leukocyte esterase (qual), Urobilinogen (qual), Urine bilirubin (qual), Microscopy

Observed values, changes, and percentage changes from baseline for hematology, clinical chemistry, and coagulation laboratory evaluations will be summarized at each visit and most extreme change. Endpoint analyses will be based on changes from baseline assessments at Days 120 and 210.

The number and percent of patients with a CTCAE toxicity grade of 3 or higher will be tabulated by laboratory evaluation with defined CTCAE grading at each visit for hematology, clinical chemistry, coagulation, and urinalysis (protein only). Patients with missing values post-baseline will be excluded from the summary. The denominator for percentages will be the number of patients with a post-baseline assessment for the laboratory parameter in question at that visit.

Hematology, clinical chemistry, coagulation, and urinalysis (protein only) shift tables displaying the shift from baseline to the worst value of CTCAE grade will be presented based on the most extreme change as it relates to the relevant CTCAE definition. CTCAE relating to “high/hyper” conditions will depend on the maximum post-baseline value while CTCAE “low/hypo” will be reported based on the minimum post-baseline value. Separate shift tables will be prepared for parameters with bi-directional toxicity grading. Lab parameters which are optional, such as gamma glutamyl transferase, will not be included in the shift tables.

Hematology and clinical chemistry laboratory data which is not graded will also be summarized in shift tables of baseline to each visit based on range categories of low (below lower limit of normal [LLN], normal, and high [above upper limit of normal [ULN]]).

Urinalysis laboratory data (excluding protein) will also be summarized in a shift table of baseline to each visit based on range categories of low (below lower limit of normal [LLN]), normal, abnormal, and high (above upper limit of normal [ULN]).

Selected hematology and clinical chemistry laboratory results will be plotted over time for treatment group for MAD phase 1b portion of the study in the safety analysis set. The mean and SD of the change from baseline for selected hematology and clinical chemistry laboratory results will be plotted over time for the POC Phase 2 portion of the study in the safety analysis set.

All laboratory parameters as well as a summary of central laboratory tests and chest x-ray will be provided in patient data listings for all randomized patients. By patient listings of clinical chemistry and hematology data will be presented in the safety analysis set for the DMC and interim analysis only.

10.3. Other Safety Evaluations

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

The C-SSRS is a low-burden measure of the spectrum of suicidal ideation and behavior that was developed by Columbia University researchers for the National Institute of Mental Health Treatment of Adolescent Suicide Attempters Study to assess severity and track suicidal events through any treatment. It is a clinical interview providing a summary of both ideation and behavior that can be administered during any evaluation or risk assessment to identify the level and type of suicidality present. The C-SSRS can also be used during treatment to monitor for clinical worsening.[i] The C-SSRS evaluation will be performed as specified in the protocol Schedule of Assessments.

The C-SSRS has 11 binary (yes/no) outcomes corresponding to five categories of suicidal ideation, five categories of suicidal behavior, and the category of self-injurious behavior without suicidal intent:

Suicidal ideation is present if there is a "yes" response to any of the 5 suicidal ideation category questions. Suicidal behavior is present if there is a "yes" response to any of the 5 suicidal behavior category questions. Suicidal ideation or behavior is present if there is a "yes" response to any of the 10 suicidal or ideation or behavior category questions.

The number and percentage of patients with any post-baseline occurrence of the following ideations and behaviors will be presented overall, by study phase and treatment group:

- Suicidal ideation (overall and by individual question),
- Suicidal behavior (overall and by individual question),
- Suicidal ideation or behavior,

A similar tabulation of the occurrence of any of the above ideations and behaviors at baseline or as part of their lifetime history/past 12-month history (collected at the screening visit) will be presented overall, by study phase, and treatment group.

A shift table displaying the shift from worst pre-treatment category to the worst post-baseline category will be presented overall, by study phase, and treatment group. Best to worst C-SSRS category is defined in the following order: no suicidal ideation or behavior, suicidal ideation, and suicidal behavior. The worst post-baseline shift is calculated based on all post-baseline visits, including unscheduled visits. Worst pre-treatment will include all C-SSRS assessments prior to first dose date/time (screening or baseline). The denominator will be the number of patients with at least one non-missing baseline and post-baseline C-SSRS assessment.

Missing data will not be imputed.

All C-SSRS individual items will be presented in data listings for screening and post-screening (since last visit), separately. A separate listing will be presented for patients with suicidal ideation, suicidal behavior, or non-suicidal self-injurious behavior for all randomized patients.

10.3.2. Vital Signs

Vital signs include: heart rate (beats/min); temperature (°C); systolic and diastolic blood pressure (mmHg). Observed values, changes, and percentage changes from baseline for vital signs will be summarized at each visit and time point, as well as for most extreme change overall, by study phase, and treatment group.

All vital signs data will be presented in patient data listings for all randomized patients.

10.3.3. 12-Lead Electrocardiogram (ECG)

ECG will be assessed as specified in the protocol Schedule of Assessments. The following ECG parameters will be collected: PR interval (msec), QRS interval (msec), RR interval (msec), QT interval (msec), and QT interval corrected for heart rate using Fridericia's formula (QTcF) (msec). ECG parameters and ECG evaluation will come directly from the database and will not be calculated during analysis.

Observed values, changes, and percentage changes from baseline for ECG parameters will be summarized at each visit, as well as for most extreme change overall, by study phase, and treatment group.

The number and percent of patients with the following parameters at any post-baseline visit and at each visit will be summarized:

- Abnormal, not clinically significant ECGs
- Abnormal, clinically significant ECGs

The number and percent of patients with each ECG evaluation result will be summarized by visit. The worst ECG evaluation at any post-baseline visit will also be summarized in the order of: Normal, Abnormal, not clinically significant, Abnormal, clinically significant.

For the summary of incidence at any post-baseline visit, the denominators for percentages will include patients with one or more non-missing value at any post-baseline visit. For the by-visit summaries, the denominators for percentages will include patients with a non-missing value at that visit.

Investigator reported ECG result shifts from baseline to each visit and worst case post-baseline will be summarized. Worst case post-baseline will be based on the most abnormal observed value on or after the randomization date.

All ECG data will be presented in a by patient data listing for all randomized patients. A separate by patient listing for patients with abnormal, clinically significant 12-lead ECGs will also be provided in the safety analysis set for DMC meetings and the interim analysis only.

10.3.4. Physical Examinations

Targeted physical examination data and full physical examination data will be presented in patient data listings. A separate by patient listing for patients with abnormal, clinically significant physical examination findings will also be provided in the safety analysis set for DMC meetings and the interim analysis only.

11. CHANGES TO THE PLANNED ANALYSIS

The following changes were made to the planned analyses:

- Time to SRI-4 Response was originally a planned endpoint for the MAD and POC portions of the study. This was removed from the POC portion of the study upon Protocol Amendment 6 (30Apr2020); however, the endpoint was inadvertently maintained in the protocol for the MAD portion of the study. This analysis will not be performed for any study part.
- Rules for applying LOCF were vague within the protocol and were further detailed in this SAP. See Section 7 for further details.
- The PP was clarified to include additional criteria beyond the protocol description. The PP will exclude patients with more than one missed dose and missing both Day 180 and Day 210 efficacy assessments (BILAG, SLEDAI-2K, or PGA).
- Cytokine, chemokine and genotyping data were originally planned to be part of the MAD and POC portion of the study; however, data for these parameters is only available for the MAD portion of the study.
- The definition of the CLASI response in the protocol is vague as the protocol defines the CLASI response as 50% improvement from baseline in “A” (total activity) or “B” (total damage) scores. The SAP further calrifies the CLASI response as 50% improvement from baseline in CLASI and distinguishes a CLASI-A response from a CLASI-B response.

12. REFERENCES

i. Posner K, Melvin G a, Stanley B, Oquendo M a, Gould M. Factors in the assessment of suicidality in youth. *CNS Spectr*. 2007;12(2):156-162.

ii. CCI [redacted]

iii. Van Vollenhoven RF, Hahn BH, Tsokos GC, et al. Efficacy and safety of ustekinumab, an IL-12 and IL-23 inhibitor, in patients with active systemic lupus erythematosus: results of a multicentre, double-blind, phase 2, randomised, controlled study. Lancet. 2018;392(10155):1330-1339.

iv. Shankar G, et al. Assessment and Reporting of the Clinical Immunogenicity of Therapeutic Proteins and Peptides – Harmonized Terminology and Tactical Recommendations. AAPS Journal, 2014; 16(4):663.

v. International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH E3 Guideline: Structure and content of clinical study reports questions & answers (R1). 2012. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_QAs_R1_Step4.pdf. Accessed 13 Jan 2016.

vi. International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonised Tripartite Guideline: Structure and content of clinical study reports E3. Step 4. 1995. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf. Accessed 13 Jan 2016.

vii. International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH E9: Statistical principles for clinical trials. 1998. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf . Accessed 15 June 2016.

viii. FDA Guidance for Industry Systemic Lupus Erythematosus – Developing Medical Products for Treatment. 2010. https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm072063.pdf. Accessed 23February2018.

ix. Committee for Medicinal Products for Human use (CHMP). EMA Guideline on clinical investigation of medicinal products for the treatment of systemic lupus erythematosus and lupus nephritis. 2015. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/03/WC500184889.pdf. Accessed 04March2018.

x. CCI [redacted]

xi. Nilsson M, Suryawanshi S, Gassmann-Mayer C, Dubrava S, McSorley P, Jiang K. Columbia-Suicide Severity Rating Scale Scoring and Data Analysis Guide. V2 2013. https://cssrs.columbia.edu/wp-content/uploads/ScoringandDataAnalysisGuide-for-Clinical-Trials-1.pdf. Accessed 18May2018.

xii. CCI [redacted]

xiii. CCI [redacted].

xiv. CCI [redacted].

13. APPENDICES

13.1. COVID-19 Impact and Data Handling

The COVID-19 Pandemic resulted in contingency plans for BOS161721-02.

- **Safety Oversight:** In case a patient cannot return to the study site for the scheduled visit, the site staff will contact the patient remotely for safety follow up, for example, via telephone or video conferencing. The analysis of safety data will include all collected data regardless of remote data collection.

- **Central Laboratory:** In the case of courier issues that will prevent the protocol-required laboratory specimens to be sent to the study core laboratory, the site should have the safety laboratory specimens (Hematology, Chemistry, and Urinalysis) sent to their local laboratory for analysis and review by the study physician. The local laboratory data will not be entered into the clinical database or analyzed/presented in data summaries.

- **Investigational Product Dosing:** In cases when dosing cannot be performed during the protocol-designated windows, the Investigator should discuss each case with the Sponsor to determine whether it is a missed visit or whether the dosing can be performed outside protocol windows (as a protocol deviation). A minimum of 2 weeks must be maintained between consecutive doses. Patients who miss more than 2 doses of the study drug (whether consecutive or not) will be discontinued from the study and will undergo assessments as outlined in Section 5.5 of the study protocol (Premature Discontinuation). A sensitivity analysis of SRI-4 response using LOCF for missing data and/or observed case data may be performed to assess the impact of patients that did not follow the original protocol dosing schedule.

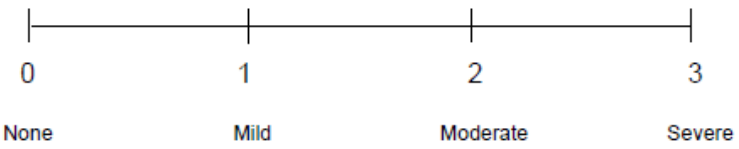
An addendum to this SAP will be developed that documents all additional data handling techniques employed to manage COVID impacted data, if needed, prior to database lock and unblinding.

13.2. Physician's Global Assessment

Subject ID:__ __ __ __ __ __ __ Patient Initials:__ __ __ Date: __ __-__ __ __-__ __ __ __ Visit: __ __
DD MON YEAR

PHYSICIAN'S GLOBAL ASSESSMENT (PGA)

How do you assess your subject's current disease activity?



Result: _____mm

Note: Please review the scale from previous visit, as applicable, and mark a vertical line on the scale above to assess the overall status of the subject's SLE signs and symptoms, and the functional capacity. Zero (0) = very good, asymptomatic and no limitation of normal activity; Three (3) = most severe possible disease ever seen in all SLE patients.

Investigator Signature	Assessor Initials	Date

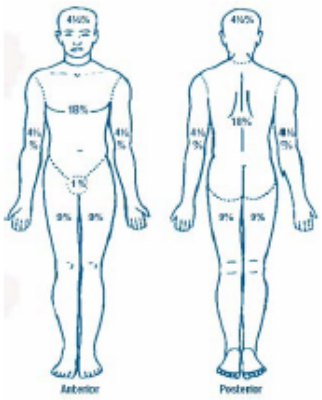
13.3. SLEDAI-2K Index

Subject Number:___ ___ ___ -___ ___ ___ ___	Protocol BOS161721-02
Screening / Randomization / _____ Week Visit & Date of Assessment: __ __ - __ __ __ - 20__ __ (Circle visit type &/or enter visit week) (Day Month Year)	

SLEDAI 2K Index Must be present at the time of visit or in the preceding 30 days (or 10 days for V1 and V2) and SLE-Related to be checked!			
Wt	(√) if Present	Descriptor	Definitions If descriptor is checked as present, please check appropriate condition(s) in definition and/or specify in the space provided.
8	☐	Seizure	Recent onset (last 28 days). Exclude metabolic, infectious or drug causes. **Specify if checked:**
8	☐	Psychosis	Altered ability to function in normal activity due to severe disturbance in the perception of reality. Include: □ hallucinations, □ incoherence, □ marked loose associations, □ impoverished thought content, □ marked illogical thinking, and □ bizarre, disorganized, or catatonic behavior. Exclude uremia and drug causes.
8	☐	Organic Brain Syndrome	Altered mental function with □ impaired orientation, memory or other intellectual function (with rapid onset and fluctuating clinical features), □ inability to sustain attention to environment, and at least 2 of the following: □ perceptual disturbance, □ incoherent speech, □ insomnia or daytime drowsiness, □ increased or decreased psychomotor activity. Exclude metabolic, infectious or drug causes.
8	☐	Visual Disturbance	Retinal changes of SLE. Include □ cytoid bodies, □ retinal hemorrhages, □ serous exudate or hemorrhages in the choroid, and □ optic neuritis. Exclude hypertension, infection, or drug causes.
8	☐	Cranial Nerve Disorder	New onset of □ sensory or □ motor neuropathy involving cranial nerves.
8	☐	Lupus Headache	Severe, persistent headache. May be migrainous, but must be nonresponsive to narcotic analgesia. (Must have been severe enough to warrant lumbar puncture and/or MRI or head CT, perhaps hospitalization, and is thought to be due to active lupus cerebritis.) **Specify if checked:**
8	☐	CVA	New onset of cerebrovascular accident(s) (CVA). Exclude arteriosclerosis. **Specify if checked:**
8	☐	Vasculitis	□ Ulceration, □ gangrene, □ tender finger nodules, □ periungual infarction, □ splinter hemorrhages, or □ biopsy or □ angiogram proof of vasculitis.

Subject Number___ ___ ___-___ ___ ___ ___			Protocol BOS161721-02
Screening / Randomization / _____ Week Visit & Date of Assessment: ___ ___ - ___ ___ ___ - 20____ (Circle visit type &/or enter visit week) (Day Month Year)			
4	☐	Arthritis	≥ 2 joints with pain and signs of inflammation (□ tenderness, □ swelling or □ effusion). **(Complete Joint Count if checked.)**
4	☐	Myositis	Proximal muscle aching/weakness associated with □ elevated creatine phosphokinase (CK)/aldolase or □ EMG changes or □ a biopsy showing myositis.
4	☐	Urinary Casts	Heme-granular or red blood cell casts. (See labs)
4	☐	Hematuria	>5 RBC/high power field. Exclude stone, infection, or other cause. (See labs)
4	☐	Proteinuria	>0.5 gm/24-hour equivalent. (See labs)
4	☐	Pyuria	>5 WBC/high power field. Exclude infection. (See labs)
2	☐	Rash	Inflammatory-type rash. **Specify and complete CLASI if checked:**
2	☐	Alopecia	□ Abnormal, □ patchy or □ diffuse loss of hair. **(Complete CLASI if checked.)**
2	☐	Mucosal Ulcers	□ Oral or □ □nasal ulcerations. **(Complete CLASI if checked.)**
2	☐	Pleurisy	Pleuritic chest pain with □ pleural rub or □ effusion, or □ pleural thickening.
2	☐	Pericarditis	Pericardial pain with at least one (1) of the following: □ rub or □ effusion, or □ ECG or □ □echocardiogram confirmation.
2	☐	Low Complement	Decrease in CH50, C3 or C4 below the lower limit of normal for testing laboratory. (See labs)
2	☐	Increased DNA Binding	>25% binding by Farr assay or above normal range for testing laboratory. (See labs)
1	☐	Fever	>38°C (100.4F). Exclude infectious cause. **Specify if checked:**
1	☐	Thrombocytopenia	<100,000 platelets/mm^3 (See labs)
1	☐	Leukopenia	<3,000 white blood cells/mm^3. Exclude drug causes. (See labs)
	TOTAL SCORE (Sum of all of weights next to descriptors marked present)		
Investigator Signature			Date

13.4. BILAG 2004 Index

BILAG 2004 Index Source Document Worksheet		
Subject Number:___ ___ ___ - ___ ___ ___ ___		**Protocol BOS161721-02**
Screening / Randomization / _____ Week Visit & Date of Assessment: ___ ___ - ___ ___ ___ - 20___ (Circle visit type &/or enter visit week) (Day Month Year)		
Only record items due to SLE disease activity. Each assessment refers to manifestation occurring in the last 4 weeks (compared with the previous visit). Scoring: 0 = Not Present, 1 = Improving, 2 = Same, 3 = Worse, 4 = New or Recurrence **There must be detailed documentation for all descriptors scored 1 – 4**		
CONSTITUTIONAL	Score	**Detail findings/changes for all descriptors scored 1 - 4**
1. Pyrexia – documented > 37.5° C		
2. Weight loss – unintentional > 5%		
3. Lymphandenopathy / splenomegaly		
4. Anorexia		
MUCOCUTANEOUS - **If any of these features are scored 1-4, please complete CLASI!**	Score	**Draw / specify area(s) and note approximate BSA (use also for CLASI documentation)**
5. Skin eruption – severe (Must involve >18% BSA)		
6. Skin eruption – mild (≤18% BSA)		
7. Angio–oedema – severe		
8. Angio–oedema - mild		
9. Mucosal ulceration – severe		
10. Mucosal ulceration – mild		
11. Panniculitis / Bullous lupus – severe (> 9% BSA)		
12. Panniculitis / Bullous lupus – mild (≤ 9% BSA)		
13. Major cutaneous vasculitis / thrombosis		
14. Digital infracts or nodular vasculitis		
15. Alopecia – severe		
16. Alopecia – mild		
17. Peri-ungual erythema / chilblains		
18. Splinter hemorrhages		
NEUROPSYCHIATRIC	Score	**Detail findings/changes for all descriptors scored 1 - 4**
19. Aseptic meningitis		
20. Cerebral vasculitis		
21. Demyelinating syndrome		
22. Myelopathy		
23. Acute confusional state		
24. Psychosis		
25. Acute inflammatory demyelinating polyradiculoneuropathy		
26. Mononeuropathy (single / multiplex)		
27. Cranial neuropathy		
28. Plexopathy		
29. Polyneuropathy		
30. Seizure disorder		
31. Status epilepticus		
32. Cerebrovascular disease (not due to vasculitis)		
33. Cognitive dysfunction		
34. Movement disorder		
35. Autonomic disorder		
36. Cerebellar ataxia (isolated)		
37. Lupus headache – severe unremitting		
38. Headache from IC hypertension		
MUSCULOSKELETAL	Score	**Detail findings/changes for all descriptors scored 1 - 4**
39. Myositis – severe		
40. Myositis – mild		
41. Arthritis – severe		
42. Arthritis – moderate / Tendonitis / Tenosynovitis		
43. Arthritis – mild / Arthralgia / Myalgia		

Subject Number: ___ ___ ___ -___ ___ ___ ___	Protocol BOS161721-02
Screening / Randomization / _____ Week Visit (Circle visit type &/or enter visit week) & Date of Assessment: __ __ - __ __ __ - 20___ (Day Month Year)	

CARDIORESPIRATORY	Score	Detail findings/changes for all descriptors scored 1 - 4
44. Myocarditis – mild		
45. Myocarditis / Endocarditis + Cardiac failure		
46. Arrhythmia		
47. New valvular dysfunction		
48. Pleurisy / Pericarditis		
49. Cardiac tamponade		
50. Pleural effusion with dyspnoea		
51. Pulmonary haemorrhage / vasculitis		
52. Interstitial alveolitis / pneumonitis		
53. Shrinking lung syndrome		
54. Aortitis		
55. Coronary vasculitis		
GASTROINTESTINAL	Score	Detail findings/changes for all descriptors scored 1 - 4
56. Lupus peritonitis		
57. Abdominal serositis or ascites		
58. Lupus enteritis / colitis		
59. Malabsorption		
60. Protein-losing enteropathy		
61. Intestinal pseudo-obstruction		
62. Lupus hepatitis		
63. Acute lupus cholecystitis		
64. Acute lupus pancreatitis		
OPHTHALMIC	Score	Detail findings/changes for all descriptors scored 1 - 4
65. Orbital inflammation / proptosis		
66. Keratitis – severe		
67. Keratitis - mild		
68. Anterior uveitis		
69. Posterior uveitis / retinal vasculitis – severe		
70. Posterior uveitis / retinal vasculitis – mild		
71. Episcleritis		
72. Scleritis – severe		
73. Scleritis – mild		
74. Retinal / choroidal vaso-occlusive disease		
75. Isolated cotton wool-spots (cytoid bodies)		
76. Optic neuritis		
77. Anterior ischaemic optic neuropathy		

RENAL	Value	SLE related?	HAEMATOLOGICAL	Value	SLE related?
78. Systolic BP > 140mmHg		Y / N	90. Haemoglobin (g/dl)		Y / N
79. Diastolic BP > 90 mmHg		Y / N	91. Total white cell count		Y / N
80. Accelerated hypertension	Yes / No		92. Neutrophils		Y / N
81. Urine dipstick (protein) (+=1, ++=2, +++=3)		Y / N	93. Lymphocytes		Y / N
82. Urine albumin-creatinine ratio (mg/mmol)	Not assessed		94. Platelets		Y / N
83. Urine protein-creatinine ratio (mg/mmol)		Y / N	95. TTP	0 1 2 3 4	
84. 24 hour urine protein		Y / N	96. Evidence of active haemolysis	Yes / No	
85. Nephrotic syndrome	Yes / No		97. Coombs' test positive (isolated)	Yes / No	
86. Creatinine (plasma/serum) (µmol/l)		Y / N	Specify lab findings / additional comments:		
87. GFR		Y / N			
88. Active urinary sediment	Yes / No				
89. Active nephritis	Yes / No				
Investigator Signature	Date				

13.5. ACR-28 Joint Count Assessment

Site/Subject #: ___-____	Protocol BOS161721-02
Visit _____	Date of Assessment: ____/_____/_____ (DD / MMM / YYYY)

28 JOINT COUNT ASSESSMENT

Please circle all that apply.
Query the subject at the start of the joint count about pain (prior to assessment of tenderness and swelling) as to whether they have experienced or are experiencing pain in any of the 28 joints. Artificial joints and joints that have ever had surgery performed on them should be scored as "NA."

	Joint #	Right Side						Left Side					
		Tenderness			Swelling			Tenderness			Swelling		
Upper Extremity													
Shoulder	1	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
Elbow	2	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
Wrist	3	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
MCP I	4	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
MCP II	5	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
MCP III	6	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
MCP IV	7	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
MCP V	8	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
IP	9	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
PIP II	10	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
PIP III	11	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
PIP IV	12	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
PIP V	13	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
Lower Extremity													
Knee	14	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA

Are there any other joints (not listed above) that are significantly affected by SLE? No Yes
If Yes, list and describe: _____

_____

Based on clinical exam/impression and patient report for the last 4 weeks*:
Is there loss of functional range of motion sufficient to interfere with instrumental activities of daily living (eg, household chores, preparing meals, working, etc.)? No Yes

Is there significant impairment of basic activities of daily living (ADLs) so that the patient requires assistance or an assistive device due to the active inflammation? No Yes

If Yes, activities affected:
Grooming Dressing Ambulating Toileting Feeding oneself

_____

INVESTIGATOR SIGNATURE: _____ **DATE:** ____ / _____ / _____
DD MMM YYYY

13.6. SLICC/ACR Damage Index

Subject Number: ___ ___ ___-___ ___ ___ ___	Protocol BOS161721-02
SLICC / ACR Damage Index	
Date of Assessment: ____ - ____ - 20____ (Day Month Year)	**Please circle Study Visit: Day 1 or Day 180**
Item	**Score**
Ocular (either eye, by clinical assessment)	
Any cataract ever	☐ 1
Retinal change or optic atrophy	☐ 1
Neuropsychiatric	
Cognitive Impairment (eg memory deficit, difficulty with calculation, poor concentration, difficulty in spoken or written language, impaired performance levels) or major psychosis	☐ 1
Seizures requiring therapy for 6 months	☐ 1
Cerebrovascular accident ever (score 2 if > 1)	☐ 1 ☐ 2
Cranial or peripheral neuropathy (excluding optic)	☐ 1
Transverse myelitis	☐ 1
Renal	
Estimated or measured glomerular filtration rate < 50%	☐ 1
Proteinuria > 3.5gm/24 hours	☐ 1
OR	
End-stage renal disease (regardless of dialysis or transplantation)	☐ 3
Pulmonary	
Pulmonary hypertension (right ventricular prominence, or loud P2)	☐ 1
Pulmonary fibrosis (physical and radiographic)	☐ 1
Shrinking lung	☐ 1
Pleural fibrosis (radiographic)	☐ 1
Pulmonary Infarction (radiographic)	☐ 1
Cardiovascular	
Angina or coronary artery bypass	☐ 1
Myocardial Infarction ever (score 2 if > 1)	☐ 1 ☐ 2
Cardiomyopathy (ventricular dysfunction)	☐ 1
Valvular disease (diastolic, murmur, or systolic murmur > 3/6)	☐ 1
Pericarditis for 6 months, or pericardiectomy	☐ 1
Peripheral Vascular	
Claudication for 6 months	☐ 1
Minor tissue loss (pulp space)	☐ 1
Significant tissue loss ever (eg loss of digit or limb) (score 2 if > 1 site)	☐ 1 ☐ 2
Venous thrombosis with swelling, ulceration, or venous stasis	☐ 1
Gastrointestinal	
Infarction or resection or bowel below duodenum, spleen, liver, or gall bladder ever, for any cause (score 2 if > 1 site)	☐ 1 ☐ 2
Mesenteric insufficiency	☐ 1
Chronic peritonitis	☐ 1
Stricture or upper gastrointestinal tract surgery ever	☐ 1
Musculoskeletal	
Muscle atrophy or weakness	☐ 1
Deforming or erosive arthritis (including reducible deformities, excluding avascular necrosis	☐ 1
Osteoporosis with fracture or vertebral collapse (excluding avascular necrosis)	☐ 1
Avascular necrosis (score 2 is > 1)	☐ 1 ☐ 2
Osteomyelitis	☐ 1
Skin	
Scarring chronic alopecia	☐ 1
Extensive scarring or panniculum other than scalp and pulp space	☐ 1
Skin ulceration (excluding thrombosis) for > 6 months	☐ 1
Premature gonadal failure	☐ 1
Diabetes (regardless of treatment)	☐ 1
Malignancy (excluding dysplasia) (score 2 if > 1 site)	☐ 1 ☐ 2

13.7. CLASI

BOS161721-02 Subject: _ _ _ _ _ _ _ Visit: _____ Date: _ _-_ _ _-20_ _
Day Month Year

Cutaneous LE Disease Area and Severity Index (CLASI)

Select the score in each anatomical location that describes the most severely affected cutaneous lupus-associated lesion

	activity		damage		
Anatomical Location	Erythema	Scale/ Hypertrophy	Dyspigmentation	Scarring/ Atrophy/ Panniculitis	Anatomical Location
	0-absent 1-pink; faint erythema 2- red; 3-dark red; purple/violaceous/ crusted/ hemorrhagic	0-absent; 1-scale 2-verrucous/ hypertrophic	0-absent, 1-dyspigmentaton	0 – absent 1 – scarring 2 – severely atrophic scarring or panniculitis	
Scalp				See below	Scalp
Ears					Ears
Nose (incl. malar area)					Nose (incl. malar area)
Rest of the face					Rest of the face
V-area neck (frontal)					V-area neck (frontal)
Post. Neck &/or shoulders					Post. Neck &/or shoulders
Chest					Chest
Abdomen					Abdomen
Back, buttocks					Back, buttocks
Arms					Arms
Hands					Hands
Legs					Legs
Feet					Feet

Mucous membrane

Mucous membrane lesions (examine if patient confirms involvement)	
0-absent; 1-lesion or ulceration	

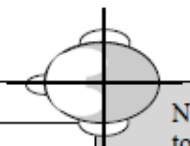
Dyspigmentation

Report duration of dyspigmentation after active lesions have resolved (verbal report by patient – tick appropriate box)

☐ Dyspigmentation usually lasts less than 12 months (dyspigmentation score above remains)

☐ Dyspigmentation usually lasts at least 12 months (dyspigmentation score is doubled)

Alopecia



Recent Hair loss (within the last 30 days / as reported by patient)	
1-Yes 0-No	

NB: if scarring and non-scarring aspects seem to coexist in one lesion, please score both

Divide the scalp into four quadrants as shown. The dividing line between right and left is the midline. The dividing line between frontal and occipital is the line connecting the highest points of the ear lobe. A quadrant is considered affected if there is a lesion within the quadrant.

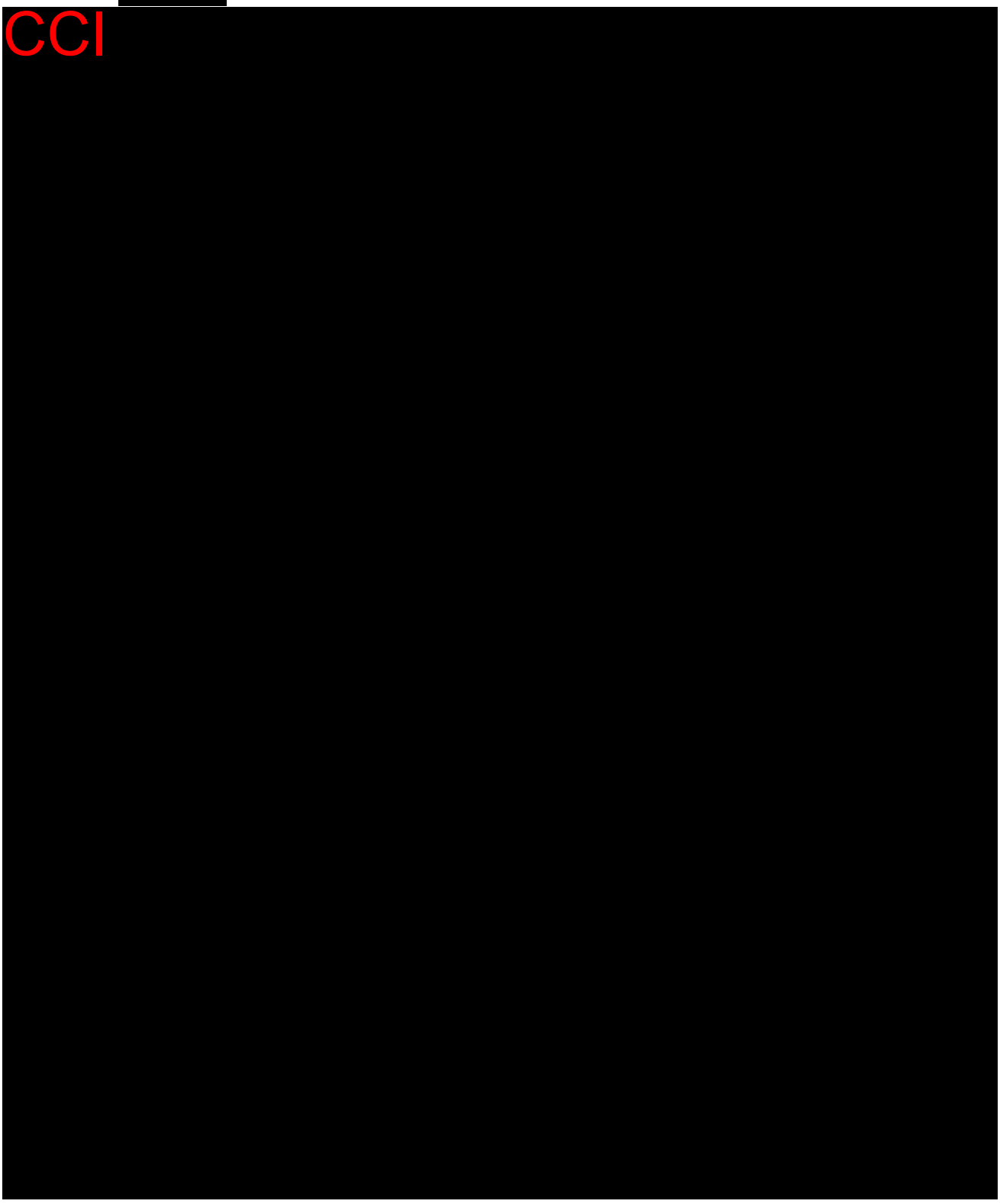
Alopecia (clinically not obviously scarred)		Scarring of the scalp (judged clinically)	
0-absent 1-diffuse; non-inflammatory 2-focal or patchy in one quadrant; 3-focal or patchy in more than one quadrant		0- absent 3- in one quadrant 4- two quadrants 5- three quadrants 6- affects the whole skull	

Total Activity Score
(For the activity score please add up the scores of the left side i.e. for Erythema, Scale/Hypertrophy, Mucous membrane involvement and Alopecia)

Total Damage Score
(For the damage score, please add up the scores of the right side, i.e. for Dyspigmentation, Scarring/Atrophy/Panniculitis and Scarring of the Scalp)

13.8. CCI

CCI



CCI

CCI

CCI

CCI

CCI

[redacted]

CCI

[redacted]

CCI

CCI

CCI

13.9. C-SSRS

13.9.1. C-SSRS (Baseline/Screening)

SUICIDAL IDEATION		
Ask questions 1 and 2. If both are negative, proceed to "Suicidal Behavior" section. If the answer to question 2 is "yes", ask questions 3, 4 and 5. If the answer to question 1 and/or 2 is "yes", complete "Intensity of Ideation" section below.	**Lifetime: Time He/She Felt Most Suicidal**	**Past 12 Months**
1. Wish to be Dead Subject endorses thoughts about a wish to be dead or not alive anymore, or wish to fall asleep and not wake up. ***Have you wished you were dead or wished you could go to sleep and not wake up?*** If yes, describe:	**Yes** ☐ **No** ☐	**Yes** ☐ **No** ☐
2. Non-Specific Active Suicidal Thoughts General non-specific thoughts of wanting to end one's life/commit suicide (e.g., *"I've thought about killing myself"*) without thoughts of ways to kill oneself/associated methods, intent, or plan during the assessment period. ***Have you actually had any thoughts of killing yourself?*** If yes, describe:	**Yes** ☐ **No** ☐	**Yes** ☐ **No** ☐
3. Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act Subject endorses thoughts of suicide and has thought of at least one method during the assessment period. This is different than a specific plan with time, place or method details worked out (e.g., thought of method to kill self but not a specific plan). Includes person who would say, *"I thought about taking an overdose but I never made a specific plan as to when, where or how I would actually do it... and I would never go through with it."* ***Have you been thinking about how you might do this?*** If yes, describe:	**Yes** ☐ **No** ☐	**Yes** ☐ **No** ☐
4. Active Suicidal Ideation with Some Intent to Act, without Specific Plan Active suicidal thoughts of killing oneself and subject reports having some intent to act on such thoughts, as opposed to *"I have the thoughts but I definitely will not do anything about them."* ***Have you had these thoughts and had some intention of acting on them?*** If yes, describe:	**Yes** ☐ **No** ☐	**Yes** ☐ **No** ☐
5. Active Suicidal Ideation with Specific Plan and Intent Thoughts of killing oneself with details of plan fully or partially worked out and subject has some intent to carry it out. ***Have you started to work out or worked out the details of how to kill yourself? Do you intend to carry out this plan?*** If yes, describe:	**Yes** ☐ **No** ☐	**Yes** ☐ **No** ☐
INTENSITY OF IDEATION		
The following features should be rated with respect to the most severe type of ideation (i.e.,1-5 from above, with 1 being the least severe and 5 being the most severe). Ask about time he/she was feeling the most suicidal. Lifetime - ***Most Severe Ideation:*** _____ (*Type # (1-5)*) _____ (*Description of Ideation*) Past 12 Months - ***Most Severe Ideation:*** _____ (*Type # (1-5)*) _____ (*Description of Ideation*)	Most Severe	Most Severe
Frequency ***How many times have you had these thoughts?*** (1) Less than once a week (2) Once a week (3) 2-5 times in week (4) Daily or almost daily (5) Many times each day	___	___
Duration ***When you have the thoughts how long do they last?*** (1) Fleeting - few seconds or minutes (2) Less than 1 hour/some of the time (3) 1-4 hours/a lot of time (4) 4-8 hours/most of day (5) More than 8 hours/persistent or continuous	___	___
Controllability ***Could/can you stop thinking about killing yourself or wanting to die if you want to?*** (1) Easily able to control thoughts (2) Can control thoughts with little difficulty (3) Can control thoughts with some difficulty (4) Can control thoughts with a lot of difficulty (5) Unable to control thoughts (0) Does not attempt to control thoughts	___	___
Deterrents ***Are there things - anyone or anything (e.g., family, religion, pain of death) - that stopped you from wanting to die or acting on thoughts of committing suicide?*** (1) Deterrents definitely stopped you from attempting suicide (2) Deterrents probably stopped you (3) Uncertain that deterrents stopped you (4) Deterrents most likely did not stop you (5) Deterrents definitely did not stop you (0) Does not apply	___	___
Reasons for Ideation ***What sort of reasons did you have for thinking about wanting to die or killing yourself? Was it to end the pain or stop the way you were feeling (in other words you couldn't go on living with this pain or how you were feeling) or was it to get attention, revenge or a reaction from others? Or both?*** (1) Completely to get attention, revenge or a reaction from others (2) Mostly to get attention, revenge or a reaction from others (3) Equally to get attention, revenge or a reaction from others and to end/stop the pain (4) Mostly to end or stop the pain (you couldn't go on living with the pain or how you were feeling) (5) Completely to end or stop the pain (you couldn't go on living with the pain or how you were feeling) (0) Does not apply	___	___

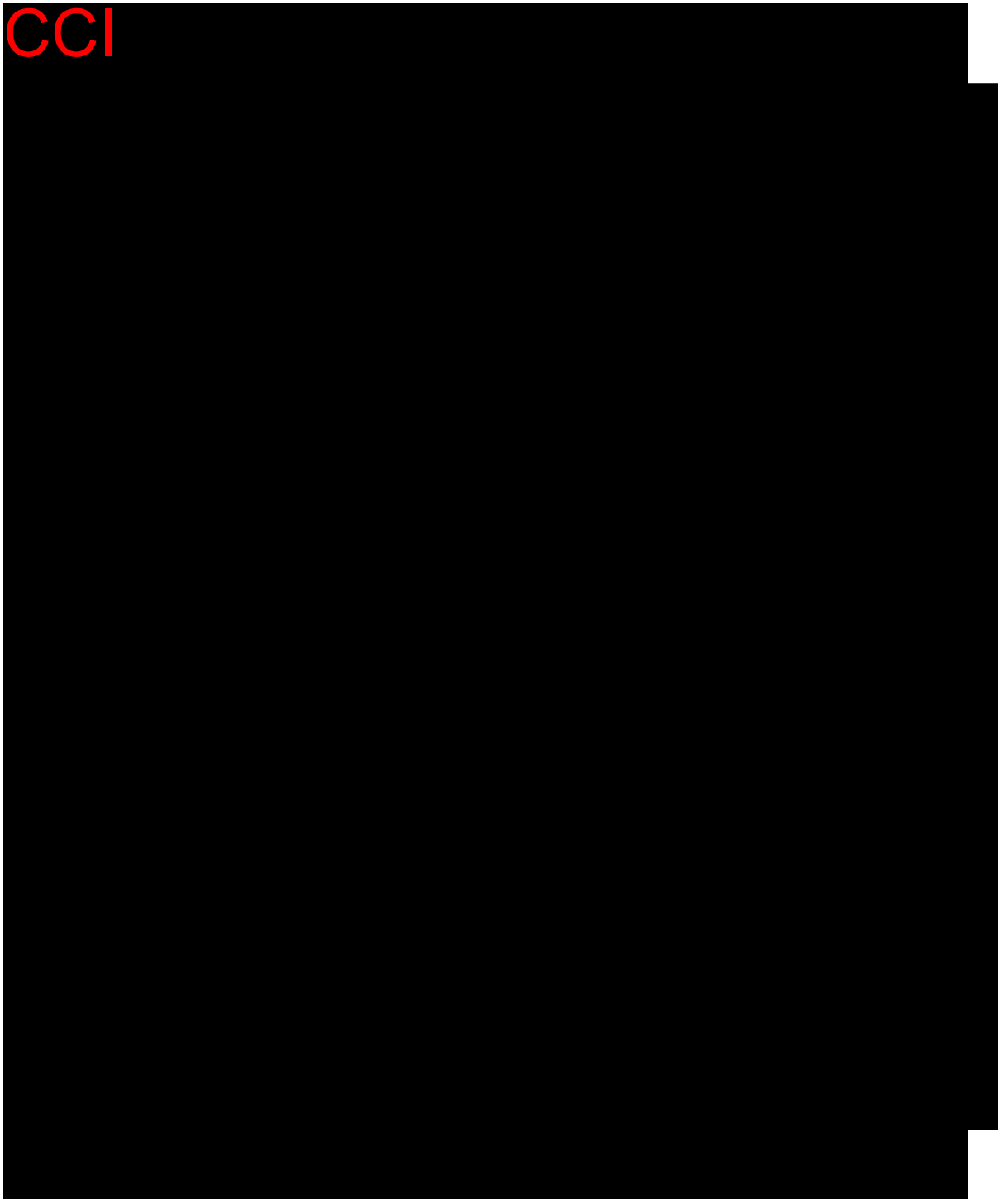
SUICIDAL BEHAVIOR *(Check all that apply, so long as these are separate events; must ask about all types)*	**Lifetime**	**Past 1 Year**
Actual Attempt: A potentially self-injurious act committed with at least some wish to die, *as a result of act*. Behavior was in part thought of as method to kill oneself. Intent does not have to be 100%. If there is ***any*** intent/desire to die associated with the act, then it can be considered an actual suicide attempt. ***There does not have to be any injury or harm***, just the potential for injury or harm. If person pulls trigger while gun is in mouth but gun is broken so no injury results, this is considered an attempt. Inferring Intent: Even if an individual denies intent/wish to die, it may be inferred clinically from the behavior or circumstances. For example, a highly lethal act that is clearly not an accident so no other intent but suicide can be inferred (e.g., gunshot to head, jumping from window of a high floor/story). Also, if someone denies intent to die, but they thought that what they did could be lethal, intent may be inferred. ***Have you made a suicide attempt?*** ***Have you done anything to harm yourself?*** ***Have you done anything dangerous where you could have died?*** ***What did you do?*** ***Did you_____as a way to end your life?*** ***Did you want to die (even a little) when you_____?*** ***Were you trying to end your life when you_____?*** ***Or did you think it was possible you could have died from_____?*** ***Or did you do it purely for other reasons / without ANY intention of killing yourself (like to relieve stress, feel better, get sympathy, or get something else to happen)?*** (Self-Injurious Behavior without suicidal intent) If yes, describe:	Yes ☐ No ☐ Total # of Attempts _____	Yes ☐ No ☐ Total # of Attempts _____
Has subject engaged in Non-Suicidal Self-Injurious Behavior?	Yes ☐ No ☐	Yes ☐ No ☐
Interrupted Attempt: When the person is interrupted (by an outside circumstance) from starting the potentially self-injurious act *(if not for that, actual attempt would have occurred)*. Overdose: Person has pills in hand but is stopped from ingesting. Once they ingest any pills, this becomes an attempt rather than an interrupted attempt. Shooting: Person has gun pointed toward self, gun is taken away by someone else, or is somehow prevented from pulling trigger. Once they pull the trigger, even if the gun fails to fire, it is an attempt. Jumping: Person is poised to jump, is grabbed and taken down from ledge. Hanging: Person has noose around neck but has not yet started to hang - is stopped from doing so. ***Has there been a time when you started to do something to end your life but someone or something stopped you before you actually did anything?*** If yes, describe:	Yes ☐ No ☐ Total # of interrupted _____	Yes ☐ No ☐ Total # of interrupted _____
Aborted Attempt: When person begins to take steps toward making a suicide attempt, but stops themselves before they actually have engaged in any self-destructive behavior. Examples are similar to interrupted attempts, except that the individual stops him/herself, instead of being stopped by something else. ***Has there been a time when you started to do something to try to end your life but you stopped yourself before you actually did anything?*** If yes, describe:	Yes ☐ No ☐ Total # of aborted _____	Yes ☐ No ☐ Total # of aborted _____
Preparatory Acts or Behavior: Acts or preparation towards imminently making a suicide attempt. This can include anything beyond a verbalization or thought, such as assembling a specific method (e.g., buying pills, purchasing a gun) or preparing for one's death by suicide (e.g., giving things away, writing a suicide note). ***Have you taken any steps towards making a suicide attempt or preparing to kill yourself (such as collecting pills, getting a gun, giving valuables away or writing a suicide note)?*** If yes, describe:	Yes ☐ No ☐	Yes ☐ No ☐
Suicidal Behavior: Suicidal behavior was present during the assessment period?	Yes ☐ No ☐	Yes ☐ No ☐

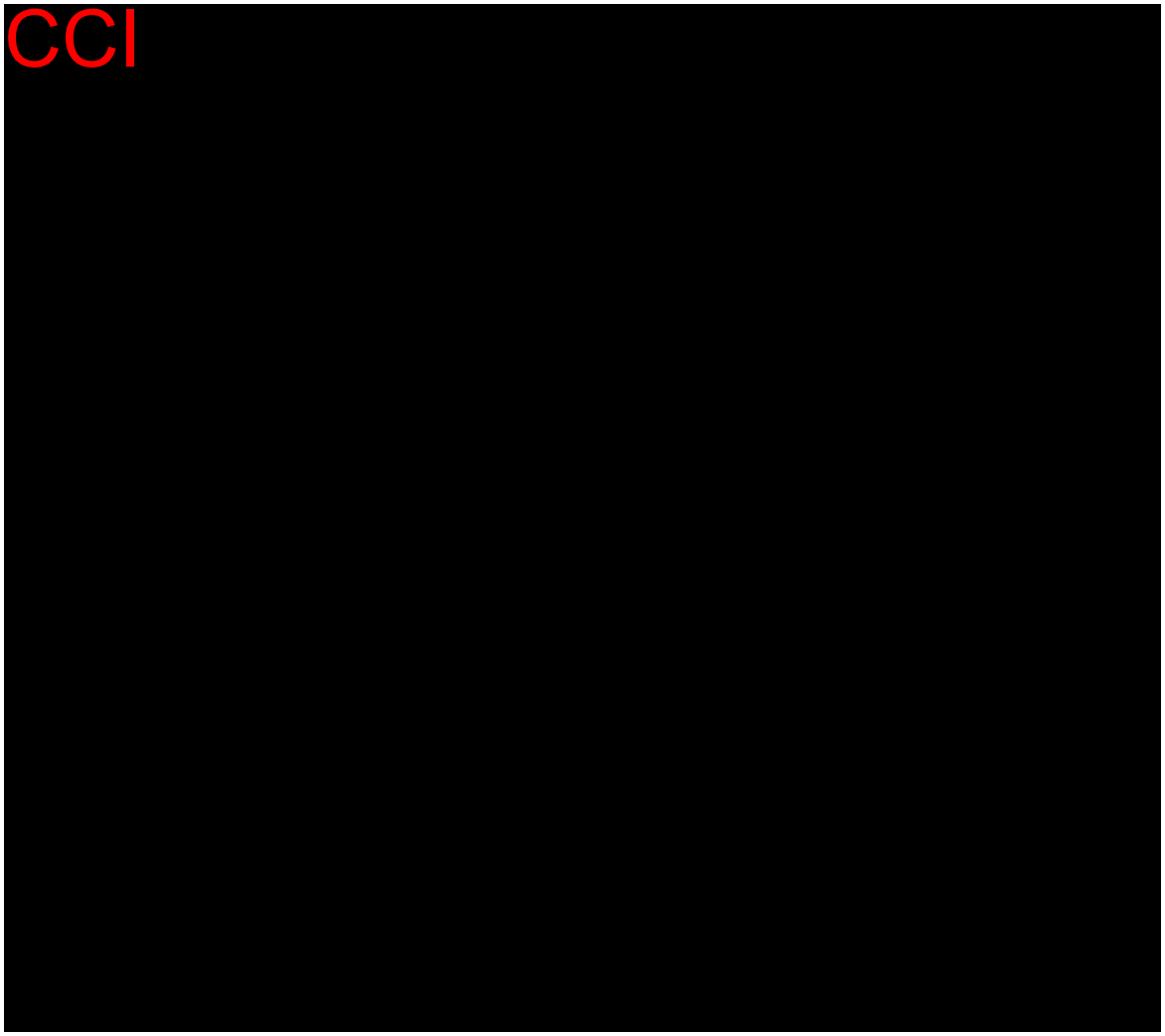
Answer for Actual Attempts Only	Most Recent Attempt Date:	Most Lethal Attempt Date:	Initial/First Attempt Date:
Actual Lethality/Medical Damage: 0. No physical damage or very minor physical damage (e.g., surface scratches). 1. Minor physical damage (e.g., lethargic speech; first-degree burns; mild bleeding; sprains). 2. Moderate physical damage; medical attention needed (e.g., conscious but sleepy, somewhat responsive; second-degree burns; bleeding of major vessel). 3. Moderately severe physical damage; *medical* hospitalization and likely intensive care required (e.g., comatose with reflexes intact; third-degree burns less than 20% of body; extensive blood loss but can recover; major fractures). 4. Severe physical damage; *medical* hospitalization with intensive care required (e.g., comatose without reflexes; third-degree burns over 20% of body; extensive blood loss with unstable vital signs; major damage to a vital area). 5. Death	*Enter Code* _____	*Enter Code* _____	*Enter Code* _____
Potential Lethality: Only Answer if Actual Lethality=0 Likely lethality of actual attempt if no medical damage (the following examples, while having no actual medical damage, had potential for very serious lethality: put gun in mouth and pulled the trigger but gun fails to fire so no medical damage; laying on train tracks with oncoming train but pulled away before run over). 0 = Behavior not likely to result in injury 1 = Behavior likely to result in injury but not likely to cause death 2 = Behavior likely to result in death despite available medical care	*Enter Code* _____	*Enter Code* _____	*Enter Code* _____

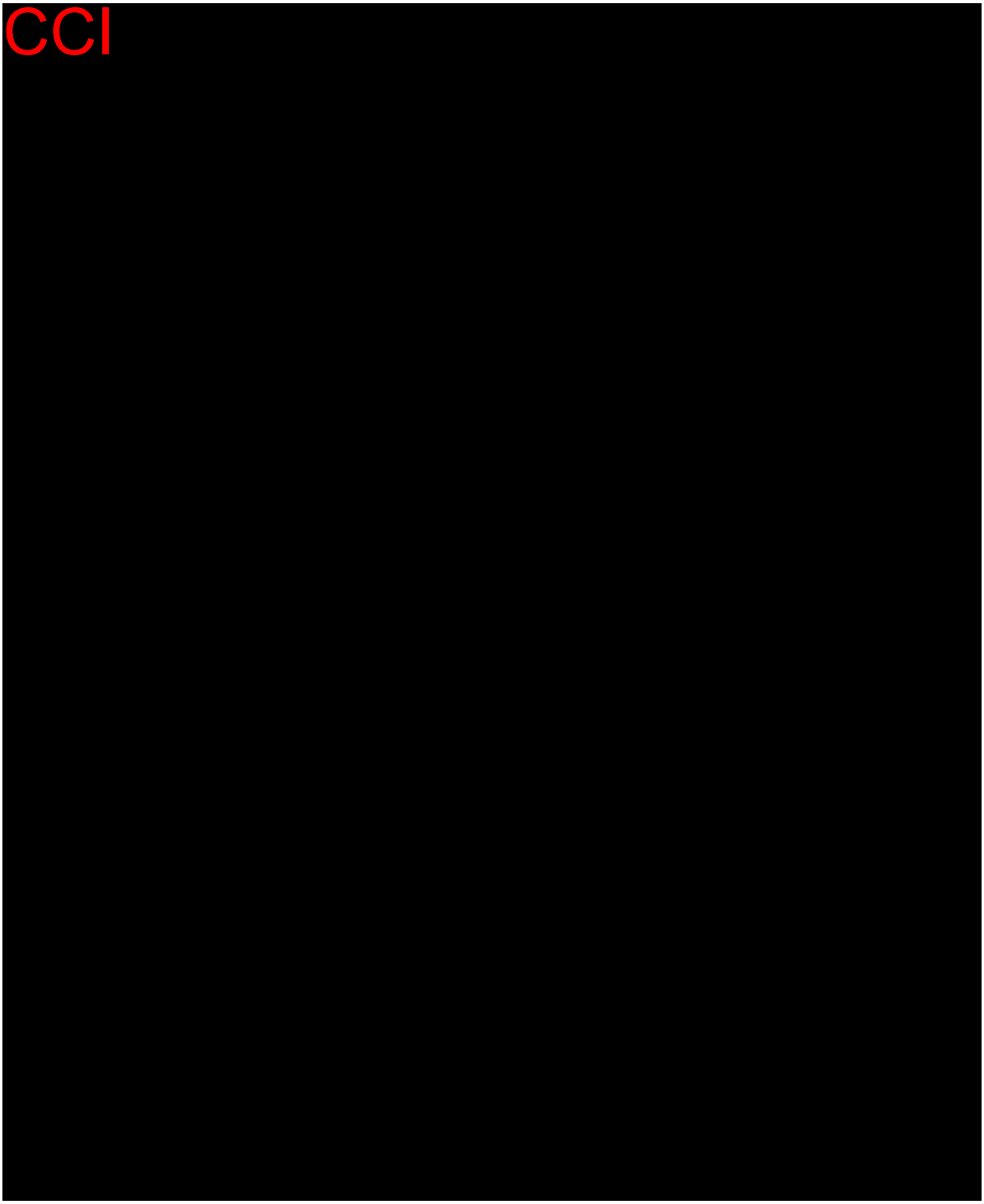
13.9.2. C-SSRS (Since Last Visit)

SUICIDAL IDEATION	
Ask questions 1 and 2. If both are negative, proceed to "Suicidal Behavior" section. If the answer to question 2 is "yes", ask questions 3, 4 and 5. If the answer to question 1 and/or 2 is "yes", complete "Intensity of Ideation" section below.	**Since Last Visit**
1. Wish to be Dead Subject endorses thoughts about a wish to be dead or not alive anymore, or wish to fall asleep and not wake up. ***Have you wished you were dead or wished you could go to sleep and not wake up?*** If yes, describe:	**Yes** ☐ **No** ☐
2. Non-Specific Active Suicidal Thoughts General non-specific thoughts of wanting to end one's life/commit suicide (e.g., *"I've thought about killing myself"*) without thoughts of ways to kill oneself/associated methods, intent, or plan during the assessment period. ***Have you actually had any thoughts of killing yourself?*** If yes, describe:	**Yes** ☐ **No** ☐
3. Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act Subject endorses thoughts of suicide and has thought of at least one method during the assessment period. This is different than a specific plan with time, place or method details worked out (e.g., thought of method to kill self but not a specific plan). Includes person who would say, *"I thought about taking an overdose but I never made a specific plan as to when, where or how I would actually do it.....and I would never go through with it".* ***Have you been thinking about how you might do this?*** If yes, describe:	**Yes** ☐ **No** ☐
4. Active Suicidal Ideation with Some Intent to Act, without Specific Plan Active suicidal thoughts of killing oneself and subject reports having some intent to act on such thoughts, as opposed to *"I have the thoughts but I definitely will not do anything about them".* ***Have you had these thoughts and had some intention of acting on them?*** If yes, describe:	**Yes** ☐ **No** ☐
5. Active Suicidal Ideation with Specific Plan and Intent Thoughts of killing oneself with details of plan fully or partially worked out and subject has some intent to carry it out. ***Have you started to work out or worked out the details of how to kill yourself? Do you intend to carry out this plan?*** If yes, describe:	**Yes** ☐ **No** ☐
INTENSITY OF IDEATION	
The following features should be rated with respect to the most severe type of ideation (i.e.,1-5 from above, with 1 being the least severe and 5 being the most severe). ***Most Severe Ideation:*** _____ _____ ***Type # (1-5)*** ***Description of Ideation***	Most Severe
Frequency ***How many times have you had these thoughts?*** (1) Less than once a week (2) Once a week (3) 2-5 times in week (4) Daily or almost daily (5) Many times each day	___
Duration ***When you have the thoughts how long do they last?*** (1) Fleeting - few seconds or minutes (2) Less than 1 hour/some of the time (3) 1-4 hours/a lot of time (4) 4-8 hours/most of day (5) More than 8 hours/persistent or continuous	___
Controllability ***Could/can you stop thinking about killing yourself or wanting to die if you want to?*** (1) Easily able to control thoughts (2) Can control thoughts with little difficulty (3) Can control thoughts with some difficulty (4) Can control thoughts with a lot of difficulty (5) Unable to control thoughts (0) Does not attempt to control thoughts	___
Deterrents ***Are there things - anyone or anything (e.g., family, religion, pain of death) - that stopped you from wanting to die or acting on thoughts of committing suicide?*** (1) Deterrents definitely stopped you from attempting suicide (2) Deterrents probably stopped you (3) Uncertain that deterrents stopped you (4) Deterrents most likely did not stop you (5) Deterrents definitely did not stop you (0) Does not apply	___
Reasons for Ideation ***What sort of reasons did you have for thinking about wanting to die or killing yourself? Was it to end the pain or stop the way you were feeling (in other words you couldn't go on living with this pain or how you were feeling) or was it to get attention, revenge or a reaction from others? Or both?*** (1) Completely to get attention, revenge or a reaction from others (2) Mostly to get attention, revenge or a reaction from others (3) Equally to get attention, revenge or a reaction from others and to end/stop the pain (4) Mostly to end or stop the pain (you couldn't go on living with the pain or how you were feeling) (5) Completely to end or stop the pain (you couldn't go on living with the pain or how you were feeling) (0) Does not apply	___

SUICIDAL BEHAVIOR *(Check all that apply, so long as these are separate events; must ask about all types)*	**Since Last Visit**
Actual Attempt: A potentially self-injurious act committed with at least some wish to die, *as a result of act*. Behavior was in part thought of as method to kill oneself. Intent does not have to be 100%. If there is ***any*** intent/desire to die associated with the act, then it can be considered an actual suicide attempt. ***There does not have to be any injury or harm***, just the potential for injury or harm. If person pulls trigger while gun is in mouth but gun is broken so no injury results, this is considered an attempt. Inferring Intent: Even if an individual denies intent/wish to die, it may be inferred clinically from the behavior or circumstances. For example, a highly lethal act that is clearly not an accident so no other intent but suicide can be inferred (e.g., gunshot to head, jumping from window of a high floor/story). Also, if someone denies intent to die, but they thought that what they did could be lethal, intent may be inferred. ***Have you made a suicide attempt?*** ***Have you done anything to harm yourself?*** ***Have you done anything dangerous where you could have died?*** ***What did you do?*** ***Did you_____as a way to end your life?*** ***Did you want to die (even a little) when you_____?*** ***Were you trying to end your life when you_____?*** ***Or Did you think it was possible you could have died from_____?*** ***Or did you do it purely for other reasons / without ANY intention of killing yourself (like to relieve stress, feel better, get sympathy, or get something else to happen)?*** (Self-Injurious Behavior without suicidal intent) If yes, describe: **Has subject engaged in Non-Suicidal Self-Injurious Behavior?**	**Yes** ☐ **No** ☐ Total # of Attempts _____ **Yes** ☐ **No** ☐
Interrupted Attempt: When the person is interrupted (by an outside circumstance) from starting the potentially self-injurious act *(if not for that, actual attempt would have occurred)*. Overdose: Person has pills in hand but is stopped from ingesting. Once they ingest any pills, this becomes an attempt rather than an interrupted attempt. Shooting: Person has gun pointed toward self, gun is taken away by someone else, or is somehow prevented from pulling trigger. Once they pull the trigger, even if the gun fails to fire, it is an attempt. Jumping: Person is poised to jump, is grabbed and taken down from ledge. Hanging: Person has noose around neck but has not yet started to hang - is stopped from doing so. ***Has there been a time when you started to do something to end your life but someone or something stopped you before you actually did anything?*** If yes, describe:	**Yes** ☐ **No** ☐ Total # of interrupted _____
Aborted Attempt: When person begins to take steps toward making a suicide attempt, but stops themselves before they actually have engaged in any self-destructive behavior. Examples are similar to interrupted attempts, except that the individual stops him/herself, instead of being stopped by something else. ***Has there been a time when you started to do something to try to end your life but you stopped yourself before you actually did anything?*** If yes, describe:	**Yes** ☐ **No** ☐ Total # of aborted _____
Preparatory Acts or Behavior: Acts or preparation towards imminently making a suicide attempt. This can include anything beyond a verbalization or thought, such as assembling a specific method (e.g., buying pills, purchasing a gun) or preparing for one's death by suicide (e.g., giving things away, writing a suicide note). ***Have you taken any steps towards making a suicide attempt or preparing to kill yourself (such as collecting pills, getting a gun, giving valuables away or writing a suicide note)?*** If yes, describe:	**Yes** ☐ **No** ☐
Suicidal Behavior: Suicidal behavior was present during the assessment period?	**Yes** ☐ **No** ☐
Suicide:	**Yes** ☐ **No** ☐
Answer for Actual Attempts Only	Most Lethal Attempt Date:
Actual Lethality/Medical Damage: 0. No physical damage or very minor physical damage (e.g., surface scratches). 1. Minor physical damage (e.g., lethargic speech; first-degree burns; mild bleeding; sprains). 2. Moderate physical damage; medical attention needed (e.g., conscious but sleepy, somewhat responsive; second-degree burns; bleeding of major vessel). 3. Moderately severe physical damage; *medical* hospitalization and likely intensive care required (e.g., comatose with reflexes intact; third-degree burns less than 20% of body; extensive blood loss but can recover; major fractures). 4. Severe physical damage; *medical* hospitalization with intensive care required (e.g., comatose without reflexes; third-degree burns over 20% of body; extensive blood loss with unstable vital signs; major damage to a vital area). 5. Death	*Enter Code* _____
Potential Lethality: Only Answer if Actual Lethality=0 Likely lethality of actual attempt if no medical damage (the following examples, while having no actual medical damage, had potential for very serious lethality: put gun in mouth and pulled the trigger but gun fails to fire so no medical damage; laying on train tracks with oncoming train but pulled away before run over). 0 = Behavior not likely to result in injury 1 = Behavior likely to result in injury but not likely to cause death 2 = Behavior likely to result in death despite available medical care	*Enter Code* _____

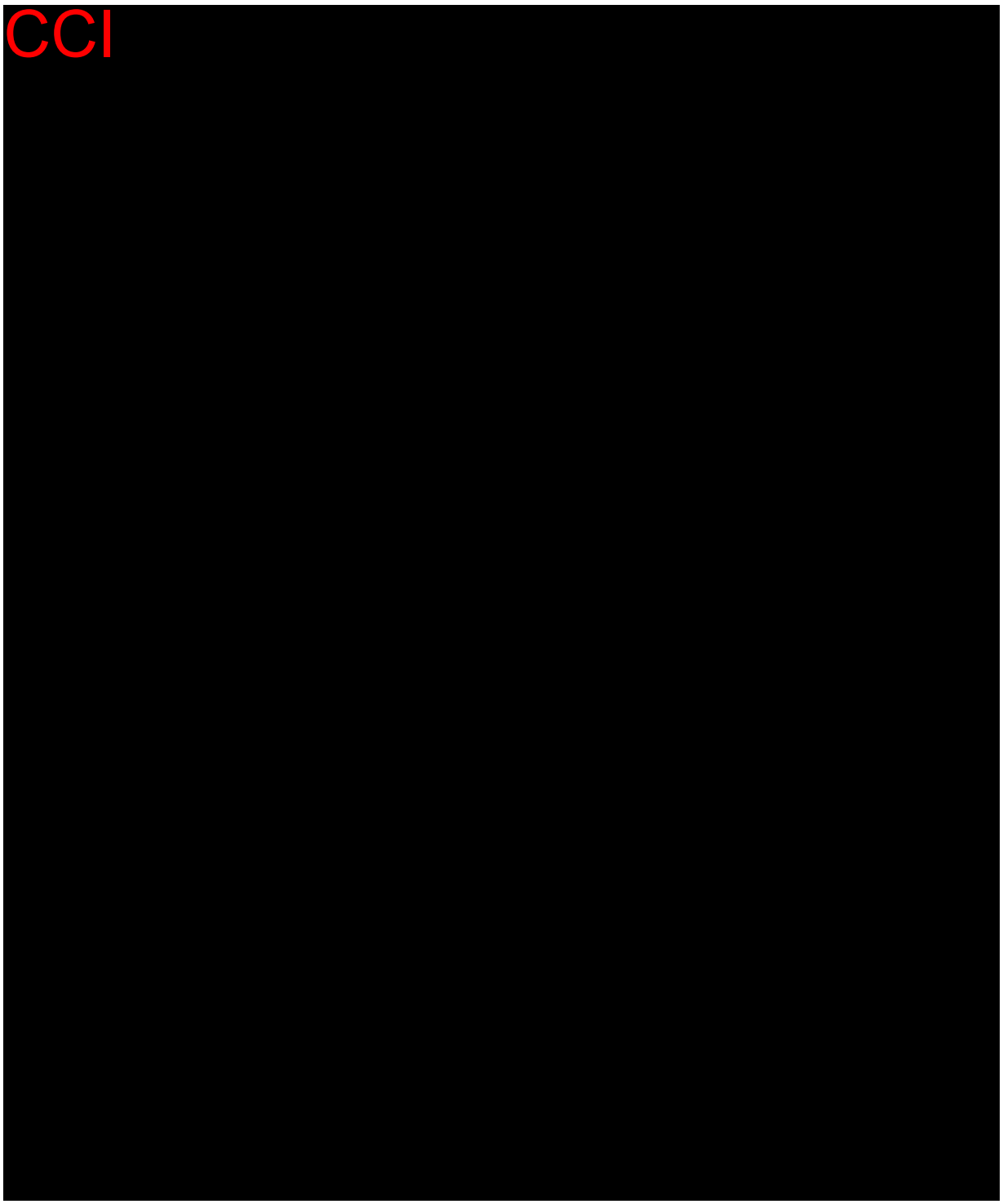

CCI


CCI


CCI

CCI

[redacted]


CCI


CCI

13.11. CTCAE Grading Details from the Central Laboratory

CTCAE CHECKS

NCI Common Terminology Criteria (CTC) for Adverse Events (AE)
Laboratory Tests

The NCI CTC for AE is a descriptive terminology which can be utilized for AE reporting. A grading (severity) scale is provided for each AE term. (CTCAE, Version 4.03)

Grade I: Mild AE
Grade II: Moderate AE
Grade III: Severe AE
Grade IV: Life-threatening or disabling AE

LLN = Lower Limit of Normal range
ULN = Upper Limit of Normal range

NOTES: The ranges in this table are adjusted to the precision and units used by BARC. In case the grading ranges are in absolute values, the ranges are listed in the different units in which the parameter can be reported. In case the grading ranges are expressed as a factor of the lower and/or upper limit of the normal ranges, no units are shown. It is possible that in some cases, the lower or upper limit of normal range is at the same time a Grade I AE.
Please note CTC checks only apply on parameters taken within the scope of the same visit.
More information about NCI CTC can be accessed on the CTEP website "http://ctep.cancer.gov/reporting/ctc.html"

Parameter	Grade I	Grade II	Grade III	Grade IV	Units
		HEMATOLOGY			
CD4 count	500 - < LLN[4]	200 - < 500	50 - < 200	< 50	/mm³
Hemoglobin	10.0 - < LLN	8.0 - < 10.0	< 8.0	-	g/dL
	100 - < LLN	80 - < 100	< 80	-	g/L
	ULN + > 0.0-2.0	ULN + > 2.0-4.0	ULN + > 4.0	-	g/dL
	ULN + > 0-20	ULN + > 20-40	ULN + > 40	-	g/L
Leukocytes	3.0 - < LLN	2.0 - < 3.0	1.0 - < 2.0	< 1.0	10^3/µL or 10^9/L
	-	-	> 100.0	-	10^3/µL or 10^9/L
Lymphocytes	0.80 - < LLN	0.50 - < 0.80	0.20 - < 0.50	< 0.20	10^3/µL or 10^9/L
	-	> 4.00 – 20.00	> 20.00	-	10^3/µL or 10^9/L
Neutrophils	1.50 - < LLN	1.00 - < 1.50	0.50 - < 1.00	< 0.50	10^3/µL or 10^9/L
Platelets	75 - < LLN	50 - < 75	25 - < 50	< 25	10^3/µL or 10^9/L
		COAGULATION			
Fibrinogen	0.75xLLN - < LLN	0.5xLLN - < 0.75xLLN	0.25xLLN - < 0.5xLLN	< 0.25xLLN	
PT (INR)	> ULN - 1.5xULN	> 1.5xULN – 2.5xULN	> 2.5xULN	-	
APTT	> ULN - 1.5xULN	> 1.5xULN – 2.5xULN	> 2.5xULN	-	
		BIOCHEMISTRY			
Albumin	3.0 - < LLN	2.0 - < 3.0	< 2.0	-	g/dL
	30 - < LLN	20 - < 30	< 20	-	g/L
Alkaline phosphatase	> ULN - 2.5xULN	> 2.5xULN - 5xULN	> 5xULN - 20xULN	> 20xULN	
ALT	> ULN – 3xULN	> 3xULN - 5xULN	> 5xULN - 20xULN	> 20xULN	
Amylase	> ULN - 1.5xULN	> 1.5xULN - 2xULN	> 2xULN - 5xULN	> 5xULN	
AST	> ULN – 3xULN	> 3xULN - 5xULN	> 5xULN - 20xULN	> 20xULN	
Bilirubin, total	> ULN - 1.5xULN	> 1.5xULN - 3xULN	> 3xULN - 10xULN	> 10xULN	

Parameter	Grade I	Grade II	Grade III	Grade IV	Units
Corrected calcium	8.0 - < LLN	7.0 - < 8.0	6.0 - < 7.0	< 6.0	mg/dL
	2.00 - < LLN	1.75 - < 2.00	1.50 - < 1.75	< 1.50	mmol/L
	> ULN - 11.5	> 11.5 - 12.5	> 12.5 - 13.5	> 13.5	mg/dL
	> ULN - 2.90	> 2.90 - 3.10	> 3.10 - 3.40	> 3.40	mmol/L
Cholesterol, total	> ULN - 300	> 300 - 400	> 400 - 500	> 500	mg/dL
	> ULN - 7.75	> 7.75 - 10.34	> 10.34 - 12.92	> 12.92	mmol/L
Creatine kinase	> ULN - 2.5xULN	> 2.5xULN - 5xULN	> 5xULN - 10xULN	> 10xULN	
Creatinine (serum)	> ULN - 1.5xULN	> 1.5xULN - 3xULN	> 3xULN - 6xULN	> 6xULN	
eGFR	60 - < LLN[4]	30 - 59	15 - 29	< 15	mL/min (/1.73m^2)
GGT	> ULN - 2.5xULN	> 2.5xULN - 5xULN	> 5xULN - 20xULN	> 20xULN	
Glucose (2h PP; random)	55 - < LLN[1]	40 - < 55	30 - < 40	< 30	mg/dL
	3.00 < LLN[1]	2.20- < 3.00	1.70- < 2.20	< 1.70	mmol/L
Glucose (fasting)	55 - < LLN	40 - < 55	30 - < 40	< 30	mg/dL
	3.00 < LLN	2.20- < 3.00	1.70- < 2.20	< 1.70	mmol/L
	> ULN - 160	> 160 - 250	> 250 - 500	> 500	mg/dL
	> ULN - 8.90	> 8.90 - 13.90	> 13.90 - 27.80	> 27.80	mmol/L
Haptoglobin	< LLN	-	-	-	
Lipase	> ULN - 1.5xULN	> 1.5xULN - 2xULN	> 2xULN - 5xULN	> 5xULN	
Magnesium	1.2 - < LLN	0.9 - < 1.2	0.7 - < 0.9	< 0.7	mg/dL
	0.50 - < LLN	0.40 - < 0.50	0.30 - < 0.40	< 0.30	mmol/L
	> ULN – 3.0	-	> 3.0 – 8.0	> 8.0	mg/dL
	> ULN - 1.23	-	> 1.23 - 3.30	> 3.30	mmol/L
Phosphorus	2.5 - < LLN	2.0 - < 2.5	1.0 - < 2.0	< 1.00	mg/dL
	0.80 - < LLN	0.60 - < 0.80	0.30 - < 0.60	< 0.30	mmol/L
Potassium	3.0 - < LLN[2]	3.0 - < LLN[2]	2.5 - < 3.0	< 2.5	mmol/L
	> ULN - 5.5	> 5.5 - 6.0	> 6.0 - 7.0	> 7.0	mmol/L
Sodium	130 - < LLN	-	120 - < 130	< 120	mmol/L
	> ULN - 150	> 150 - 155	> 155 - 160	> 160	mmol/L
Triglycerides	150 - 300	> 300 - 500	> 500 - 1000	> 1000	mg/dL
	1.71 - 3.42	> 3.42 - 5.70	> 5.70 - 11.40	> 11.40	mmol/L
Troponin I	> 0.01 - < 0.30 (♀)		≥ 0.30 (♀ & ♂)		ng/mL or µg/L
	> 0.03 - < 0.30 (♂)				ng/mL or µg/L
Troponin T	> 0.01 - <0.10	-	≥ 0.10	-	ng/mL or mg/L
Uric acid	> ULN - 10.0[3]	-	> ULN - 10.0[3]	> 10.0	mg/dL
	> ULN - 590[3]	-	> ULN - 590[3]	> 590	µmol/L
URINALYSIS					
Creatinine clearance	60 - < LLN	30 - 59	15 - 29	< 15	mL/min
Blood	+(++)	-	-	-	
Protein	+	++ or +++ (≥ 18y)	-	-	
	< 1000	1000 – 3499 (≥18y)	≥ 3500 (≥ 18y)	-	mg/24h
Protein/creatinine ratio (< 18y)		0.5 – 1.9	> 1.9		

[1] Glucose: BARC reference range has no LLN, therefore 74 mg/dL (4.11 mmol/L) is applied as LLN in Grade I flagging (reference Tietz Clinical guide to laboratory tests 4th ed.)
[2] Potassium: Grade II when intervention is indicated, else Grade I
[3] Uric acid: Grade III, when in combination with physiologic consequences, else Grade I
[4] CD4 count & eGFR: LLN is lower or equal to lower threshold for Grade I flagging, therefore only the exact lower threshold value is flagged as Grade I

13.12. Tables, Listings, and Figures

Table, listing, and figure shells will be provided in a separate living mock shell document.

13.13. Other SAP(s) Describing Analyses not Covered in the SAP

SAPs describing analyses (pharmacokinetics [PK], PK/pharmacodynamics[PD], etc.) not covered within this SAP will be included within this Appendix.

BOS161721-02: A RANDOMIZED DOUBLE-BLIND PHASE 1b/2 COMBINED STAGGERED MULTIPLE DOSE ESCALATION STUDY OF BOS161721 IN SYSTEMIC LUPUS ERYTHEMATOSUS (SLE) PATIENTS ON A BACKGROUND OF LIMITED STANDARD OF CARE

Statistical Analysis Plan

VERSION 3.0
DATE OF PLAN:

13Nov2020

STUDY DRUG:
BOS161721

PREPARED FOR:

Boston Pharmaceuticals, Inc.

Approval Signature Page: Array Biostatistics, LLC

Document Author: PPD, MS Director of Biostatistics	Date
Document Reviewer: PPD, PhD Vice President, Biostatistical Consulting	Date

Approval Signature Page: Boston Pharmaceuticals, Inc.

PPD
Clinical Operations Lead

Date

PPD, MD
Vice President, Clinical Development

Date

PPD
Sr. Director, Biostatistics

Date

Contents

ABBREVIATIONS

ACR	American College of Rheumatology
ADA	Anti-drug antibody
AE	Adverse event
AESI	Adverse event of special interest
ANA	Antinuclear antibody
ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
APL	Antiphospholipid
ATC	Anatomic therapeutic class
AUC	Area under the curve
BICLA	BILAG-based Composite Lupus Assessment
BILAG	British Isles Lupus Assessment Group
BP	Blood pressure
C3	Complement C3
C4	Complement C4
C_L	Systematic clearance
CLASI	Cutaneous Lupus Area and Severity Index
C_{max}	Maximum plasma concentration
CS	Corticosteroid
CSR	Clinical Study Report
C-SSRS	Columbia Suicide Severity Rating Scale
CTCAE	Common Terminology Criteria for Adverse Events
CV	Coefficient of variation
DLT	Dose-limiting toxicity
DMC	Data Monitoring Committee
dsDNA	Double-stranded deoxyribonucleic acid
ECG	Electrocardiogram
eCRF	Electronic case report form
EWB	Emotional well being
CCI	[redacted]
FAS	Full analysis set
FS	Fatigue subscale
FWB	Functional well being
HR	Hazard ratio
IA	Interim Analysis
IEC	Independent Ethics Committee
IL-21	interleukin 21
IM	Intramuscular
IRB	Institutional Review Board
ISR	Injection site reaction
LLN	Lower limit of normal
LLOQ	Lower limit of quantification
LOCF	Last observation carried forward

LS	Least squares
MAD	Multiple Ascending Dose
MCS	Mental component summary
MedDRA	Medical Dictionary for Medical Affairs
NAb	Neutralizing antibody
NCI	National Cancer Institute
PCS	Physical component summary
PD	Pharmacodynamics
PGA	Physician's Global Assessment
PK	Pharmacokinetics
POC	Proof of concept
PP	Per protocol set
pSTAT3	Phosphorylated signal transducer and activator of transcription 3
PT	Preferred Term
PWB	Physical well being
Q1	25^{th} percentile/first quartile
Q3	75^{th} percentile/third quartile
SAD	Single ascending dose
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SAS	Statistical analysis system
SC	Subcutaneously
SD	Standard deviation
CCI	[redacted]
SI	International System of Units
SLE	Systemic Lupus Erythematosus
SLEDAI-2K	SLE Disease Activity Index 2000
SLICC	Systemic Lupus International Collaborating Clinics
Sm	Smith
SOC	System Organ Class
SRI-4	SLE Responder Index 4
SRI-5	SLE Responder Index 5
SRI-6	SLE Responder Index 6
SS	Safety analysis set
SSA	Sjögren syndrome-A (Ro)
SSB	Sjögren syndrome-B (La)
SWB	Social/Family well being
$t_{1/2}$	Terminal elimination half-life
TEAE	Treatment emergent adverse event
T_{max}	First time to maximum concentration
TTMF	Time to medication failure
ULN	Upper limit of normal
V_d	Volume of distribution
WHO	World Health Organization

1. INTRODUCTION

The Statistical Analysis Plan (SAP) details the planned analyses that will be included in the Clinical Study Report (CSR) of study number BOS161721-02: A Randomized Double-Blind Phase 1b/2 Combined Staggered Multiple Dose Escalation Study of BOS161721 in Systemic Lupus Erythematosus (SLE) Patients on a Background of Limited Standard of Care.

SAPs describing analyses (pharmacokinetics [PK], PK/pharmacodynamics[PD], etc.) not covered within this SAP will be included within Appendix 13.13 of this document. The author, finalization date, signatories, and page numbering of this SAP are independent of those SAPs and vice versa. Only the last final versions will be included in the appendix just prior to study conclusion.

The content of this SAP is based on the protocol Version 7.0 (Amendment 6) dated 30April2020.

SAP Revision Chronology:	**Date:**	**Comments:**
V1.0	05SEP2018	Original
V2.0	29NOV2018	Modifications due to Protocol Version 4 (Amendment 3) • Modified evaluable window for efficacy assessments to be inclusive of 35 days since prior assessment. • Modified structure of table shells to combine analyses (i.e., CLASI total activity and total damage; ACR-28 sum of swelling and sum of tenderness; CCI [redacted] biomarker and pharmacodynamic parameters) • Modified structure of listing shells CCI [redacted] • Clarified Last Observation Carried Forward (LOCF) versus observed analyses in the output titles within the shells
V3.0	13NOV2020	Modifications due to Protocol Version 7 (Amendment 6) including: • Sample size changed based on recent published literature • LOCF algorithm updated • Modified planned sensitivity analyses based on missing data • Added, amended, and removed secondary and exploratory endpoints per protocol amdement 6 • Defined efficacy subgroups • Added COVID-19 related data handling details • Modified evaluable window for efficacy assessments • Modified table, listing, and figure shells in accordance with text changes

2. STUDY OBJECTIVES AND ENDPOINTS

2.1. Multiple Ascending Dose (MAD) Phase 1b Study

Objectives	Endpoints
Primary	
• To assess safety, tolerability, and immunogenicity of repeat doses of BOS161721 (20, 60, and 120 mg) administered subcutaneously (SC) in adult patients with moderately to severely active SLE on limited background standard of care treatment, in order to estimate the optimal dose.	Safety Endpoints • Incidence and severity of adverse events (AEs) and serious adverse events (SAEs), related AEs, AEs leading to study drug discontinuation, AEs by severity and relatedness • Injection site reactions • Columbia Suicide Severity Rating Scale (C-SSRS) • 12-lead electrocardiograms (ECGs) parameter results at each visit and change from baseline • Vital signs (blood pressure [BP], heart rate, and temperature) parameter results at each visit and change from baseline • Clinical laboratory results and change from baseline • Physical examinations changes from baseline • Anti-drug antibodies (ADAs) • Study drug exposure/compliance

Objectives	Endpoints
Secondary	
• To characterize the PK of BOS161721 and select the optimal dose of BOS161721 based on safety, PK, and PD effects in patients with moderate to severely active SLE.	Pharmacokinetic (PK) Endpoints • BOS161721 concentration by visit and time point • Maximum observed concentration (C_{max}), first time to maximum concentration (T_{max}), area under the concentration-time curve (AUC), terminal half-life ($t_{½}$), systematic clearance (C_L), volume of distribution (V_d) Pharmacodynamic (PD) Endpoints • Results and changes (or shifts) from baseline to each visit in phosphorylated signal transducer and activator of transcription 3 (pSTAT3), C3 and C4 levels, and leukocyte immunophenotype • Results and changes (or shifts) from baseline in anti-double-stranded DNA (dsDNA), antinuclear antibodies (ANA), anti-Sjögren syndrome A and B (SSA, SSB), Smith (Sm), and antiphospholipid (APL) autoantibodies at each visit • Results and changes (or shifts) from baseline in abrogation of IL-21 gene signature at each indicated visit
Exploratory	
• CCI [redacted]	CCI [redacted]

Objectives	Endpoints
	- CCI • CCI • CCI
• CCI	• CCI
• CCI	• CCI

[1] See Section 11

2.2. Proof of Concept (POC) Phase 2 Study

Objectives	Endpoints
Primary	
• To demonstrate a superior effect of BOS161721 at the chosen dose compared with placebo for response on the SLE SRI-4	Primary Efficacy Endpoint • The proportion of patients with a SRI-4 response at Day 210
Secondary	
• To demonstrate a superior effect of BOS161721 at the chosen dose compared with placebo for response on clinical indicators of SLE activity, in adult patients with moderately to severely active SLE on limited background standard of care treatment	Secondary Efficacy Endpoints • The proportion of patients with: - SRI-4 response at each visit - SRI-5 and SRI-6 response at each visit - a sustained reduction of oral corticosteroid (CS) (≤ 7.5 mg/day and < Day 0 dose) between Day 150 and Day 210 - new or recurrent BILAG flares (≥ 1 qualifying BILAG A or > 1 qualifying BILAG B) through Day 210 - PGA worsening - a BICLA response - a CLASI response - medication failures • Change and percentage change from baseline at each visit in: - CLASI - PGA - Total number of swollen joints, tender joints, and active joints (swelling and tenderness in the same joint) in the ACR-28 joint count - SLEDAI-2K - SLICC/ACR damage index • TTMF • Group mean percent reduction in corticosteroid administration from baseline Day 0 dose through Day 210 in patients receiving ≥ 7.5 mg/day prednisone equivalent at Day 0

Objectives	Endpoints
	• Duration of longest SRI-4 response • Time to first BILAG flare (≥ 1 new or recurrent BILAG A or > 1 new or recurrent BILAG B) relative to baseline through Day 210
Safety	
• To assess safety and tolerability of repeat doses of BOS161721 at the chosen dose administered SC in adult patients with moderately to severely active SLE on limited background standard of care treatment	Safety Endpoints • Incidence and severity of AEs and SAEs, related AEs, AEs leading to study drug discontinuation, AEs by severity and relatedness • Injection site reactions • C-SSRS • 12-lead ECGs parameter results at each visit and change from baseline • Vital signs (BP, heart rate, and temperature) parameter results at each visit and change from baseline • Clinical laboratory results and change from baseline • Physical examinations changes from baseline • ADAs • Study drug exposure/compliance

Objectives	Endpoints
Exploratory	
• CCI	• CCI
• CCI	CCI
• CCI	• CCI
• CCI	• CCI

Endpoints will be discussed in the SAP based on their order in the phase 2 POC portion of the study. The phase 1b MAD endpoints overlap with and will be discussed within the context of the phase 2 POC endpoints.

3. STUDY DESIGN

3.1. Study Design and Population

This is a Phase 1b/2 combined, randomized, multicenter, double-blind, placebo-controlled trial to study the clinical efficacy, safety, and PK of multiple subcutaneous doses of BOS161721 in adult

patients with moderately to severely active SLE. After successfully completing a screening phase, eligible patients will be randomized to a specified dose of BOS161721 or placebo. The dosing schedule will be monthly.

The trial will consist of 2 double-blinded portions: MAD Phase 1b and POC Phase 2. Patients may receive a total of 7 SC monthly doses of study drug on Days 0, 30, 60, 90, 120, 150, and 180, followed by safety follow-up visit at Days 210, 240, and 270.

Investigators will assess SLE activity in accordance with accepted evaluation tools (further details provided in the SLE Assessment Reference Guide).

SLE disease activity assessment data will be centrally reviewed by the CRO medical monitor(s), independent subject matter expert reviewers/adjudicators and the Sponsor to ensure the scores are clinically meaningful, compliant with specific definitions and compatible across various disease scoring tools. The scope of responsibility includes, but is not limited to, review and confirmation of BILAG organ system disease grades, BILAG A and B flares, and clinical components of SLEDAI-2Kscore at screening and during the study as well as cross-validation of the instruments used in this study to assess disease activity. Further details on the content, processes, procedures and methods of disease activity data reports will be outlined in the Medical Data Review Plan and the BILAG-2004 Review Guidelines.

3.1.1. Multiple Ascending Dose Phase 1b

The MAD portion will consist of 3 cohorts:

- Cohort 1 (20 mg SC) will include 6 patients
    - 5 patients will receive BOS161721 (active group) and 1 patient will receive placebo (placebo group)
- Cohorts 2 (60 mg SC) and 3 (120 mg SC) will include 12 patients each
    - 9 patients in the active group and 3 in the placebo group

Doses selected for each of the 3 cohorts is based on a 90-day safety, tolerability, PK and PD data review from the Phase 1 single ascending dose (SAD) study (BOS161721-01) in healthy patients. All doses selected for the MAD part of the study are projected not to exceed the mean exposure of that achieved at the highest dose in the SAD study.

The MAD portion of the study design is staggered, where after the 6 patients in Cohort 1 have received 2 doses and have completed 2 weeks of follow up- after the second dose, Cohort 2 begins dosing after the Data Monitoring Committee (DMC) evaluation of the safety and tolerability data from Cohort 1. Similarly, after 8 of the 12 patients in Cohort 2 have received 2 doses and have completed 2 weeks of follow up- after the second dose, Cohort 3 begins dosing after the DMC evaluation of the safety and tolerability data from Cohorts 1 and 2. Each cohort will continue at their assigned dose level through their respective 6-month treatment periods (See

Study Schematic in the study protocol). If patients discontinue the study in a cohort prior to adequate safety follow-up, he/she may be replaced.

Criteria for dose escalation are further described in the DMC Charter.

3.1.2. Proof of Concept Phase 2

The dose for the POC portion of the study is 120 mg administered SC monthly (a total of 7 doses). The rationale for the BOS161721-02 phase 2 POC dose selection was based on cumulative safety, tolerability, immunogenicity, PK, and PD data available from an interim analysis (IA) performed during the MAD phase 1b portion of the trial.

The data cut-off for this IA occurred on CCI and included all 6 patients and 7 doses from Cohort 1 (20 mg), 12 patients and 6 doses from Cohort 2 (60 mg), and 12 patients and 4 doses from Cohort 3 (120 mg).

The safety analysis focused on incidence and severity of all AEs, SAEs, and pre-determined adverse events of special interest (AESI). The DMC and designated unblinded Boston Pharmaceuticals team met on CCI and did not identify any untoward safety signals at any BOS161721 dose levels.

Because there were no safety, tolerability, or immunogenicity trends observed at the time of the IA, the phase 2 POC dose selection was made based on available PK and PD data. pSTAT3 levels were assessed as the primary PD biomarker of IL-21R signaling levels. This is because IL-21R signaling, upon IL-21 binding, initially involves phosphorylation of JAK1/JAK3 which dissociate from the receptor complex, and phosphorylate STAT3 which translocates to the nucleus and drives IL-21-regulated gene expression. CCI

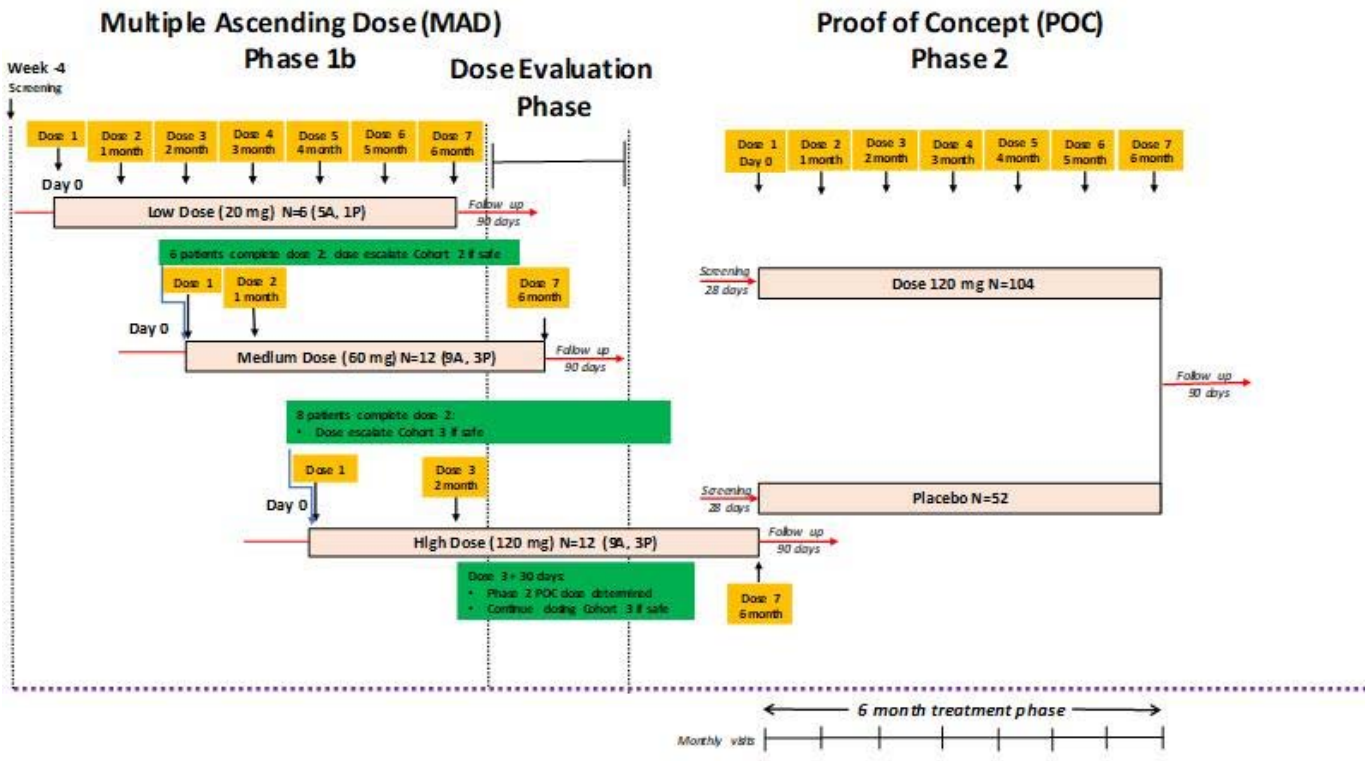
The 120 mg dose was communicated to site Investigators participating in the POC phase 2 portion, and to Institutional Review Boards (IRBs) and the Independent Ethics Committee (IEC).

For the POC part of the study, approximately 110 additional patients will be randomized to active or placebo groups in a 2:1 ratio.

As in the MAD part of the study, each patient in the POC portion may receive a total of 7 SC monthly doses of study drug on Days 0, 30, 60, 90, 120, 150, and 180. Assessments will be completed according to the Schedule of Assessments detailed in the study protocol.

DMC safety reviews will be conducted periodically throughout the study as described in the DMC charter.

Study Diagram for MAD Phase 1b and POC Phase 2



A = active drug (BOS161721); P = placebo

3.2. Randomization and Blinding

This is a randomized, double-blind study.

Patients meeting all inclusion and exclusion criteria will be centrally randomized to either placebo or BOS161721 using an interactive web response system according to a randomization list generated by an independent, non-study statistician. Randomization will be performed separately for each study portion and separately for each cohort in the Phase 1b and 2 portions as follows:

Phase/Cohort	Number of Patients	Randomization Ratio BOS161721:Placebo
Phase 1b/Cohort 1	6	5:1
Phase 1b/Cohort 2	12	3:1
Phase 1b/Cohort 3	12	3:1
Phase 2	*Approximately 110	2:1

*Additional patients may be enrolled to ensure sufficient numbers of patients are in the full analysis set (FAS).

Eligible patients will be assigned to the study portion which is active at time of enrollment. Similarly, patients in the Phase 1b MAD will be assigned to the cohort which is active. Each patient will be assigned a unique randomization number which will not be reused.

All patients, investigators, and study team participants will be blinded to treatment assignment. An independent biostatistician not otherwise involved on the study will be unblinded and prepare materials for the IA, ad hoc analyses as needed, and the DMC safety reviews. The DMC will review unblinded data during safety reviews and the IA. A limited team at Boston Pharmaceuticals will review unblinded results from the Phase 1b MAD portion during the IA to determine the dose that will be used for the Phase 2 POC portion of the study. Details regarding maintenance of the blinding and content of data reviews will be described in the DMC charter or related study documentation.

3.3. Sample Size Considerations

Sample size in the Phase 1b part of the study is based on operational consideration.

The sample size in the Phase 2 POC study is based on the primary endpoint, SRI-4 Response at Day 210. Based on efficacy data from a recently completed study with Ustekinumab[iii], where the primary endpoint was also SRI-4, the sample size in the phase 2 part of the BOS161721-02 study is set at 96 evaluable patients. In the Ustekinumab study of 102 patients, a statistically significant difference in SRI-4 response was shown in favor of Ustekinumab (62% of Ustekinumab treated patients achieved SRI-4 response vs. 33% in patients receiving placebo). The Ustekinumab phase 2 study is an appropriate comparator to BOS161721-02 because both studies share the same primary endpoint of SRI-4 response and BOS161721 and Ustekinumab have overlapping mechanisms of action (T follicular helper and Th17 cell biology).

Approximately 110 additional patients will be randomized in a 2:1 ratio to BOS161721 versus placebo to achieve approximately 96 evaluable patients in the FAS of the phase 2 portion. This assumes 14 patients will discontinue prior to having received treatment or having completed a post-baseline efficacy measure. Enrollment and randomization will be monitored in a blinded fashion and may be increased if needed to ensure at least 96 evaluable patients are randomized into the FAS (see Section 6.1).

A total of 96 evaluable patients randomized provides 80% power to detect a treatment difference of 29% in SRI-4 response rates at Day 210, based on a targeted 2-sided significance level of 10% and using a 2-sided Pearson's chi-squared test. This assumes a response rate of 62% for BOS161721 and 33% for placebo.

3.4. Data Monitoring Committee (DMC)

This study will use an external DMC. The DMC is an independent committee established to provide safety oversight and to advise the sponsor regarding actions necessary to protect study participants and the scientific integrity of the study. The recommendations made by the DMC

(i.e., dose escalation, etc.) will be forwarded to the sponsor for final decision. The sponsor will forward such decisions, which may include summaries of safety data which are not endpoints, to regulatory authorities, as appropriate.

The sponsor will appoint a DMC for the periodic review of available study data. The members of the DMC serve in an individual capacity and provide their expertise and recommendations. The primary responsibilities of the DMC are to (1) periodically review and evaluate the accumulated study data for participant safety, study conduct, and progress, (2) make recommendations to the sponsor concerning the continuation, modification, or termination of the study and (3) suggest dose for POC portion of the study as described in the study protocol.

The DMC considers study-specific data as well as relevant background knowledge about the disease, test agent, or patient population under study. The DMC is responsible for defining its deliberative processes, including event triggers that would call for an unscheduled review, stopping guidelines, unmasking (unblinding), and voting procedures prior to initiating any data review. The DMC is also responsible for maintaining the confidentiality of its internal discussions and activities as well as the contents of reports provided to it.

The DMC will have access to unblinded treatment information during the clinical trial. Details regarding management and process of this committee are found in the DMC Charter.

The DMC may recommend termination of BOS161721 treatment arm or the entire BOS161721 MAD/POC trial for any safety concern that is felt to outweigh potential benefits. The recommendation must be supported by the sponsor as indicated in the DMC Charter.

3.5. Interim Analysis (IA)

One IA is planned for this study and will be conducted at the time of dose selection for POC. The POC dose will be selected based on cumulative available safety, tolerability, PK and PD data based on the MAD data. Since there is no IA during the POC part of the study, there is no impact on the type 1 error.

3.6. Timing of Analyses

The following analyses are planned:

- An IA will be performed during the last cohort of the MAD portion to determine dose selection for the POC portion.
- The final analysis will be performed when all patients have either completed the POC safety follow up or withdrawn from the study.
- Safety analyses will be performed for DMC reviews throughout the study to evaluate accumulating safety data and make dose escalation decisions. The frequency and details of the content and format of the safety review meetings are described in the DMC charter and SAP and applicable mock shells.

4. DATA ANALYSIS CONSIDERATIONS

All analyses will be conducted based on SAS 9.3 or higher.

All data in the database will be presented in by-patient data listings.

Unless otherwise stated, all listings will be sorted by study phase, treatment group, patient number, and assessment date (and time, if available). If assessment date is missing, chronological visit will be used for sorting.

Unless stated otherwise, continuous data will be summarized by treatment group based on number of patients (n), mean, median, standard deviation (SD), 25th percentile/first quartile (Q1), 75th percentile/third quartile (Q3), minimum value, and maximum value.

The geometric % coefficient of variation (CV) is calculated as 100*sqrt(exp(SDlog^2)-1) where SDlog is the SD of log-transformed values.

Unless stated otherwise, categorical data will be summarized by treatment group using n and percentage based on the number of non-missing values.

- The number of missing values will be presented as a separate category with no percentage, but only if one or more patients are missing data.
- Counts of zero will be presented without percentages.

Statistics will be presented in the summary tables based on the following:

- Minimum and maximum: same number of significant digits as the raw value
- Mean, median, Q1, and Q3: one additional decimal place to that reported for minimum and maximum
- SD: two additional decimal places than presented for the minimum and maximum
- Percentages: reported to one decimal place with the exception of 100% which will be displayed without decimals.
- P-values will be reported to four decimal places. If the value is below 0.0001 it will be noted as < 0.0001; if the value above 0.9999 it will be noted as > 0.9999.

Unless otherwise noted, statistical inference will be based on a 10% significance level (i.e. 90% confidence intervals will be produced).

Statistical testing will be performed on data from the Phase 2 study.

Numbering for data displays will be based on ICH E3.

4.1. Stratification and Covariates

The effects of noncompliance, background therapy use, treatment discontinuations, premature withdrawal from study and covariates will be assessed to determine the impact on the general applicability of results from this study. Exploratory analyses of the data may be conducted as deemed appropriate to evaluate factors for which analyses are not controlled.

Continuous efficacy endpoints will be assessed via analysis of variance (ANOVA) or analysis of covariance (ANCOVA) when applicable, adjusting for the baseline value.

Sensitivity analyses of the primary efficacy endpoint may be performed with covariates defined based on the subgroups in section 4.2.

4.2. Evaluation of Subgroups

CCI

CCI [redacted]

[redacted]

The patient will be excluded from subgroup analyses if they are missing the subgroup parameter at baseline.

4.3. Multiple Comparisons and Multiplicity

The Type 1 error rate will not be adjusted for multiplicity.

5. GENERAL DATA HANDLING CONVENTIONS

5.1. Reference Dates

- Screening date is defined as the electronic case report form (eCRF) provided date on which a patient was screened for trial entry.
- Informed consent date is the date the patient signed the informed consent form.
- Randomization date is defined as the date on which the patient is randomized to study treatment.
- Treatment start date is defined as the date of first dose of study drug.
- Treatment end date is defined as the date of last dose of study drug.
- Age will not be calculated and will come directly from the eCRF. The eCRF uses the informed consent date as its reference date for age calculation.
- Safety data, such as AEs and laboratory assessments will use the treatment start date as a reference date.
- Efficacy data will use the randomization date as a reference date.
- Study day will be based on treatment start date as a reference date.

5.2. Study Day and Duration Variables

Reference date calculations will be defined as follows, assuming non-missing dates:

- date of interest – reference date + 1 when the date of interest is on or after the reference date;
- otherwise, date of interest – reference date.

If either date is missing or incomplete, reference date calculations will not be performed. Date imputation will be performed as identified in Section 5.5.

Study day will either have a negative value if collected before dosing or a positive value if collected on or after the day of drug dosing; there will be no study day zero.

Duration of time is dependent on reference dates and will be calculated in a manner similar to that of the reference date calculation. Duration on study is defined as the end of study date – informed consent date + 1. Duration of treatment is defined as treatment end date – treatment start date + 1, where treatment end date is the date of last dose of study drug.

Time-to-event endpoints are followed until first event or censoring. As a result, event time will be calculated as: event or censoring date – reference date + 1. These are further described in Section 7.

When reporting time-to-event data (TTMF and time to BILAG A flare or >1 BILAG B flare) or duration outcomes, the results (in days) above will be converted to an appropriate unit. When reporting in months it will be divided by 30.4375; for reporting in weeks it will be divided by 7; and for reporting in years it will be divided by 365.25.

5.3. Study Time Periods

Where applicable, data reporting will be classified by the following study periods for analysis:

- Pre-treatment is defined as the period prior to a patient's treatment start date.
- On-treatment is defined as the period between a patient's treatment start date and treatment end date, inclusive.
- Post-treatment is defined as the period of time following the on-treatment period.

5.4. Baseline, Post-Baseline Changes, and LOCF

Unless stated otherwise, baseline and post-baseline change values will be based on the following:

- Baseline will be based on the last non-missing value collected prior to or on the date/time of first dose of study drug.
- Post-baseline values will be those collected after the date/time of first dose of study drug.
- Change from baseline is defined as: value – baseline value.
- Percentage change from baseline is defined as: [(value – baseline value) / baseline value]*100. Patients with baseline values of zero are excluded due to division by zero.
- Most extreme change: The maximum most extreme change will be the maximum post-baseline change value; the minimum most extreme change will be the smallest post-baseline change value. This calculation will consider all assessments collected after the first dose of study drug, scheduled or unscheduled.
- LOCF:
    - Imputation rules for assessments that include multiple components:

        - In case of missing laboratory components, lab values from the most recent unscheduled visit <=14 days after the nominal visit will be utilized. If lab values are not available from an unscheduled visit <=14 days after the visit, laboratory data will be carried forward from the most recent laboratory assessment <=30 (scheduled or unscheduled) days prior to the nominal visit.
        - In case of missing non-laboratory components, no imputation will be performed.
    - In case of missing the entire score for an assessment, imputation will be based on LOCF. LOCF will only be implemented with the most recent prior visit (including unscheduled assessments). If the most recent prior visit is missing, occurred more than 30 days before (current visit date – most recent prior visit date > 30), or is not evaluable, LOCF is not implemented. See Section 7 for further details. Additional details regarding LOCF for components for specific assessments are described in Section 7.
    - Observed values do not incorporate any imputations and are presented as collected.

5.5. Imputation of Partial Dates

Adverse Events

- If the AE start date is completely missing, or if the patient was not treated, no imputation will be conducted.
- If the AE start date is missing day and month, do the following:
    - If the treatment start date is missing or the AE start year does not fall in the same year as that of the treatment start date or if the AE record contains information to indicate that the event ended before the treatment start date (e.g. the AE end date month and year are earlier than the treatment start date or the full AE end date is known and occurs earlier than the treatment start date), then set the AE start month and day to January 1st.
    - Otherwise, set the AE start date to the treatment start date.
- If only the AE start day is missing, do the following:
    - If the study treatment start date is missing or the AE start month and year does not fall in the same month and year as that of the treatment start date or if the AE contains information to indicate that the event ended before the treatment start date, then set the AE start day to the 1st day of the month of the AE start date.
    - Otherwise, set the AE start date to the treatment start date.
- AE end dates will not be imputed.

Prior, Concomitant, and Subsequent Medications (Excluding Corticosteroids)

- The imputation rules for AE start dates will be used for medication start dates.
- Medication stop dates will be imputed as follows:

- If the stop date is only missing the day, then the stop day is the last day of the month
- If the stop date is missing both the day and month and the year matches the last study date, then the stop month and day is the earlier of the last study date and December 31
- If the stop date is missing both the day and month and the year is earlier than the last study date, then the stop month and day is December 31
- If the stop date is completely missing, no imputation is performed, and the medication will be classified as a concomitant medication for patients who were treated.

Corticosteroids

Due to the criticality of the corticosteroid usage, handling of partial start/stop dates will be performed on a case-by-case basis as it will require medical review of the patient records. Such determinations will be made in a blinded fashion prior to database lock and approved by the Sponsor.

Diagnosis Date

- If the diagnosis date is completely missing, no imputation will be conducted.
- If the month and day are completely missing but year is present, do the following:
  - Set missing month and day to July 1st
  - If July 1st of the non-missing year is after informed consent date, set month and year to January 1st
- If only the day is missing, set to the 1st of the month

5.6. Multiple Assessments and Visit Windows

Nominal visits (e.g. those identified by the study eCRF) will be the basis of summarization and statistical analysis unless otherwise specified. Unscheduled data may be included in summaries of most extreme, baseline, LOCF, and endpoint values; summaries of specific abnormalities any time post-baseline; and patient data listings

5.7. Treatment Group Display

Treatment groups will be displayed with the following columns. See the mock shells for additional details and a visual representation of the treatment group display. All overall columns will be excluded from efficacy outputs.

- Phase 1b MAD: Cohort 1 BOS 20mg
- Phase 1b MAD: Cohort 2 BOS 60mg
- Phase 1b MAD: Cohort 3 BOS 120mg

- Phase 1b MAD: Placebo
- Phase 1b MAD: Total BOS (All patients treated with BOS161721, excludes placebo)
- Phase 2 POC: BOS 120mg
- Phase 2 POC: Placebo
- Overall: Total BOS 120mg (All patients treated with BOS161721 120mg in MAD and POC)
- Overall: Total BOS (Patients treated with BOS161721 in MAD and POC, excludes placebo)
- Overall: Total Placebo (Placebo patients in MAD and POC, excludes patients treated with BOS161721)

5.8. Missing Data

Missing data imputation for the efficacy endpoints are discussed in the corresponding efficacy sections. AE, concomitant medication, and corticosteroid date imputations are described in Section 5.5. Missing data and other analysis conventions for noncompartmental analysis of pharmacokinetics will follow standard imputation methods, as described in a separate Clinical Pharmacology SAP. Otherwise, missing data will not be imputed.

6. STUDY PATIENT DATA

6.1. Analysis Populations/Sets

FAS: Defined as all patients who receive at least one dose (partial or complete) of study treatment and have at least one evaluable post-baseline efficacy evaluation (BILAG, PGA, SLEDAI-2K, CLASI, ACR-28, SLICC/ACR, CCI , or CCI). FAS analyses will be conducted on the basis of the assigned (randomized) treatment. The FAS will be used as the basis for the primary efficacy analysis.

Safety Analysis Set (SS): Defined as all patients who receive at least one dose (partial or complete) of study treatment and have at least one evaluable post-baseline safety evaluation (adverse events, concomitant medications, vital signs, physical examination, ECG, or laboratory assessment). Safety analyses will be conducted on the basis of actual treatment received as reported on the eCRF. The SS will be the basis of all safety reporting.

Per Protocol Set (PP): Defined as all patients from the FAS except those with major deviations to the protocol deemed to impact the analysis of the primary endpoint. These deviations will be identified based on blinded data review prior to study unblinding. Patients with more than one missed dose will be excluded from the PP. PP analyses will be conducted on the basis of the assigned (randomized) treatment. Additional information on inclusion into this population can be found in Section 6.3.

Pharmacokinetic Analysis Set (PK): Defined as all patients in the FAS with sufficient data for the calculation of PK parameters. This analysis population will not be used for analyses corresponding to this SAP.

6.2. Patient Disposition

Disposition data will be summarized for all randomized patients. The number of patients in each analysis set population, the number of patients who discontinued study treatment including reasons, and the number of patients who discontinued the study including reasons will be summarized. The number of patients who completed the study, duration on study treatment in days and duration on study in days will also be summarized. Data will be presented overall, by study phase, and treatment group.

A by-patient listing of patient disposition data including reason for discontinuation, if applicable, will be presented for all randomized patients. A by-patient listing of randomization details will also be provided for all randomized patients.

A by-patient listing of inclusion/exclusion criteria and screen failures will be produced for all enrolled patients. Screen failures will otherwise not be included in any analyses.

6.3. Protocol Deviations

Protocol deviations will be identified and classified as major or minor before database lock and unblinding. Major protocol deviations may include but are not limited to:

- Violation of Inclusion/Exclusion Criteria
- Use of prohibited therapies
- Incorrect treatment

Major protocol deviations that necessitate exclusion from the efficacy analysis in the PP population will be identified prior to database lock and unblinding.

Protocol deviations will be summarized by any deviation, deviation category, and major/minor designation. Protocol deviations will be summarized overall, by study phase, and treatment group for all randomized patients.

A listing of protocol deviations, including protocol version under which the deviation occurred, will be provided for all randomized patients. A separate listing of COVID-19 related deviations also will be provided.

6.4. Demographic and Baseline Characteristics

Patient demographics and baseline characteristics will be summarized in the SS. These will include age (years), sex (Male / Female), ethnicity (Hispanic or Latino / Not Hispanic or Latino), race (American Indian or Alaska Native / Asian / Black or African American / Native Hawaiian or Pacific Islander / White / Other), geographic regions (North America / South or Central America / Eastern Europe), baseline height (cm), baseline weight (kg), and baseline BMI (kg/m^2). Age is reported as collected in the clinical database and will also be categorized as a categorical variable (≤45, >45-<65, ≥65) for reporting. Patient eligibility for central review will be summarized. Childbearing potential and birth control methods for female and male patients will be summarized for the DMC meetings and interim analysis only. Demographics and baseline characteristics will be summarized overall, by study phase, and treatment group.

Baseline disease characteristics will also be summarized in the FAS. The SLEDAI total score, PGA, CLASI total activity score, CLASI total damage score, ACR-28 sum of swelling, ACR-28 sum of tenderness, ACR-28 sum of active joints (swelling and tenderness in the same joint) and SLICC/ACR damage index total score will be summarized at baseline. Frequencies and percentages of patients with SLEDAI presence by component, ACR-28 loss of functional range interfering with life and ACR-28 impairment of basic activities from inflammation as well as BILAG grades by body organ system will be presented. Frequency and percentage of patients with each organ system disease on BILAG SLE history will be summarized. Frequency and percentage of patients with SLICC criteria for SLE will be summarized. Frequencies and percentages of patients with CLASI Activity Score (<8 vs. >=8), SLEDAI-2K (<10 vs. >=10) and baseline corticosteroid use (0, <7.5, >= 7.5, <10, >=10 mg/day), will be presented.

Frequencies and percentages of patients will also be presented for the following characteristics:

- Positive anti-ds-DNA
- Low C3 and/or C4
- Positive anti-dsDNA and normal C3 and normal C4
- Negative anti-dsDNA and low C3 and/or C4
- Positive anti-dsDNA and low C3 and/or C4
- Negative anti-dsDNA and normal C3 and normal C4
- Other

In addition, frequencies and percentages of patients will be presented for the following characteristics:

- SLEDAI-2K ≥ 10 and negative** anti-dsDNA and normal C3 and normal C4
- SLEDAI-2K ≥ 10 and positive** anti-dsDNA and normal C3 and normal C4
- SLEDAI-2K ≥ 10 and negative** anti-dsDNA and low C3 and/or C4
- SLEDAI-2K < 10 and positive** anti-dsDNA and low C3 and/or C4
- SLEDAI-2K < 10 and positive** anti-dsDNA and normal C3 and normal C4

- SLEDAI-2K < 10 and negative** anti-dsDNA and low C3 and/or C4
- SLEDAI-2K < 10 and negative** anti-dsDNA and normal C3 and normal C4
- SLEDAI-2K ≥ 10 and positive** anti-dsDNA and low C3 and/or C4

** Positive anti-dsDNA is defined as ≥ 15 IU/mL. Negative anti-dsDNA is defined as < 15 IU/mL

Time from diagnosis to informed consent in years defined as (informed consent date - date of diagnosis as recorded on the screening BILAG assessment)/365.35 will be summarized. Partial dates of diagnosis will be handled as defined in section 5.5. Baseline disease characteristics will be summarized overall by study phase and treatment group.

Demographics and baseline characteristics will be listed for all randomized patients. A separate by-patient listing will be created for childbearing potential and birth control for all randomized patients.

The following conversions and equations will be used as applicable:

Height (in cm) = height (in inches) * 2.54

Weight (in kg) = weight (in lbs) * 0.4536

BMI (kg/m^2) = weight(kg)/[height$(m)^2$]

6.5. Medical History

The incidence of medical history will be summarized by system organ class (SOC) and preferred term (PT) overall, by study phase, and treatment group in the SS. Patients will be counted once per SOC (likewise for PT within SOC). Medical history terms will be coded based on the Medical Dictionary for Regulatory Affairs (MedDRA). Medical history will be reported in descending order of overall incidence of SOC and then in descending order of overall incidence for PT within the SOC.

Medical history will be presented in data listings for all randomized patients. Medical history coded terms will be provided, including the SOC and PT.

The MedDRA version used is defined in the study data management plan.

6.6. Prior and Concomitant Medication and Procedures

The incidence of medication use will be summarized by World Health Organization (WHO) Drug Dictionary anatomic therapeutic class (ATC) Level 2 classification (i.e. therapeutic main group) and preferred name. A patient will be counted only once at each level of reporting. Prior medications are those which have been identified to have been discontinued prior to the first dose of study medication. Concomitant medications are those which have been identified to have been taken at any point during the period of time between the first dose of study drug and

last dose of study drug + 90 days, including medications which started prior to first dose of study drug that are ongoing at first dose. Subsequent medications are those which have been started after the last dose date + 90 days. The data management plan specifies the version of WHO Drug used.

Partial dates will be imputed according to Section 5.5 for the determination of prior, concomitant, and subsequent medications.

Concomitant medication use will be summarized and presented overall, by study phase, and treatment group in the SS.

A separate summary will be prepared for corticosteroids. Corticosteroids will be presented overall, by study phase and treatment group in the SS. The average daily dose of CS (prednisone equivalent) at each visit including summary statistics as well as categorical summaries for <5 mg/day, ≥5 mg/day to ≤10 mg/day, and >10 mg/day will be presented. Average daily dose is calculated as described in Section 7.2.4. Prednisone equivalence is also provided in the same section.

All prior and concomitant medication data will be provided together in a by-patient listing including the verbatim and preferred drug name and WHO ATC Class (ATC Level 2) for all randomized patients. Past biological, immunosuppressant, antimalarial and corticosteroid SLE medication history data will be listed separately for all randomized patients.

Concomitant procedures will be presented in a data listing for all randomized patients.

6.7. Study Drug Exposure and Compliance

The total dose administered (mL) and the number of doses received will be summarized as continuous variables. The number and percentage of patients that received any dose and the number and percentage of patients who received the total planned dose at each nominal, protocol specified, visit (Day 0 / Day 30 / Day 60 / Day 90 / Day 120 / Day 150 / Day 180) will be summarized. The number and percentage of patients who received 1-7 doses (mutually exclusive) will also be summarized categorically. The denominator for the number of patients dosed and the number of doses received will be the number of patients in the SS. The denominator for the number of patients who received the total planned dose will be the number of patients dosed at the corresponding nominal visit.

The percent compliance will be summarized descriptively and is defined as 100 * (sum of doses received / 7). Percent compliance will also be calculated categorically as <70%, ≥70% to <85%, ≥85% to <100%, ≥100%.

Relative dose intensity will be summarized descriptively and is defined as 100 * (total actual dose summed across all visits)/total planned dose summed across all planned visits) per the below table.

Dose Level (mg)	Dose Level (mL)*	Total Planned Dose Across All 7 Visits (mL)	Relative Dose Intensity Calculation
20	0.4	2.8	(Total actual dose in mL summed across all visits / 2.8) X 100
60	1.2	8.4	(Total actual dose in mL summed across all visits / 8.4) X 100
120	2.4	16.8	(Total actual dose in mL summed across all visits / 16.8) X 100

* Dose Level in mL is based on 50mg/mL administration.

The duration of exposure in days will also be summarized descriptively and is defined as the last dose date – first dose date +1.

Study drug exposure and compliance will be presented overall, by study phase, and treatment group in the SS.

By-patient listings of BOS161721 and placebo dosing data will be produced for all patients in the SS.

7. EFFICACY

7.1. Primary Efficacy Endpoint and Analyses

Safety will be the primary endpoint for the Phase 1b MAD portion of the study and is discussed in Section 10. Efficacy will be evaluated for the Phase 1b MAD portion of the study as an exploratory endpoint and details will be included in relation to the Phase 2 POC portion of the study within this section. Efficacy analyses will be presented by study phase and treatment group.

7.1.1. The proportion of patients with a SRI-4 response at Day 210

The primary efficacy endpoint for the POC is the proportion of patients who achieve a SRI-4 response at Day 210 in the FAS.

The SRI-4 is a composite index of SLE disease improvement that consists of scores derived from the SLEDAI-2K, the BILAG 2004 Index, and the PGA. SRI-4 response is defined as:

1) ≥ 4-point reduction from baseline in SLEDAI-2K total score

AND

2) No new severe disease activity (BILAG A organ score) or more than 1 new moderate organ score (BILAG B) compared with baseline

AND

3) No deterioration from baseline in the PGA by ≥ 30mm.

The test for treatment group difference in proportion of patients who achieve a SRI-4 response at Day 210 will be based on Pearson's chi-square test. The difference in percentages of patients who achieve SRI-4 response (BOS161721 vs. Placebo) will be presented with its corresponding (Wald) 90% CI. The number and percentage of SRI-4 responders, along with an exact (Clopper-Pearson) 90% CI, will be presented by treatment group. If an expected count for the Pearson chi-square test falls below 5, a Fisher's exact test will be used to assess the treatment difference and the CI will be based on an exact method (Santner Snell). The analyses will be repeated for each of the three components contributing to the SRI-4 response. The primary analysis of the primary efficacy endpoint will be performed in the FAS and presented by treatment group. The analysis will be repeated in the PP analysis set. SRI-4 Response at Day 210 will also be repeated by all subgroups in Section 4.2 and treatment group for POC Phase 2 only.

The superiority of BOS161721 relative to placebo will be evaluated. If the proportion of SRI-4 responders for the selected POC dose in BOS161721 arm is higher than that of the POC placebo arm, then BOS161721 will be considered superior to placebo if the 2-sided p-value is less than 0.10.

Evaluability/Non-Responders:

Medication Failures: Patients who received prohibited medications or undergo unallowable corticosteroid (CS) usage will be considered "medication failures" and will be treated as non-responders at time points on and following the first date of prohibited medication or unallowable CS usage for the primary efficacy analysis. Prohibited medications are detailed in Appendix 4 of the study protocol. Prohibited medications and unallowable CS usage will be recorded as protocol deviations in the clinical database.

A patient is considered to have unallowable CS usage if any of the following is not met:

- A maximum daily dose of 30 mg/day of oral CS will be acceptable for eligibility for the study. For patients whose only SLE treatment is steroids, their stable steroid dose must be at least 10 mg/day and no more than 30 mg/day for a minimum of 6 weeks at time of randomization on Day 0
- Topical steroids may be used, but the dose must be stable for at least 6 weeks prior to Day 0, and be maintained at a constant dose throughout the study duration until the rashes resolve. PRN topical steroids are not permitted
- Once the patient has received the first dose of study drug (Day 0), tapering of oral steroids will be highly encouraged and should be continually evaluated during the

protocol-allowed tapering windows (Day 0 through Day 150) with the target of achieving a CS daily dose of < 7.5 mg and < Day 0 dose between Day 150 and Day 210

- Between Day 150 and Day 210 (i.e., within 60 days of primary endpoint assessments), oral CS doses must be held constant
- CS Burst for SLE-related Indications
    - After Day 0 (first dose of study drug), a maximum of 1 oral CS "burst" equivalent to ≤ 40 mg/day prednisone for increased SLE disease activity will be allowed between Day 0 and Day 120, which must be tapered down to the baseline (Day 0) CS dose or lower within 14 days of initiation of the "burst". Any "burst" continuing after Day 120 or occurring after Day 120 is considered a protocol deviation.
    - Alternatively, a single intramuscular (IM) dose of methylprednisolone (< 40 mg) is permitted during this period.
- CS Burst for Non-SLE-related Indications
    - A single treatment of oral prednisone equivalent of ≤ 40 mg/day for 14 days is permitted for a non-SLE indication, though it must be completed prior to Day 120. No long acting steroid injections are permitted.
    - Note: Treatment with inhaled CS are allowed for the treatment of non-SLE-related indications only (e.g., for asthma).

Any other increase from baseline of CS dose or any other systemic use of CS of any kind (including intra-articular and intravenous administration) are not permitted from Day 0 through Day 210 and will result in the patient being considered a medication failure and will be treated as non-responders for the primary efficacy analysis. There is no restriction of CS usage after Day 210.

The determination of medication failures will be reviewed using blinded data and finalized prior to unblinding. Once a patient is determined to be a non-responder, the patient will remain a non-responder for the remainder of the study.

Out of Window Assessments:

- If a patient has an efficacy assessment (i.e. BILAG, SLEDAI-2K, PGA, CLASI, etc.) more than 30 days since the previous assessment (current visit – prior visit > 30), the corresponding result will not contribute to LOCF analyses.
- Since the Day 180 nominal visit is assigned as "Day 180/End of Treatment" in the clinical database, Day 180 nominal assessments must be >=159 days (180 – 21 day protocol allowed treatment window) since first dose to be considered for "Day 180" nominal assessments and also to be considered for LOCF to Day 210 in the case of missing Day 210 data for all efficacy analyses.

Early Treatment Withdrawals/Discontinuations:

For early treatment withdrawal/discontinuation prior to study day 159 (180 days – 21 day treatment window), patient will be considered a non-responder after the date of early treatment

withdrawal/discontinuation assuming the patient was not considered a medication failure prior to or on the date of early withdrawal/discontinuation for the primary efficacy analysis.

For early treatment withdrawal/discontinuation on or after study day 159 (180 days – 21 days) such that patient has data for nominal visit Day 180 but not for nominal visit Day 210, Day 180 data will be carried forward to Day 210 for assessment of SRI-4 response.

For early treatment withdrawal/discontinuation on or after study day 159 (180 days – 21 days) such that patient does not have data for nominal visit Day 180 and Day 210, the patient will be considered a non-responder on Day 210.

For early treatment withdrawal/discontinuation on Day 210 such that patient has data for nominal visit Day 210, the Day 210 data will be used for assessment of SRI-4 response.

Missing Data:

- **Pre-Treatment:** If a patient is missing data during the pre-treatment period for SLEDAI-2K, PGA, or BILAG, the patient will be excluded from the primary efficacy analysis.

- **Overall SRI-4 Components:** Missing data at Day 210 for overall components for SLEDAI-2K, BILAG, and PGA will be addressed by employing a LOCF analysis within each overall component as described in Section 5.4. For example, if a patient has data for SLEDAI-2K and BILAG at Day 210 but is missing PGA at Day 210, the PGA component would employ LOCF as described in Section 5.4. If the overall component is missing from Day 180 through Day 210 and the patient is not a medication failure on or prior to Day 210 and did not withdraw/discontinue treatment prior to Day 180, the SRI-4 response at Day 210 will be considered not evaluable.

- **Individual Items Contributing to Overall SRI-4 Components:** Data collection and scoring will be performed in the clinical database for BILAG, SLEDAI-2K, and PGA. If missing individual items contributing to any of the overall components occurs for Day 210, the overall component will be recalculated per the Table 1.

Table 1

SRI-4 Component	Method for Carrying Forward Individual Items Contributing to SRI-4 Component Result/Score
BILAG Grade	BILAG Grading is evaluated and entered into the clinical database by the medical monitoring team. Grades are adjudicated by independent subject matter experts. The medical monitoring team and the adjudication team will utilize BILAG index results to determine BILAG Grades. Missing data will be handled by the medical monitoring team as follows. If lab values are missing in the BILAG

	Index at Day 210, lab values from the most recent unscheduled visit <=14 days after Day 210 will be utilized for determination of the BILAG grade in the applicable organ system and entered into the clinical database. If lab values are not available from an unscheduled visit <=14 days after Day 210, laboratory data will be carried forward from the most recent laboratory assessment (scheduled or unscheduled) <=30 days prior to the Day 210 nominal visit for determination of BILAG grade. If the lab components for BILAG are missing from Day 210 and no unscheduled laboratory assessments occur within 14 days after or 30 days prior to the Day 210 assessment, the lab components from Day 180 will be utilized if Day 180 occurred within 30 days of Day 210. Otherwise, no individual BILAG Index results will be carried forward. If the BILAG grade for a given organ system is missing at Day 210, the grade will be carried forward from the Day 180 assessment. If the Day 180 BILAG grade is also missing, occurred more than 30 days before Day 210 (Day 210 visit date – Day 180 visit date > 30), or is not evaluable, the Day 210 grade will be considered not evaluable.
SLEDAI-2K	If lab values are missing at Day 210, lab values will be imputed based on the below hierarchy of determination: (1) If any unscheduled laboratory assessments occurred <=14 days after the Day 210 visit, the laboratory values at the most recent unscheduled laboratory assessment will be used for the determination of missing laboratory SLEDAI-2K criteria. For example, if platelet count is missing at Day 210 and an unscheduled laboratory assessment occurs 10 days after the Day 210 visit date, the platelet count from the unscheduled laboratory assessment would be used to determine if the patient had thrombocytopenia at Day 210 for SLEDAI-2K. (2) Else if #1 does not occur or lab values are not available at the unscheduled visit noted in #1, the laboratory data will be carried forward from the most recent laboratory assessment (scheduled or unscheduled) <=30 days prior to the Day 210 nominal visit to make the SLEDAI-2K determination. For example, if a patient is missing thrombocytopenia and leukopenia SLEDAI-2K components at Day 210, they have an unscheduled lab assessment 10 days earlier with platelet counts populated and WBC not populated, and they have a

	Day 180 assessment within 30 days of the Day 210 visit with both platelets and WBC present, the platelet value from the unscheduled visit would be carried forward to the Day 210 for thrombocytopenia component designation and the WBC value would be carried forward from the Day 180 assessment to Day 210 for the Leukopenia component designation. (3) If the lab components for SLEDAI-2K are missing from Day 210 and no unscheduled laboratory assessments occur within 14 days after or 30 days prior to the Day 210 assessment, the lab components from Day 180 will be carried forward, if available and if Day 180 occurred within 30 days of Day 210. Otherwise, no non-lab individual SLEDAI-2K components will be carried forward. If non-lab individual components are missing at Day 210, the entire SLEDAI-2K score from Day 180 will be carried forward. If Day 180 is missing, occurred more than 30 days before Day 210, or is not evaluable, LOCF is not implemented and the Day 210 score will be considered not evaluable. In the case of missing SLEDAI-2K lab components at Day 210, where any SLEDAI-2K lab component is carried forward (or back) per above, the SLEDAI-2K total score will be recalculated outside of the clinical database as the sum of: (1) All of the non-missing weights for each descriptor at Day 210 (2) All individual lab component weights carried forward from the Day 180 visit or unscheduled visit, as applicable. If any components are still missing at Day 210 after carrying forward available lab components, the Day 210 SLEDAI-2K score will not be recalculated and will be considered not evaluable. Note: Clinical input may be obtained prior to database lock and unblinding to make a final determination of SLEDAI-2K criteria in the case of carrying forward (or back) individual lab component scores.
PGA	Individual components are not applicable for PGA and will not be carried forward. Total PGA VAS measurement will be carried forward from the Day 180 visit in the case of missing PGA measurement at Day 210. If total PGA score is also missing from Day 180, occurred more than 30 days before Day 210, or is not evaluable, LOCF is not

	implemented and the Day 210 visit will be considered not evaluable.

Table 2 summarizes the primary analysis and details several sensitivity analyses which may be performed:

Table 2

Analysis	**Analysis Population**	**Criteria**
Primary Analysis LOCF for missing data, Medication Failures=Non-Responder	FAS	The primary analysis of the primary efficacy endpoint will be performed based on LOCF methods for missing data. (1) Medication failures are considered non-responders on and after the date of medication failure (2) Early withdrawals/discontinuations prior to study day 159 will be considered non-responders after the date of early withdrawal/discontinuation a. If patient is a medication failure prior to withdrawal/discontinuation, patient will be considered a non-responder on and after the date of medication failure. b. In case of early withdrawal/discontinuation after study day 159 but before Day 210 such that patient has Day 180 data but is lacking Day 210 data in a patient without medication failure, the Day 180 data will be carried forward to Day 210. (3) Patients with missing Day 180 and Day 210 data without prior medication failure and without early withdrawal/discontinuation prior to study day 159 are considered not evaluable and are excluded from analysis
Sensitivity Analysis 1: Observed (no LOCF for missing data), Medication Failures=Non-Responder	FAS	A sensitivity analysis of the primary efficacy endpoint may be performed based on purely observed data. No LOCF imputations will be performed for missing data. Additional criteria are as follows: (1) Medication failures are considered non-responders on and after the date of medication failure (2) Early withdrawals/discontinuations prior to Day 210 will be considered non-responders after the date of early withdrawal/discontinuation a. If patient is a medication failure prior to withdrawal/discontinuation, patient will be

		considered a non-responder on and after the date of medication failure. (3) Patients with missing Day 210 data without prior medication failure and without early withdrawal/discontinuation prior to Day 210 are considered not evaluable and are excluded from the sensitivity analysis If lab component scores are carried forward from Day 180 to Day 210 in order to assign a Day 210 grade in the primary analysis, the Day 210 assessment will be considered not evaluable for this sensitivity analysis.
Sensitivity Analysis 2: LOCF for missing data, Exclude Medication Failures	FAS	A sensitivity analysis of the primary efficacy endpoint may be performed based on LOCF methods for missing data detailed for the primary efficacy analysis. Additional/modified criteria are as follows: (1) Medication failures are considered not evaluable and are excluded from the sensitivity analysis on and after the date of medication failure (2) Early withdrawals/discontinuations prior to study day 159 will be considered not evaluable and will be excluded from analysis after the date of early withdrawal/discontinuation a. If patient is a medication failure prior to withdrawal/discontinuation, patient will be considered not evaluable and excluded on and after the date of medication failure. b. In case of early withdrawal/discontinuation after study day 159 but before Day 210 such that patient has Day 180 data but is lacking Day 210 data in a patient without medication failure, the Day 180 data will be carried forward to Day 210. (3) Patients with missing Day 180 and Day 210 data are considered not evaluable and are excluded from the sensitivity analysis
Sensitivity Analysis 3: Observed (no LOCF for missing data), Exclude Medication Failures	FAS	A sensitivity analysis of the primary efficacy endpoint may be performed based on purely observed data. No LOCF imputations will be performed for missing data. Additional criteria are as follows: (1) Medication failures are considered not evaluable and are excluded from the sensitivity analysis on and after the date of medication failure (2) Early withdrawals/discontinuations prior to Day 210 will be considered not evaluable and are excluded from the sensitivity analysis after the date of early withdrawal/discontinuation

		a. If patient is a medication failure prior to withdrawal/discontinuation, patient will be considered not evaluable and excluded on and after the date of medication failure. (3) Patients with missing Day 210 data are considered not evaluable and are excluded from the sensitivity analysis If lab component scores are carried forward from Day 180 to Day 210 in order to assign a Day 210 grade in the primary analysis, the Day 210 assessment will be considered not evaluable for this sensitivity analysis.
Sensitivity Analysis 4: LOCF for missing data, Medication Failures=Non-Responder, Logistic Regression	FAS	CCI [redacted]

Additional sensitivity analyses may be performed to consider the effect of missing data.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

A summary of disposition of SRI-4 response per the primary analysis at day 210 will be presented by study phase and treatment group in the FAS and PP. Sensitivity analysis 2 (LOCF, Exclude Medication Failures) will also be performed for disposition of SRI-4 response at day 210. These summaries will include the number of SRI-4 responders and the number of SRI-4 non-responders with a breakdown of reasons for non-response (early withdrawal, medication failure, <4-point reduction from baseline in SLEDAI-2K total score, new BILAG 1A/2B, and PGA worsening). Denominators for percentages will be the number of patients eligible for SRI-4 response analysis (patients not deemed not evaluable) defined as those with data for PGA, SLEDAI-2K and BILAG at Day 210 (whether or not due to LOCF, FAS analysis).

A listing of SRI Response data at Days 30, 60, 90, 120, 150, 180 and 210 will be presented in the FAS. A listing of SRI-4 Response at Day 210 (Primary and Sensitivity Analyses) will also be presented in the FAS. A listing of patients with medication failures, early withdrawals/discontinuations and those excluded from the primary efficacy analysis will also be provided for all randomized patients.

7.2. Secondary Efficacy Endpoints and Analyses

PK and PD measures are the secondary endpoints for the Phase 1b MAD part of the study and exploratory endpoints for the POC part of the study. Those analyses are described in Section 8.

7.2.1. Proportion of Patients with SRI-4 Response at Each Visit

The primary analysis methods described in Section 7.1.1 and Table 2 will be repeated for each nominal visit and will be presented by study phase and treatment group in the FAS. Other analyses described in Section 7.1.1 may also be performed. The 3 components will not be summarized separately by visit. In case of missing data, LOCF methods will be implemented as described in section 5.4 and Table 1. LOCF will only be implemented with the most recent prior nominal visit. If the most recent prior nominal visit is missing, occurred more than 30 days before, or is not evaluable, LOCF is not implemented. For example, if Day 30 SLEDAI assessment is complete but Day 60 and Day 90 SLEDAI assessments are missing, Day 60 will implement LOCF from Day 30 assuming Day 30 is within 30 days of Day 60 and Day 60 SLEDAI will be included in the analysis but Day 90 will not carry forward from Day 60 and thus Day 90 will be considered not evaluable and will be excluded from the analysis. Similarly, if Day 150 assessments are complete but Day 180 and Day 210 are missing, Day 180 will implement LOCF from Day 150 assuming Day 150 is within 30 days of Day 180 and will be included in the analysis but Day 210 will not carry forward from Day 180, will be considered not evaluable, and will be excluded from the analysis.

If a patient is a medication failure, the patient is considered to be a non-responder at timepoints on or after the first date of prohibited medication/unallowable CS usage.

If a patient discontinues/withdraws early, the patient will be considered a non-responder after the date of discontinuation/withdrawal (unless the patient had a prior medication failure in which case the patient is considered a non-responder on and after the date of medication failure).

The denominators for analyses will be the number of patients in the FAS with an evaluable post-baseline SRI-4 assessment defined as those with data for PGA, SLEDAI-2K and BILAG (whether observed or imputed via LOCF depending on the analysis) at a given visit.

A graphical display of SRI-4 response by study phase, treatment group, and visit in the FAS will be presented.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.2. Proportion of Patients with SRI-5 Response at Each Visit

The SRI-5 is a composite index of SLE disease improvement that consists of scores derived from the SLEDAI-2K, the BILAG 2004 Index and the PGA. Response based on the SRI-5 is defined by:

1) ≥ 5-point reduction from baseline in SLEDAI-2K total score

AND

2) No new severe disease activity (BILAG A organ score) or more than 1 new moderate organ score (BILAG B) compared with baseline

AND

3) No deterioration from baseline in the PGA by ≥ 30mm.

The primary analysis described in Section 7.1.1 and Table 2 will be repeated for each nominal visit as described in Section 7.2.1 based on the SRI-5 response definition and will be presented by study phase and treatment group in the FAS. The 3 components will not be summarized separately by visit for SRI-5 response. Other analyses described in Section 7.1.1 may also be performed.

The denominators for analyses will be the number of patients in the FAS with an evaluable post-baseline SRI-5 assessment defined as those with data for PGA, SLEDAI-2K and BILAG (whether observed or imputed via LOCF depending on the analysis) at a given visit.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.3. Proportion of Patients with SRI-6 Response at Each Visit

The SRI-6 is a composite index of SLE disease improvement that consists of scores derived from the SLEDAI-2K, the BILAG 2004 Index and the PGA. Response based on the SRI-6 is defined by:

1) ≥ 6-point reduction from baseline in SLEDAI-2K total score

AND

2) No new severe disease activity (BILAG A organ score) or more than 1 new moderate organ score (BILAG B) compared with baseline

AND

3) No deterioration from baseline in the PGA by ≥ 30mm.

The primary analysis described in Section 7.1.1 and Table 2 will be repeated for each nominal visit as described in Section 7.2.1 based on the SRI-6 response definition and will be presented by study phase and treatment group in the FAS. The 3 components will not be summarized separately by visit for SRI-6 response. Other analyses described in Section 7.1.1 may also be performed.

The denominators for analyses will be the number of patients in the FAS with an evaluable post-baseline SRI-6 assessment defined as those with data for PGA, SLEDAI-2K and BILAG (whether observed or imputed via LOCF depending on the analysis) at a given visit.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.4. Oral Corticosteroid (CS) Reduction

The average daily CS dose will be presented over time starting with Day 0 as baseline. Oral CS usage for all visits will be based on the average daily dose, defined as the sum of the corticosteroid doses taken each day from the date of the previous visit to the day before the date of the current visit (converted to prednisone equivalent dose) / the duration between the two visits. For example, the average daily dose at Day 0 (Baseline) is the sum of daily doses of CS from the screening visit to the day prior to Day 0 divided by the duration between the screening visit to the day prior to Day 0, inclusive. Similarly, the average daily dose at Day 30 is the sum of daily doses of CS from Day 0 to the day prior to Day 30 divided by the duration between Day 0 to the day prior to Day 30, inclusive. Partial dates for CS use will be imputed based on details in Section 5.5. Oral CS will be converted to prednisone dose equivalence for analysis of reduction of oral CS use. Common prednisone dose equivalents are provided in Table 3.

Table 3

Medication	**Prednisone Dose Equivalence**
Prednisone	1 mg
Cortisone	5 mg
Hydrocortisone	4 mg
Prednisolone	1 mg
Methylprednisolone	0.8 mg
Triamcinolone	0.8 mg
Budesonide	0.25 mg
Dexamethasone	0.16 mg
Bethamethasone	0.16 mg

7.2.4.1. Proportion of Patients with a sustained reduction of oral CS between Day 120 and Day 210 (MAD)

The proportion of patients with a sustained reduction of oral corticosteroid, defined as ≤ 10 mg/day prednisone equivalent and < Day 0 dose, between Day 120 and Day 210 determined based on actual daily CS usage on and in between the nominal visit dates will be summarized for MAD Phase 1b by treatment group in the FAS. To qualify as sustained reduction, the actual daily CS usage needs to meet the definition above every day between nominal visit dates for Day 120 and Day 210 inclusive. LOCF and imputation for missing values will not be performed. Otherwise, statistical methods described in Section 7.1.1 will be implemented. Patients who withdraw prior to Day 210 or are designated as a medication failure on or prior to Day 210 or miss the Day 120 and/or the day 210 visit will not be considered to have sustained reduction in CS. Denominators for percentages are the number of patients taking oral CS at baseline.

7.2.4.2. Proportion of Patients with a sustained reduction of CS between Day 150 and Day 210 (POC)

The proportion of patients with a sustained reduction of oral corticosteroid, defined as ≤ 7.5 mg/day prednisone equivalent and < Day 0 dose, between Day 150 and Day 210 determined based on actual daily CS usage on and in between the nominal visit dates will be summarized for POC Phase 2 by treatment group in the FAS. To qualify as sustained reduction, the actual daily CS usage needs to meet the definition above every day between nominal visit dates for Day 150 and Day 210 inclusive. LOCF and imputation for missing values will not be performed. Otherwise, statistical methods described in Section 7.1.1 will be implemented. Patients who withdraw prior to Day 210 or are designated as a medication failure on or prior to Day 210 or miss the Day 150 and/or Day 210 visit will not be considered to have sustained reduction in CS. Denominators for percentages are the number of patients taking oral CS at baseline.

7.2.4.3. Percent Reduction in CS Administration from Day 0 through Day 210 (POC)

The percent reduction in CS administration from Day 0 through Day 210 determined based on the average daily CS usage will be summarized for POC Phase 2 by treatment group and visit. Average daily dose will be calculated for nominal visits as defined in Section 7.2.4. This will be performed for patients in the FAS who received ≥ 7.5 mg/day prednisone equivalent at Day 0. Patients considered medication failures will be excluded from the percent reduction analysis at timepoints on and after the first date of medication failure.

7.2.5. Proportion of new or recurrent BILAG A flare or >1 BILAG B flare through Day 210

The proportion of patients with a new or recurrent BILAG A flare or >1 BILAG B flares through Day 210 will be summarized by study phase and treatment group in the FAS. Statistical methods described in Section 7.1.1 will be implemented. BILAG flares will be adjudicated, recorded in the clinical database, and summarized as recorded in the clinical database. No LOCF methods will be implemented. When discrepancies between the two-reviewers occurs, the adjudicated flare determination will be used for analyses. The proportion of new or recurrent BILAG A flare or >1 BILAG B flare will be summarized overall and by nominal visit. The denominator for percentages of patients with a new or recurrent flare will be patients in the FAS with a post-baseline BILAG assessment at the given visit.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

Results as recorded in the clinical database will be presented in separate data listings for BILAG SLE history, BILAG-2004 index, and BILAG grading for all randomized patients.

7.2.6. Physician's Global Assessment (PGA)

The PGA is used to assess investigator's general impression on the patient's overall status of SLE disease activity via visual analogue scale (100 mm) with 0 being "very good, asymptomatic and no limitation of normal activities" with 100 mm being "most severe possible disease ever seen in all SLE patients."

7.2.6.1. Proportion of Patients with PGA Worsening at Each Visit

The proportion of patients with PGA worsening, defined as an increase of >30mm from baseline, will be summarized by study phase and treatment group in the FAS overall (at any post-baseline nominal visit) and by nominal visit. Statistical methods described in Section 7.1.1 will be implemented. In case of missing data, LOCF methods will be implemented as described in section 5.4 and Table 1 and in Section 7.2.1 for the "at each visit" analyses.

The analyses described in Section 7.1.1 and Table 2 (Primary Analysis methods) will be repeated for each nominal visit as described in Section 7.2.1 based on the PGA worsening definition and will be presented by study phase and treatment group in the FAS.

The denominators for analyses will be the number of patients in the FAS with an evaluable post-baseline PGA assessment (overall analysis) or with evaluable PGA assessment (whether or not imputed via LOCF depending on the analysis) at a given visit (nominal visit analysis).

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.6.2. Results, Changes, and Percentage Change from Baseline in PGA

The analysis of PGA will be performed separately using an ANCOVA model with change or percentage change from baseline to each nominal visit as the dependent variable, treatment group as the factor, and the corresponding baseline PGA result as a covariate. The p-value for the difference between the BOS161721 and placebo groups at each nominal visit is based on the F-test.

Descriptive statistics for PGA will be summarized separately at baseline and each visit. Change and percentage change from baseline will be summarized at each post-baseline visit. The least squares (LS) mean estimate, its standard error and 90% CI will be tabulated by treatment group and for the difference between treatment groups (BOS161721 versus placebo) for change and percentage change at each visit. In case of missing data, LOCF methods will be implemented as described in Section 5.4 and Table 1 and in Section 7.2.1.

Two separate approaches will be taken based on handling for medication failures and treatment withdrawals:

Approach 1: Corresponding to Primary Analysis Methods (Medication Failure LOCF): If a patient is a medication failure, the most recent assessment prior to the medication failure date will be carried forward to all subsequent assessments. In the case of patients who

withdraw/discontinue early without prior medication failure, the most recent data on or prior to withdrawal/discontinuation will be carried forward to all subsequent assessments. In case of early withdrawal/discontinuation after Day 180 but before Day 210 such that patient has Day 180 data but is lacking Day 210 data in a patient without medication failure, the Day 180 data will be carried forward to Day 210.

Approach 2: Corresponding to Sensitivity Analysis 2 Methods (Exclude Medication Failure): If a patient is a medication failure, all assessments on and after medication failure date will be considered not evaluable and will be excluded. In the case of patients who withdraw/discontinue early without prior medication failure, all assessments after the early withdrawal date will be considered not evaluable and will be excluded. In case of early withdrawal/discontinuation after Day 180 but before Day 210 such that patient has Day 180 data but is lacking Day 210 data in a patient without medication failure, the Day 180 data will be carried forward to Day 210.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

PGA observed results will be presented in a data listing for all randomized patients.

7.2.7. Proportion of Patients with BICLA Response

The BICLA is a responder index developed to measure response to therapy, and it includes scores from the BILAG, SLEDAI-2K, and PGA. BICLA response is defined as:

1) at least 1 gradation of improvement in baseline BILAG 2004 scores in all body systems with moderate disease activity at entry (e.g., all A [severe disese] score falling to B [moderate], C [mild] or D [no activity] or all B [moderate disease] scores falling to C [mild], or D [no activity])

AND

2) no new BILAG A or more than 1 new BILAG B scores

AND

3) no worsening of total SLEDAI-2K score from baseline

AND

4) ≤ 10% deterioration in PGA score

AND

5) no medication failure.

Statistical methods described in Section 7.1.1 will be implemented. In case of missing data, LOCF methods will be implemented as described in section 5.4 and Table 1 and in Section 7.2.1 for the “at each visit” analysis. The primary analysis and sensitivity analysis 2 (LOCF for

missing data, Exclude Medication Failures) methods described in Section 7.1.1 and Table 2 will be repeated overall and as described in Section 7.2.1 for Days 210, 240, and 270 based on the BICLA response definition and will be presented by study phase and treatment group in the FAS.

Other analyses described in Section 7.1.1 may also be performed.

The denominators for analyses will be the number of patients in the FAS with an evaluable post-baseline BICLA assessment defined as those with data for PGA, SLEDAI-2K and BILAG (whether observed or imputed via LOCF depending on the analysis) (overall analysis) or at a given visit (nominal visit analysis).

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.8. Cutaneous Lupus Erythematosus Area and Severity Index (CLASI)

The CLASI is a comprehensive tool for assessment of disease activity and damage in cutaneous lupus, shown to be valid, reliable, and sensitive to changes in disease activity.

The total activity score will be calculated by summing the scores on the left side of the form (erythema, scale/hypertrophy, mucous membrane, alopecia – recent hair loss and alopecia clinically not obviously scarred).

The total damage score will be calculated by:

(1) Sum the responses for “Dyspigmentation” in all anatomical locations

(2) If a response for “Report duration of dyspigmentation after active lesions have resolved” is “Dyspigmentation usually lasts at least 12 months”, then multiply the sum in #1 by 2.

(3) If a response for “Report duration of dyspigmentation after active lesions have resolved” is “Dyspigmentation usually lasts less than 12 months”, then the sum in #1 will not change.

(4) If the sum in #1 is 0 and a response for “Report duration of dyspigmentation after active lesions have resolved” is missing, then the sum in #1 will not change.

(5) Determine the Dyspigmentation Score based on the results from 2, 3, or 4.

(6) Sum the remaining scores on the right side of the form (scarring/atrophy/panniculitis, alopecia - scarring of the scalp judged clinically).

(7) Sum the scores from #5 and #6 to obtain the total damage score.

If any individual scores which contribute to the total activity score or the total damage score are missing, the total score in question will not be calculated. If total activity score or total damage score are missing at a visit, LOCF methods will be implemented as described in Section 5.4 and Section 7.2.1 for the “at each visit” analyses.

The mean and SD of the change from baseline in CLASI total activity and total damage scores will be presented graphically over time by study phase and treatment group.

7.2.8.1. Proportion of Patients with CLASI Response

CLASI response is defined as 50% improvement from baseline in CLASI. A CLASI-A response applies to "A" (total activity) and a CLASI-B response applies to "B" (total damage) score.

Statistical methods described in Section 7.1.1 will be implemented. In case of missing data, LOCF methods will be implemented as described in Section 5.4 and 7.2.1 for the "at each visit" analysis. Individual components of the CLASI will not be carried forward; however, the total activity score or total damage score may be carried forward from the most recent prior visit. If the most recent prior visit is missing, occurred more than 30 days before (current visit date – most recent prior visit date > 30), or is not evaluable, LOCF is not implemented. The primary analysis methods (LOCF for missing data, medication failure=non-responder) as described in Section 7.1.1 and Table 2 will be performed at each visit based on the CLASI response definition and will be presented by study phase and treatment group in the FAS.

CCI [redacted]

The analysis will also be repeated in line with Sensitivity Analysis 2 (LOCF for missing data, exclude medication failures) in Section 7.1.1.

Other analyses described in Section 7.1.1 may also be performed.

The denominators for analyses will be the number of patients in the FAS with an evaluable post-baseline CLASI assessment (whether observed or imputed via LOCF depending on the analysis) at a given visit.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.8.2. Results, Changes, and Percentage Change from Baseline in CLASI

The analysis of CLASI total activity score and CLASI total damage score will be performed separately using an ANCOVA model with change or percentage change from baseline to each nominal visit as the dependent variable, treatment group as the factor, and the corresponding baseline CLASI score as a covariate. The p-value for the difference between the BOS161721 and placebo groups at each nominal visit is based on the F-test.

Descriptive statistics for the activity score and damage score will be summarized separately at baseline and each visit. Change and percentage change from baseline will be summarized at each post-baseline visit. The LS Mean estimate, its standard error and 90% CI will be tabulated by treatment group and for the difference between treatment groups (BOS161721 versus placebo) for change and percentage change at each visit. Statistical methods described in Section 7.1.1 will be implemented. In case of missing data, LOCF methods will be implemented

as described in Section 5.4 and Section 7.2.1 for the "at each visit" analyses. Individual components of the CLASI will not be carried forward; however, the total activity score or total damage score may be carried forward from the most recent prior visit. If the most recent prior visit is missing, occurred more than 30 days before (current visit date – most recent prior visit date > 30), or is not evaluable, LOCF is not implemented. The analyses will be performed in line with the two approaches outlined in Section 7.2.6.2. CCI [redacted]

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

CLASI observed results will be presented in data listings for all randomized patients.

7.2.9. Results, Changes, and Percentage Changes from Baseline in Swollen, Tender, and Active Joints ACR-28

The ACR-28 joint count evaluates the number of tender and swollen joints in the shoulder, elbow, wrist, hand, knee joints. Joints of the feet are excluded. The sum of active joints is also evaluated and is defined as the total number of joints with swelling and tenderness in the same joint. The joint sums for tender and swollen joints are recorded in the database. The sum of active joints will be derived.

The analysis of ACR-28 Swollen and Tender Joints for the sum of tenderness (left and right), the sum of swelling (left and right), and sum of active joints (left and right) will be performed separately using an ANCOVA model with change and percentage change from baseline to each nominal visit as the dependent variable, treatment group as the factor, and the baseline ACR-28 score as a covariate. The p-value for the difference between the BOS161721 and placebo groups at each nominal visit is based on the F-test.

Descriptive statistics for the tenderness, swelling, and derived active joint sums will be summarized separately at baseline and each visit. Change and percentage change from baseline will be summarized at each post-baseline visit. The LS Mean estimate, its standard error and 90% CI will be tabulated by treatment group and for the difference between treatment groups (BOS161721 versus placebo) for change and percentage change at each visit.

Statistical methods described in Section 7.1.1 will be implemented. In case of missing data, LOCF methods will be implemented as described in Section 5.4 and in Section 7.2.1 for the "at each visit" analyses. Individual components of ACR-28 will not be carried forward; however, the total ACR-28 score may be carried forward from the most recent prior visit. If the most recent prior visit is missing, occurred more than 30 days before (current visit date – most recent prior visit date > 30), or is not evaluable, LOCF is not implemented. The analyses will be performed in line with the two approaches outlined in Section 7.2.6.1.

The mean and SD of the change from baseline in ACR-28 sum of tenderness, sum of swelling, and sum of active joints based on primary analysis methods will be presented graphically over time by study phase and treatment group.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

Swollen, tender, and active joints ACR-28 observed results will be presented in data listings for all randomized patients.

7.2.10. Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K)

The SLEDAI-2K is a validated instrument that measures disease activity in SLE patients at the time of the visit and in the previous 30 days. It is a global index and includes 24 clinical and laboratory variables that are weighted by the type of manifestation, but not by severity. The total score falls between 0 and 105, with higher scores representing increased disease activity. The SLEDAI-2K has been shown to be a valid and reliable disease activity measure in multiple patient groups. A SLEDAI 2K of 6 or more generally represents moderately to severely active disease.

7.2.10.1. Proportion of Patients with a SLEDAI-2K Response

The proportion of patients with a SLEDAI-2K response, defined as ≥ 4-point reduction from baseline in SLEDAI-2K total score, will be summarized by study phase and treatment group in the FAS overall and by nominal visit. Statistical methods described in Section 7.1.1 will be implemented. In case of missing data, LOCF methods will be implemented as described in section 5.4 and Table 1 and in Section 7.2.1 for the "at each visit" analyses.

The analyses described in Section 7.1.1 and Table 2 (Primary Analysis methods and Sensitivity Analysis 2 methods) will be repeated for each nominal visit as described in Section 7.2.1 based on the SLEDAI-2K response definition and will be presented by study phase and treatment group in the FAS.

The denominators for analyses will be the number of patients in the FAS with an evaluable post-baseline SLEDAI-2K assessment (overall analysis) or with evaluable SLEDAI-2K assessment (whether observed or imputed via LOCF depending on the analysis) at a given visit (nominal visit analysis).

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.10.2. Results, Changes, and Percentage Change from Baseline in SLEDAI-2K

The analysis of SLEDAI-2K will be performed using an ANCOVA model with change and percent change from baseline to each nominal visit as the dependent variable, treatment group as the factor, and the baseline SLEDAI-2K score as a covariate. The SLEDAI-2K total score is calculated and provided in the clinical database. In case of missing data, SLEDAI-2K will be re-calculated outside the clinical database after application of imputation rules.

Descriptive statistics for the SLEDAI-2K total score will be summarized separately at baseline and each visit. Change and percentage change from baseline will be summarized at each post-

baseline visit. The LS mean estimate, its standard error and 90% CI will be tabulated by treatment group and for the difference between treatment groups (BOS161721 versus placebo) for change and percentage change at each visit. The p-value for the difference between the BOS161721 and placebo groups at each nominal visit is based on the F-test.

Statistical methods described in Section 7.1.1 will be implemented,. In case of missing data, LOCF methods will be implemented as described in section 5.4 and Table 1 and in Section 7.2.1 for the "at each visit" analyses. The analyses will be performed in line with the two approaches outlined in Section 7.2.6.1.

Approach 1 corresponding to the primary analysis methods (LOCF for missing data, medication failure = non-responder) will be repeated for subgroups CCI as defined below:

- CCI

7.2.11. Results, Changes, and Percentage Change from Baseline in SLICC/ACR Damage Index

The SLICC/ACR damage index is a validated instrument to assess damage, defined as irreversible impairment, continuously persistent for 6 months (ascertained by clinical assessment), occurring since the onset of lupus, and it is based on a weighted scoring system. This index records damage occurring in patients with SLE regardless of cause, with

demonstrated content, face, criterion, and discriminant validity. It will be performed on Days 0 and 180.

The analysis of SLICC/ACR Damage Index total score will be performed using an ANCOVA model with change and percentage change from baseline to Day 180 as the dependent variable, treatment group as the factor, and the baseline SLICC/ACR Damage Index total score as a covariate. The SLICC/ACR Damage Index total score is calculated and provided in the clinical database. A missing total SLICC/ACR Damage Index score will not be imputed. Observed results will be presented for the change and percentage change from baseline analysis. The p-value for the difference between the BOS161721 and placebo groups at each nominal visit is based on the F-test.

Descriptive statistics for the SLICC/ACR Damage Index total score will be summarized separately at baseline and Day 180. Change and percentage change from baseline will be summarized at Day 180. The LS Mean estimate, its standard error and 90% CI will be tabulated by treatment group and for the difference between treatment groups (BOS161721 versus placebo) for change and percentage change at Day 180.

In case of medication failure or early withdrawal/discontinuation prior to Day 180, the Day 180 assessment will be considered not evaluable and will not contribute to the analysis at that visit.

These analyses will also be performed for the MAD Phase 1b with the exception of the statistical testing.

Observed SLICC/ACR Damage Index data will be presented in a by-patient listing for all randomized patients. SLICC Criteria for SLE at screening will also be presented in a by-patient listing for all randomized patients.

7.2.12. Medication Failure

7.2.12.1. Proportion of Patients with Medication Failures

Patients who received prohibited medications or had unallowable CS usage as described in Section 4.6 of the study protocol will be considered “medication failures” and will be treated as non-responders at time points on and following the first date of prohibited medication or unallowable CS usage for the primary efficacy analysis and selected sensitivity and secondary efficacy analyses. The determination of medication failures and medication failure dates will be reviewed using blinded data and finalized prior to unblinding.

The number and proportion of patients with medication failures will be summarized by study phase and treatment group in the FAS overall and by visit. A patient will be counted as a medication failure at each visit occurring on and after the date in which a patient is considered a medication failure. Statistical methods described in Section 7.1.1 will be implemented for observed data only. LOCF and missing data imputation will not be performed.

The denominator is the number of patients in FAS (overall analysis) and the number of patients with non-missing visit data (for the by visit analysis).

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.12.2. Time to Medication Failure (TTMF)

Medication failure is defined in Section 4.6 of the study protocol and discussed in Section 7.1.1. TTMF will be computed as the event date (as described below) – randomization date + 1.

Kaplan-Meier methods will be used to estimate TTMF for each treatment group. Estimates of median TTMF will be provided along with 90% confidence intervals. Q1 and Q3 will also be tabulated. Comparisons between treatment groups will use the two-sided log-rank test. In addition, hazard ratio (HR) estimates with 90% CIs will be calculated from a Cox proportional hazards regression analysis that includes treatment as a factor. Censoring rules are defined as in Table 4.

Table 4

Situation	**Date of Event or Censoring**	**Outcome**
Patient is designated a medication failure (received prohibited meds or unallowable CS burst)	Date of first dose of medication leading to medication failure designation. These dates will be obtained from the date of deviation on the protocol deviation eCRF form.	Event
Patient withdraws from study (no medication failure)	Censored at date of withdrawal	Censored
Patient completes the study (no medication failure)	Censored at last study visit	Censored
Patient dies during study (no medication failure)	Censored at the date of death	Censored

Kaplan-Meier curves of TTMF will be plotted over time.

These analyses will also be performed for the MAD Phase 1b with the exception of the statistical testing.

Time to medication failure data will be provided in a by-patient listing in the full analysis set.

7.2.13. Duration of Longest SRI-4 Response

Duration of longest SRI-4 response will be computed for patients who have been identified as a responder at least once based on criteria defined in Section 7.1.1. The duration of longest response will be defined as the longest period a patient meets the SRI-4 responder criteria at consecutive visits. This will be computed as the date of last consecutive SRI-4 response – date of first consecutive response + 1.

The duration of longest SRI-4 response will be analyzed using an analysis of variance (ANOVA) model with treatment group as the effect. The p-value for the difference between the BOS161721 and placebo group is based on the F-test.

The mean, SD, median, minimum, maximum, Q3, Q4, LS Mean estimate, its standard error and 90% CI will be tabulated by treatment group and for the difference between treatment groups (BOS161721 versus placebo).

If a patient has missed a visit between two scheduled visits with a response at both, the patient will be counted as having a response at the missed visit unless otherwise specified.

Duration of longest SRI-4 response will be displayed in the SRI response listing.

These analyses will also be performed for the MAD Phase 1b with the exception of the statistical testing.

7.2.14. Time to First BILAG A Flare or >1 BILAG B Flare through Day 210

Time to first BILAG A Flare or >1 BILAG B Flare will be computed as the event date (as described below) – randomization date + 1. If a patient is designated as a medication failure prior to or on their day 210 visit, they will be considered as having a BILAG flare (event) as of the date of first dose of the medication leading to the medication failure designation due to prohibited medication or unallowable CS burst.

Time to first BILAG A Flare or >1 BILAG B Flare will be analyzed with Kaplan-Meier methods described for TTMF and censoring as described in Table 5. Kaplan-Meier curves will be plotted over time.

Table 5

Situation	**Date of Event or Censoring**	**Outcome**
Patient has a BILAG A Flare or >1 BILAG B Flare (Noted as BILAG 1A/2B Flare) prior to or on their Day 210 visit	Date of first BILAG 1A/2B Flare	Event

Patient is designated a medication failure (received prohibited meds or unallowable CS burst) prior to or on their Day 210 visit	Date of first dose of medication leading to medication failure designation. These dates will be obtained from the date of deviation on the protocol deviation eCRF form.	Event
Patient withdraws from study (no BILAG 1A/2B Flare, no medication failure) prior to or on their Day 210 visit	Censored at date of last non-missing BILAG flare assessment (prior to or on withdrawal) through Day 210	Censored
Patient completes the study (no BILAG 1A/2B Flare, no medication failure)	Censored at date of last non-missing BILAG flare assessment through Day 210	Censored
Patient dies during study (no BILAG 1A/2B Flare, no medication failure) prior to or on their Day 210 visit	Censored at the day of death	Censored

An event will be counted once in a patient who is a medication failure and also experiences a BILAG flare. Date of event will be the date of the earlier of the two reasons.

These analyses will also be performed for the MAD Phase 1b with the exception of the statistical testing.

Time to first BILAG A Flare or >1 BILAG B Flare data will be provided in a by-patient listing in the full analysis set.

7.3. Exploratory Endpoints and Analyses

Efficacy will be evaluated for the Phase 1b MAD portion of the study as an exploratory endpoints and details are noted throughout sections 7.1 and 7.2. Exploratory endpoints and analyses may be included in relation to the Phase 2 POC part of the study as described within this section.

7.3.1. CCI [redacted]

CCI [redacted]

CCI [redacted]

CCI [redacted]

CCI [redacted]

8. PHARMACOKINETICS/PHARMACODYNAMICS/ IMMUNOGENICITY

Pharmacokinetics, pharmacodynamics, and immunogenicity are secondary endpoints for the Phase 1b MAD study and exploratory endpoints for the Phase 2 POC study.

PD results reported as less than the lower limit of quantification (LLOQ) will be counted as 0 to calculate summary linear statistics and will be reported as <LLOQ in listings. All observed Immunogenicity and PD data, as provided by an external vendor, will be presented in by-patient listings.

SAPs describing analyses (pharmacokinetics [PK], PK/pharmacodynamics[PD], etc.) not covered within this SAP will be included within an Appendix of this document. The author, finalization date, signatories, and page numbering of this SAP are independent of those SAPs and vice versa. Only the last final versions will be included in the appendix just prior to study conclusion.

8.1. Pharmacokinetics

BOS161721 PK concentration and parameter analyses will performed by a separate PK vendor and described under a separate SAP.

8.2. Pharmacodynamics

Descriptive statistics for the observed results of the following PD endpoints/results will be summarized separately at baseline and each visit. Change from baseline and percentage change from baseline will be summarized at each post-baseline visit. In cases where both numeric and qualitative results are provided, the summaries will be based on the numeric results. In cases where only qualitative results are provided and results given only as titers, shifts will be summarized. Graphical presentations of mean and SD of the change from baseline over time will also be presented for the parameters in table 6.

Table 6

PD Endpoint	**Applicable Study Portion(s)**	**Parameters to be Included in Summary Tables and Figures ('NA (all)' is displayed if all collected parameters are to be included)**
Clinical PD Biomarkers		
CCI	MAD and POC	NA (all)
C-Reactive Protein	MAD and POC	NA (all)
CCI	MAD and POC	NA (all)
CCI	MAD and POC	NA (all)
CCI	MAD and POC	NA (all)
CCI	MAD and POC	NA (all)
CCI	MAD and POC	NA (all)
Other Biomarkers		
Phosphorylated signal transducer and activator of transcription 3 (pSTAT3)	MAD only	% pSTAT3+ Lymphocytes –Stimulated; Ratio MEF (Stimulated / Unstimulated)

PD Endpoint	**Applicable Study Portion(s)**	**Parameters to be Included in Summary Tables and Figures ('NA (all)' is displayed if all collected parameters are to be included)**
Leukocyte immunophenotype (B and NK Cells)	MAD only	CD19+% of CD45+ Lymphocytes; IgD+C27-% of CD19+; IgD+CD27-CD38++CD24++% of CD19+; IgD-CD27-% of CD19+; IgD+CD27+% of CD19+; IgD-CD27+% of CD19+; IgD-CD27+CD38++CD138-% of CD19+; IgD-CD27+CD38++CD138+% of CD19+; CD56+% OF CD45+ Lymphocytes
Leukocyte immunophenotype (T-Cells)	MAD only	CD4+CD8-% of CD3+; CXCR5+% of CD4+CD8-; PD-1+% of CD4+CD8-; ICOS+% of CD4+CD8-; CD25+CD127-% of CD4+CD8-; CXCR5+% of CD25+CD127-; PD-1+% of CD25+CD127-; ICOS+% of CD25+CD127-; CD8+CD4-% of CD3+
CCI	MAD only	CCI
CCI	MAD and POC	CCI
CCI	MAD and POC	Analyses planned for available CCI data will be described in a separate SAP to be included in the appendix of this SAP.

CCI [redacted]

CCI [redacted]

Additional PD results may be plotted over time by treatment group and visit.

8.3. Immunogenicity

Immunogenicity data as collected on the eCRF will be provided in by-patient data listings including neutralizing antibodies (NAb) and ADA for all relevant timepoints.

All serum ADA samples that are confirmed positive in the ADA assay will be analyzed for the presence of NAb.

The proportion of patients with positive or negative ADA will be summarized by study phase, treatment group, and overall, at baseline (pre-dose), and post-baseline in the FAS. The frequency and percentage of patients with transient positive and persistent positive ADA will also be included in the summary. If warranted based on data, treatment-emergent transient positive and treatment-emergent persistent positive ADA will be summarized. The frequency and percentage of patients with confirmed positive, treatment-emergent confirmed positive, and negative ADA at each post-baseline timepoint will also be summarized. ADA designations and denominators are defined in table 7.

Table 7

Timepoint	**ADA Designations**	**Definition**	**Denominator/Evaluable Patients**
Baseline (Day 0 Pre-Dose)	Positive	Confirmed positive ADA at baseline	Number of patients with a baseline and at least one post-baseline ADA assessment
	Negative	Negative ADA at baseline	
Post-Baseline	Positive	Confirmed positive ADA at any time after baseline	Number of patients with at least one post-baseline ADA assessment
	Negative	Negative ADA at all times after baseline	
	Transient Positive[iv]	(1) Last ADA sampling timepoint is negative	

		AND 2a OR 2b: (2a) Only one post-baseline confirmed positive ADA (excluding the last sampling timepoint) exists OR (2b) Two or more post-baseline confirmed positive ADA where the first and last ADA confirmed positive samples are less than 16 weeks apart (irrespective of any negative ADA in between)	
	Persistent Positive[iv]	1 OR 2 OR 3: (1) Two or more post-baseline confirmed positive ADA where the first and last ADA confirmed positive samples are more than 16 weeks apart (irrespective of any negative ADA in between) OR (2) Confirmed positive ADA at the last post-baseline assessment OR (3) Negative ADA at the last post-baseline assessment and a prior confirmed positive ADA sample occurred less than 16 weeks before the last post-baseline negative ADA sample	
Treatment-Emergent	Positive	A post-baseline confirmed positive result exists with a titer >=4 times greater than the baseline titer OR	Number of patients with at least one post-baseline ADA assessment

		Any post-baseline confirmed positive result after a negative baseline result Note: The number of treatment-emergent transient positive and treatment-emergent persistent positive ADA will sum to the total positive treatment-emergent ADA Note: If baseline ADA data is missing, the baseline value is treated as negative for this analyses	
	Transient Positive[iv]	<u>Scenario 1:</u> (1) Negative ADA at baseline AND (2) Last ADA sampling timepoint is negative AND (3a OR 3b): (3a) Only one post-baseline confirmed positive ADA (excluding the last sampling timepoint) exists OR (3b) Two or more post-baseline confirmed positive ADA where the first and last ADA confirmed positive samples are less than 16 weeks apart (irrespective of any negative ADA in between) <u>Scenario 2:</u> (1) Confirmed positive ADA at baseline AND	

		(2) Last ADA sampling timepoint is negative or is the same as baseline ADA levels AND (3a OR 3b): (3a) Only one post-baseline confirmed positive ADA (excluding the last sampling timepoint) exists and the titer at the post-baseline timepoint is >=4 times greater than the baseline titer OR (3b) Two or more post-baseline confirmed positive ADA where the first and last ADA confirmed positive samples are less than 16 weeks apart (irrespective of any negative ADA in between) and the titers at each of the two or more post-baseline confirmed positive timepoints are >=4 times greater than the baseline titer Note: If baseline ADA data is missing, the baseline value is treated as negative for this analyses	
	Persistent Positive[iv]	<u>Scenario 1:</u> (1) Negative ADA at baseline AND (2 OR 3 OR 4): (2) Two or more post-baseline confirmed positive ADA where the first and last ADA confirmed positive samples are more than 16 weeks apart (irrespective of any negative ADA in between) OR	

		(3) Confirmed positive ADA at the last post-baseline assessment OR (4) Negative ADA at the last post-baseline assessment and a prior confirmed positive ADA sample occurred less than 16 weeks before the last post-baseline negative ADA sample <u>Scenario 2:</u> (1) Confirmed positive ADA at baseline AND (2 OR 3 OR 4): (2) Two or more post-baseline confirmed positive ADA where the first and last ADA confirmed positive samples are more than 16 weeks apart (irrespective of any negative ADA in between) and the titers at each of the two or more post-baseline confirmed positive timepoints are >=4 times greater than the baseline titer OR (3) Confirmed positive ADA at the last post-baseline assessment and the titer at the last post-baseline timepoint is >=4 times greater than the baseline titer OR (4) Negative/similar to baseline ADA at the last post-baseline assessment and a prior confirmed positive ADA sample occurred less than 16 weeks before the last post-	

		baseline negative/similar to baseline ADA sample and the titer of the confirmed positive ADA is >=4 times greater than the baseline titer Note: If baseline ADA data is missing, the baseline value is treated as negative for this analyses	
Post-baseline at each timepoint	Positive	Confirmed Positive ADA at the post-baseline timepoint	Number of patients with ADA assessment at baseline and at the post-baseline timepoint
	Treatment-Emergent	<u>Scenario 1:</u> (1) Negative ADA at baseline (2) Confirmed positive ADA at the post-baseline timepoint <u>Scenario 2:</u> (1) Confirmed positive ADA at baseline (2) Confirmed positive ADA at the post-baseline timepoint and the titer at the of the confirmed positive ADA is >=4 times greater than the baseline titer	
	Negative	Negative ADA at the post-baseline timepoint	
Overall	Positive	A patient with positive baseline (pre-dose) ADA or positive post-baseline ADA	Number of patients with a baseline or post-baseline ADA assessment
	Negative	A patient with negative baseline (pre-dose) ADA and negative post-baseline ADA at all time points	

The frequency and percentages of positive and negative neutralizing antibodies will be summarized in the FAS overall and at each relevant post-baseline timepoint (i.e., Day 15, Day 30, Day 60, Day 90, Day 180, Day 270). The denominators will be the number of patients with non-missing NAb data overall and at each timepoint, respectively.

The adverse event overview table from section 10.1 will be repeated by post-baseline positive and negative ADA as defined in table 7 in the safety analysis set. Additional adverse event analyses may be performed by post-baseline transient positive and persistent positive.

9. QUALITY OF LIFE

9.1. CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

9.2. CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

10. SAFETY

All safety analysis reporting will be based on the Safety Analysis Set. Unless otherwise noted, all safety analyses will be presented overall, by study phase, and treatment group.

10.1. Adverse Events

AEs will be collected and recorded for each patient from the date of the first dose of study drug until the end of their participation in the study, including the safety follow up period. Severity of adverse events will be graded by the investigator according to National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) version 4.03. AEs will also be assessed for relationship to study drug, and seriousness. Adverse events will be marked as AESIs, dose-limiting toxicities (DLTs), or injection site reactions (ISRs), where applicable. AEs will be considered treatment-emergent adverse events (TEAEs) if they started any time after first dose or if the year, month and day of start date were missing. Any missing severity assessments will be assumed to be Grade 3, missing relationship assessments will be assumed to be related, and missing seriousness assessments will be assumed as serious.

An overview of TEAEs will be produced, including counts and percentages of patients with any incidences of TEAE, TEAEs related to study treatment, SAEs, SAEs related to study treatment, any CTCAE grade 2 or higher TEAEs, any CTCAE grade 2 or higher TEAEs related to study treatment, any CTCAE grade 3 TEAEs, any CTCAE grade 3 or higher TEAEs related to study treatment, any CTCAE grade 4 TEAEs, any CTCAE grade 4 or higher TEAEs related to study treatment, TEAEs leading to study treatment discontinuation, treatment related TEAEs leading to study treatment discontinuation, TEAEs of special interest, related TEAEs of special interest, DLTs, related DLTs, TEAEs resulting in death, related TEAEs resulting in death, and injection site reactions. The maximum CTCAE grade (Grades 1-5) will also be included on the summary table.

Adverse events will be coded based on the MedDRA. Patients will be counted once per SOC (and likewise for PT within SOC). TEAEs will be summarized by SOC and PT in descending order of overall incidence of SOC (and PT within SOC). The MedDRA version used is defined in the study data management plan.

Summaries of adverse events by SOC and PT will include the following types:

- TEAEs;
- TEAEs related to study treatment;
- SAEs;
- Related Serious TEAEs;
- CTCAE Grade 3 or higher TEAEs;
- CTCAE Grade 3 or higher TEAEs related to study treatment;
- TEAEs leading to treatment discontinuation; and
- Dose-Limiting Toxicities
- TEAEs of Special Interest

A summary of TEAEs by SOC, PT, and maximum CTCAE grade will also be prepared. For these summaries, TEAEs will be sorted for each patient by SOC/PT and CTCAE grade; patients will be counted once within a SOC/PT based on their maximum CTCAE grade. Missing CTCAE grade will be counted as grade 3 and will also be presented in a 'Missing' row for completeness.

A summary of TEAEs by PT and maximum CTCAE grade will also be prepared with sections for any TEAE and TEAEs related to study treatment. For these summaries, patients will be counted once within a PT based on their maximum CTCAE grade. Missing CTCAE grade will be counted as grade 3. An additional summary of TEAEs by PT and maximum CTCAE grade will summarized for all TEAEs and TEAEs related to treatment for PTs with at least one patient experiencing a CTCAE grade 3 or higher and PTs where 10% or more of patients experience a CTCAE grade 1 or 2 for that PT. The summary will be sorted by the total incidence of the PT in the any TEAE section.

A comprehensive listing of all AEs will be provided in a by-patient data listing for the safety analysis set. In addition, the following listings will be provided for the safety analysis set:

- TEAEs related to study treatment;
- SAEs;
- TEAEs leading to treatment discontinuation; and
- AESIs
- CTCAE Grade 3 or Higher TEAEs

10.2. Clinical Laboratory Evaluations

Laboratory tests will be performed at times defined in the protocol Schedule of Assessments. Clinical chemistry and hematology parameters will be reported based on the International System of Units (SI). The following laboratory evaluations will be reported in data summaries, with asterisks indicating those that will be graded using NCI CTCAE. Grading will be performed by the Central Lab for relevant laboratory parameters based on the details provided in Appendix 13.11 . If a grade is not present in the central lab database for a graded lab parameter, the grade will be analyzed as grade 0.

Hematology: Hemoglobin*, Hematocrit, RBC count, RDW, MCV, MCH, MCHC, Platelet count*, WBC count*, CD4+ count*, Total neutrophils (Abs)*, Eosinophils (Abs), Monocytes (Abs), Basophils (Abs), Lymphocytes (Abs)*

Clinical chemistry: Alanine Aminotransferase*, Albumin*, Alkaline Phosphatase*, Aspartate Aminotransferase*, Bicarbonate, Bilirubin*, Blood urea nitrogen, C Reactive Protein, Calcium, Chloride, Creatine Kinase*, Creatinine*, Glucose (fasting)*, Glucose (random)*, Potassium*, Protein, Sodium*, Uric acid*, eGFR (Cockcroft-Gault)*, eGFR (MDRD)*, Total cholesterol (fasting)*, LDL-C (fasting), HDL-C (fasting), Triglycerides (fasting)*, Gamma Glutamyl Transferase*.

Coagulation: Prothrombin Intl. Normalized Ratio*, Activated Partial Thromboplastin Time*

Urinalysis: pH, Glucose (qual), Protein (qual)*, Blood (qual)*, Ketones (qual), Nitrites (qual), Leukocyte esterase (qual), Urobilinogen (qual), Urine bilirubin (qual), Microscopy

Observed values, changes, and percentage changes from baseline for hematology, clinical chemistry, and coagulation laboratory evaluations will be summarized at each visit and most extreme change. Endpoint analyses will be based on changes from baseline assessments at Days 120 and 210.

The number and percent of patients with a CTCAE toxicity grade of 3 or higher will be tabulated by laboratory evaluation with defined CTCAE grading at each visit for hematology, clinical chemistry, coagulation, and urinalysis (protein only). Patients with missing values post-baseline will be excluded from the summary. The denominator for percentages will be the number of patients with a post-baseline assessment for the laboratory parameter in question at that visit.

Hematology, clinical chemistry, coagulation, and urinalysis (protein only) shift tables displaying the shift from baseline to the worst value of CTCAE grade will be presented based on the most

extreme change as it relates to the relevant CTCAE definition. CTCAE relating to "high/hyper" conditions will depend on the maximum post-baseline value while CTCAE "low/hypo" will be reported based on the minimum post-baseline value. Separate shift tables will be prepared for parameters with bi-directional toxicity grading. Lab parameters which are optional, such as gamma glutamyl transferase, will not be included in the shift tables.

Hematology and clinical chemistry laboratory data which is not graded will also be summarized in shift tables of baseline to each visit and most extreme change based on range categories of low (below lower limit of normal [LLN], normal, and high [above upper limit of normal [ULN]]).

Urinalysis laboratory data (excluding protein) will also be summarized in a shift table of baseline to each visit based on range categories of low (below lower limit of normal [LLN]), normal, abnormal, and high (above upper limit of normal [ULN]).

Selected hematology and clinical chemistry laboratory results will be plotted over time for treatment group for MAD phase 1b portion of the study in the safety analysis set. The mean and SD of the change from baseline for selected hematology and clinical chemistry laboratory results will be plotted over time for the POC Phase 2 portion of the study in the safety analysis set.

All laboratory parameters as well as a summary of central laboratory tests and chest x-ray will be provided in patient data listings for all randomized patients. By patient listings of clinical chemistry and hematology data will be presented in the safety analysis set for the DMC and interim analysis only.

10.3. Other Safety Evaluations

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

The C-SSRS is a low-burden measure of the spectrum of suicidal ideation and behavior that was developed by Columbia University researchers for the National Institute of Mental Health Treatment of Adolescent Suicide Attempters Study to assess severity and track suicidal events through any treatment. It is a clinical interview providing a summary of both ideation and behavior that can be administered during any evaluation or risk assessment to identify the level and type of suicidality present. The C-SSRS can also be used during treatment to monitor for clinical worsening.[i] The C-SSRS evaluation will be performed as specified in the protocol Schedule of Assessments.

The C-SSRS has 11 binary (yes/no) outcomes corresponding to five categories of suicidal ideation, five categories of suicidal behavior, and the category of self-injurious behavior without suicidal intent:

Suicidal ideation is present if there is a "yes" response to any of the 5 suicidal ideation category questions. Suicidal behavior is present if there is a "yes" response to any of the 5 suicidal behavior category questions. Suicidal ideation or behavior is present if there is a "yes" response to any of the 10 suicidal or ideation or behavior category questions.

The number and percentage of patients with any post-baseline occurrence of the following ideations and behaviors will be presented overall, by study phase and treatment group:

- Suicidal ideation (overall and by individual question),
- Suicidal behavior (overall and by individual question),
- Suicidal ideation or behavior,

A similar tabulation of the occurrence of any of the above ideations and behaviors at baseline or as part of their lifetime history/past 12-month history (collected at the screening visit) will be presented overall, by study phase, and treatment group.

A shift table displaying the shift from worst pre-treatment category to the worst post-baseline category will be presented overall, by study phase, and treatment group. Best to worst C-SSRS category is defined in the following order: no suicidal ideation or behavior, suicidal ideation, and suicidal behavior. The worst post-baseline shift is calculated based on all post-baseline visits, including unscheduled visits. Worst pre-treatment will include all C-SSRS assessments prior to first dose date/time (screening or baseline). The denominator will be the number of patients with at least one non-missing baseline and post-baseline C-SSRS assessment.

Missing data will not be imputed.

All C-SSRS individual items will be presented in data listings for screening and post-screening (since last visit), separately. A separate listing will be presented for patients with suicidal ideation, suicidal behavior, or non-suicidal self-injurious behavior for all randomized patients.

10.3.2. Vital Signs

Vital signs include: heart rate (beats/min); temperature (^{o}C); systolic and diastolic blood pressure (mmHg). Observed values, changes, and percentage changes from baseline for vital signs will be summarized at each visit and time point, as well as for most extreme change overall, by study phase, and treatment group.

All vital signs data will be presented in patient data listings for all randomized patients.

10.3.3. 12-Lead Electrocardiogram (ECG)

ECG will be assessed as specified in the protocol Schedule of Assessments. The following ECG parameters will be collected: PR interval (msec), QRS interval (msec), RR interval (msec), QT interval (msec), and QT interval corrected for heart rate using Fridericia's formula (QTcF) (msec). ECG parameters and ECG evaluation will come directly from the database and will not be calculated during analysis.

Observed values, changes, and percentage changes from baseline for ECG parameters will be summarized at each visit, as well as for most extreme change overall, by study phase, and treatment group.

The number and percent of patients with the following parameters at any post-baseline visit and at each visit will be summarized:

- Abnormal, not clinically significant ECGs
- Abnormal, clinically significant ECGs

The number and percent of patients with each ECG evaluation result will be summarized by visit. The worst ECG evaluation at any post-baseline visit will also be summarized in the order of: Normal, Abnormal, not clinically significant, Abnormal, clinically significant.

For the summary of incidence at any post-baseline visit, the denominators for percentages will include patients with one or more non-missing value at any post-baseline visit. For the by-visit summaries, the denominators for percentages will include patients with a non-missing value at that visit.

Investigator reported ECG result shifts from baseline to each visit and worst case post-baseline will be summarized. Worst case post-baseline will be based on the most abnormal observed value on or after the randomization date.

All ECG data will be presented in a by patient data listing for all randomized patients. A separate by patient listing for patients with abnormal, clinically significant 12-lead ECGs will also be provided in the safety analysis set for DMC meetings and the interim analysis only.

10.3.4. Physical Examinations

Targeted physical examination data and full physical examination data will be presented in patient data listings. A separate by patient listing for patients with abnormal, clinically significant physical examination findings will also be provided in the safety analysis set for DMC meetings and the interim analysis only.

11. CHANGES TO THE PLANNED ANALYSIS

The following changes were made to the planned analyses:

- Time to SRI-4 Response was originally a planned endpoint for the MAD and POC portions of the study. This was removed from the POC portion of the study upon Protocol Amendment 6 (30Apr2020); however, the endpoint was inadvertently maintained in the protocol for the MAD portion of the study. This analysis will not be performed for any study part.
- Rules for applying LOCF were vague within the protocol and were further detailed in this SAP. See Section 7 for further details.
- The PP was clarified to include additional criteria beyond the protocol description. The PP will exclude patients with more than one missed dose.

- Cytokine, chemokine and genotyping data were originally planned to be part of the MAD and POC portion of the study; however, data for these parameters is only available for the MAD portion of the study.
- The definition of the CLASI response in the protocol is vague as the protocol defines the CLASI response as 50% improvement from baseline in “A” (total activity) or “B” (total damage) scores. The SAP further calrifies the CLASI response as 50% improvement from baseline in CLASI and distinguishes a CLASI-A response from a CLASI-B response.

12. REFERENCES

i. Posner K, Melvin G a, Stanley B, Oquendo M a, Gould M. Factors in the assessment of suicidality in youth. *CNS Spectr*. 2007;12(2):156-162.

ii. CCI [redacted]

iii. Van Vollenhoven RF, Hahn BH, Tsokos GC, et al. Efficacy and safety of ustekinumab, an IL-12 and IL-23 inhibitor, in patients with active systemic lupus erythematosus: results of a multicentre, double-blind, phase 2, randomised, controlled study. Lancet. 2018;392(10155):1330-1339.

iv. Shankar G, et al. Assessment and Reporting of the Clinical Immunogenicity of Therapeutic Proteins and Peptides – Harmonized Terminology and Tactical Recommendations. AAPS Journal, 2014; 16(4):663.

v. International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH E3 Guideline: Structure and content of clinical study reports questions & answers (R1). 2012. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_QAs_R1_Step4.pdf. Accessed 13 Jan 2016.

vi. International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonised Tripartite Guideline: Structure and content of clinical study reports E3. Step 4. 1995. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf. Accessed 13 Jan 2016.

vii. International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH E9: Statistical principles for clinical trials. 1998. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf . Accessed 15 June 2016.

viii. FDA Guidance for Industry Systemic Lupus Erythematosus – Developing Medical Products for Treatment. 2010. https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm072063.pdf. Accessed 23February2018.

ix. Committee for Medicinal Products for Human use (CHMP). EMA Guideline on clinical investigation of medicinal products for the treatment of systemic lupus erythematosus and lupus nephritis. 2015. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/03/WC500184889.pdf. Accessed 04March2018.

x. CCI [redacted]

xi. Nilsson M, Suryawanshi S, Gassmann-Mayer C, Dubrava S, McSorley P, Jiang K. Columbia-Suicide Severity Rating Scale Scoring and Data Analysis Guide. V2 2013. https://cssrs.columbia.edu/wp-content/uploads/ScoringandDataAnalysisGuide-for-Clinical-Trials-1.pdf. Accessed 18May2018.

xii. CCI [redacted]

xiii. CCI [redacted]

xiv. CCI [redacted]

13. APPENDICES

13.1. COVID-19 Impact and Data Handling

The COVID-19 Pandemic resulted in contingency plans for BOS161721-02.

- **Safety Oversight:** In case a patient cannot return to the study site for the scheduled visit, the site staff will contact the patient remotely for safety follow up, for example, via telephone or video conferencing. The analysis of safety data will include all collected data regardless of remote data collection.

- **Central Laboratory:** In the case of courier issues that will prevent the protocol-required laboratory specimens to be sent to the study core laboratory, the site should have the safety laboratory specimens (Hematology, Chemistry, and Urinalysis) sent to their local laboratory for analysis and review by the study physician. The local laboratory data will not be entered into the clinical database or analyzed/presented in data summaries.

- **Investigational Product Dosing:** In cases when dosing cannot be performed during the protocol-designated windows, the Investigator should discuss each case with the Sponsor to determine whether it is a missed visit or whether the dosing can be performed outside protocol windows (as a protocol deviation). A minimum of 2 weeks must be maintained between consecutive doses. Patients who miss more than 2 doses of the study drug (whether consecutive or not) will be discontinued from the study and will undergo assessments as outlined in Section 5.5 of the study protocol (Premature Discontinuation). A sensitivity analysis of SRI-4 response using LOCF for missing data and/or observed case data may be performed to assess the impact of patients that did not follow the original protocol dosing schedule.

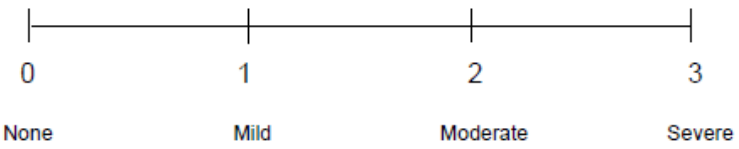
An addendum to this SAP will be developed that documents all additional data handling techniques employed to manage COVID impacted data, if needed, prior to database lock and unblinding.

13.2. Physician's Global Assessment

Subject ID:__ __ __ __ __ __ __ Patient Initials:__ __ __ Date: __ __-__ __ __-__ __ __ __ Visit: __ __
DD MON YEAR

PHYSICIAN'S GLOBAL ASSESSMENT (PGA)

How do you assess your subject's current disease activity?



Result: _____mm

Note: Please review the scale from previous visit, as applicable, and mark a vertical line on the scale above to assess the overall status of the subject's SLE signs and symptoms, and the functional capacity. Zero (0) = very good, asymptomatic and no limitation of normal activity; Three (3) = most severe possible disease ever seen in all SLE patients.

Investigator Signature	Assessor Initials	Date

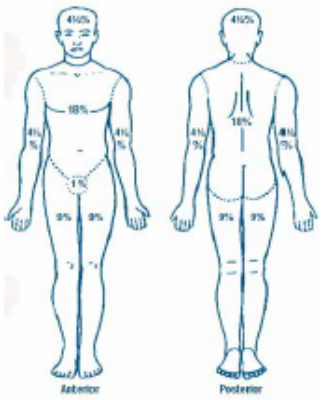
13.3. SLEDAI-2K Index

Subject Number:___ ___ ___ -___ ___ ___ ___	Protocol BOS161721-02
Screening / Randomization / _____ Week Visit & Date of Assessment: __ __ - __ __ __ - 20__ __ (Circle visit type &/or enter visit week) (Day Month Year)	

SLEDAI 2K Index Must be present at the time of visit or in the preceding 30 days (or 10 days for V1 and V2) and SLE-Related to be checked!			
Wt	(√) if Present	Descriptor	Definitions If descriptor is checked as present, please check appropriate condition(s) in definition and/or specify in the space provided.
8	☐	Seizure	Recent onset (last 28 days). Exclude metabolic, infectious or drug causes. **Specify if checked:**
8	☐	Psychosis	Altered ability to function in normal activity due to severe disturbance in the perception of reality. Include: □ hallucinations, □ incoherence, □ marked loose associations, □ impoverished thought content, □ marked illogical thinking, and □ bizarre, disorganized, or catatonic behavior. Exclude uremia and drug causes.
8	☐	Organic Brain Syndrome	Altered mental function with □ impaired orientation, memory or other intellectual function (with rapid onset and fluctuating clinical features), □ inability to sustain attention to environment, and at least 2 of the following: □ perceptual disturbance, □ incoherent speech, □ insomnia or daytime drowsiness, □ increased or decreased psychomotor activity. Exclude metabolic, infectious or drug causes.
8	☐	Visual Disturbance	Retinal changes of SLE. Include □ cytoid bodies, □ retinal hemorrhages, □ serous exudate or hemorrhages in the choroid, and □ optic neuritis. Exclude hypertension, infection, or drug causes.
8	☐	Cranial Nerve Disorder	New onset of □ sensory or □ motor neuropathy involving cranial nerves.
8	☐	Lupus Headache	Severe, persistent headache. May be migrainous, but must be nonresponsive to narcotic analgesia. (Must have been severe enough to warrant lumbar puncture and/or MRI or head CT, perhaps hospitalization, and is thought to be due to active lupus cerebritis.) **Specify if checked:**
8	☐	CVA	New onset of cerebrovascular accident(s) (CVA). Exclude arteriosclerosis. **Specify if checked:**
8	☐	Vasculitis	□ Ulceration, □ gangrene, □ tender finger nodules, □ periungual infarction, □ splinter hemorrhages, or □ biopsy or □ angiogram proof of vasculitis.

Subject Number___ ___ ___-___ ___ ___ ___			Protocol BOS161721-02
Screening / Randomization / _____ Week Visit & Date of Assessment: ___ ___ - ___ ___ ___ - 20____ (Circle visit type &/or enter visit week) (Day Month Year)			
4	☐	Arthritis	≥ 2 joints with pain and signs of inflammation (□ tenderness, □ swelling or □ effusion). **(Complete Joint Count if checked.)**
4	☐	Myositis	Proximal muscle aching/weakness associated with □ elevated creatine phosphokinase (CK)/aldolase or □ EMG changes or □ a biopsy showing myositis.
4	☐	Urinary Casts	Heme-granular or red blood cell casts. (See labs)
4	☐	Hematuria	>5 RBC/high power field. Exclude stone, infection, or other cause. (See labs)
4	☐	Proteinuria	>0.5 gm/24-hour equivalent. (See labs)
4	☐	Pyuria	>5 WBC/high power field. Exclude infection. (See labs)
2	☐	Rash	Inflammatory-type rash. **Specify and complete CLASI if checked:**
2	☐	Alopecia	□ Abnormal, □ patchy or □ diffuse loss of hair. **(Complete CLASI if checked.)**
2	☐	Mucosal Ulcers	□ Oral or □ □nasal ulcerations. **(Complete CLASI if checked.)**
2	☐	Pleurisy	Pleuritic chest pain with □ pleural rub or □ effusion, or □ pleural thickening.
2	☐	Pericarditis	Pericardial pain with at least one (1) of the following: □ rub or □ effusion, or □ ECG or □ □echocardiogram confirmation.
2	☐	Low Complement	Decrease in CH50, C3 or C4 below the lower limit of normal for testing laboratory. (See labs)
2	☐	Increased DNA Binding	>25% binding by Farr assay or above normal range for testing laboratory. (See labs)
1	☐	Fever	>38°C (100.4F). Exclude infectious cause. **Specify if checked:**
1	☐	Thrombocytopenia	<100,000 platelets/mm^3 (See labs)
1	☐	Leukopenia	<3,000 white blood cells/mm^3. Exclude drug causes. (See labs)
	TOTAL SCORE (Sum of all of weights next to descriptors marked present)		
Investigator Signature			Date

13.4. BILAG 2004 Index

BILAG 2004 Index Source Document Worksheet		
Subject Number:___ ___ ___ - ___ ___ ___ ___		**Protocol BOS161721-02**
Screening / Randomization / _____ Week Visit & Date of Assessment: ___ ___ - ___ ___ ___ - 20___ (Circle visit type &/or enter visit week) (Day Month Year)		
Only record items due to SLE disease activity. Each assessment refers to manifestation occurring in the last 4 weeks (compared with the previous visit). Scoring: 0 = Not Present, 1 = Improving, 2 = Same, 3 = Worse, 4 = New or Recurrence **There must be detailed documentation for all descriptors scored 1 – 4**		
CONSTITUTIONAL	Score	**Detail findings/changes for all descriptors scored 1 - 4**
1. Pyrexia – documented > 37.5° C		
2. Weight loss – unintentional > 5%		
3. Lymphandenopathy / splenomegaly		
4. Anorexia		
MUCOCUTANEOUS - **If any of these features are scored 1-4, please complete CLASI!**	Score	**Draw / specify area(s) and note approximate BSA (use also for CLASI documentation)**
5. Skin eruption – severe (Must involve >18% BSA)		
6. Skin eruption – mild (≤18% BSA)		
7. Angio–oedema – severe		
8. Angio–oedema - mild		
9. Mucosal ulceration – severe		
10. Mucosal ulceration – mild		
11. Panniculitis / Bullous lupus – severe (> 9% BSA)		
12. Panniculitis / Bullous lupus – mild (≤ 9% BSA)		
13. Major cutaneous vasculitis / thrombosis		
14. Digital infracts or nodular vasculitis		
15. Alopecia – severe		
16. Alopecia – mild		
17. Peri-ungual erythema / chilblains		
18. Splinter hemorrhages		
NEUROPSYCHIATRIC	Score	**Detail findings/changes for all descriptors scored 1 - 4**
19. Aseptic meningitis		
20. Cerebral vasculitis		
21. Demyelinating syndrome		
22. Myelopathy		
23. Acute confusional state		
24. Psychosis		
25. Acute inflammatory demyelinating polyradiculoneuropathy		
26. Mononeuropathy (single / multiplex)		
27. Cranial neuropathy		
28. Plexopathy		
29. Polyneuropathy		
30. Seizure disorder		
31. Status epilepticus		
32. Cerebrovascular disease (not due to vasculitis)		
33. Cognitive dysfunction		
34. Movement disorder		
35. Autonomic disorder		
36. Cerebellar ataxia (isolated)		
37. Lupus headache – severe unremitting		
38. Headache from IC hypertension		
MUSCULOSKELETAL	Score	**Detail findings/changes for all descriptors scored 1 - 4**
39. Myositis – severe		
40. Myositis – mild		
41. Arthritis – severe		
42. Arthritis – moderate / Tendonitis / Tenosynovitis		
43. Arthritis – mild / Arthralgia / Myalgia		

Subject Number: ___ ___ ___ -___ ___ ___ ___	Protocol BOS161721-02
Screening / Randomization / _____ Week Visit (Circle visit type &/or enter visit week) & Date of Assessment: __ __ - __ __ __ - 20___ (Day Month Year)	

CARDIORESPIRATORY	Score	Detail findings/changes for all descriptors scored 1 - 4
44. Myocarditis – mild		
45. Myocarditis / Endocarditis + Cardiac failure		
46. Arrhythmia		
47. New valvular dysfunction		
48. Pleurisy / Pericarditis		
49. Cardiac tamponade		
50. Pleural effusion with dyspnoea		
51. Pulmonary haemorrhage / vasculitis		
52. Interstitial alveolitis / pneumonitis		
53. Shrinking lung syndrome		
54. Aortitis		
55. Coronary vasculitis		
GASTROINTESTINAL	Score	Detail findings/changes for all descriptors scored 1 - 4
56. Lupus peritonitis		
57. Abdominal serositis or ascites		
58. Lupus enteritis / colitis		
59. Malabsorption		
60. Protein-losing enteropathy		
61. Intestinal pseudo-obstruction		
62. Lupus hepatitis		
63. Acute lupus cholecystitis		
64. Acute lupus pancreatitis		
OPHTHALMIC	Score	Detail findings/changes for all descriptors scored 1 - 4
65. Orbital inflammation / proptosis		
66. Keratitis – severe		
67. Keratitis - mild		
68. Anterior uveitis		
69. Posterior uveitis / retinal vasculitis – severe		
70. Posterior uveitis / retinal vasculitis – mild		
71. Episcleritis		
72. Scleritis – severe		
73. Scleritis – mild		
74. Retinal / choroidal vaso-occlusive disease		
75. Isolated cotton wool-spots (cytoid bodies)		
76. Optic neuritis		
77. Anterior ischaemic optic neuropathy		

RENAL	Value	SLE related?	HAEMATOLOGICAL	Value	SLE related?
78. Systolic BP > 140mmHg		Y / N	90. Haemoglobin (g/dl)		Y / N
79. Diastolic BP > 90 mmHg		Y / N	91. Total white cell count		Y / N
80. Accelerated hypertension	Yes / No		92. Neutrophils		Y / N
81. Urine dipstick (protein) (+=1, ++=2, +++=3)		Y / N	93. Lymphocytes		Y / N
82. Urine albumin-creatinine ratio (mg/mmol)	Not assessed		94. Platelets		Y / N
83. Urine protein-creatinine ratio (mg/mmol)		Y / N	95. TTP	0 1 2 3 4	
84. 24 hour urine protein		Y / N	96. Evidence of active haemolysis	Yes / No	
85. Nephrotic syndrome	Yes / No		97. Coombs' test positive (isolated)	Yes / No	
86. Creatinine (plasma/serum) (µmol/l)		Y / N	Specify lab findings / additional comments:		
87. GFR		Y / N			
88. Active urinary sediment	Yes / No				
89. Active nephritis	Yes / No				
Investigator Signature	Date				

13.5. ACR-28 Joint Count Assessment

Site/Subject #: ___-____	Protocol BOS161721-02
Visit _____	Date of Assessment: ____/_____/_____ (DD / MMM / YYYY)

28 JOINT COUNT ASSESSMENT

Please circle all that apply.
Query the subject at the start of the joint count about pain (prior to assessment of tenderness and swelling) as to whether they have experienced or are experiencing pain in any of the 28 joints. Artificial joints and joints that have ever had surgery performed on them should be scored as "NA."

	Joint #	Right Side						Left Side					
		Tenderness			Swelling			Tenderness			Swelling		
Upper Extremity													
Shoulder	1	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
Elbow	2	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
Wrist	3	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
MCP I	4	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
MCP II	5	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
MCP III	6	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
MCP IV	7	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
MCP V	8	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
IP	9	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
PIP II	10	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
PIP III	11	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
PIP IV	12	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
PIP V	13	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
Lower Extremity													
Knee	14	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA

Are there any other joints (not listed above) that are significantly affected by SLE? No Yes
If Yes, list and describe: _____

_____

Based on clinical exam/impression and patient report for the last 4 weeks*:
Is there loss of functional range of motion sufficient to interfere with instrumental activities of daily living (eg, household chores, preparing meals, working, etc.)? No Yes

Is there significant impairment of basic activities of daily living (ADLs) so that the patient requires assistance or an assistive device due to the active inflammation? No Yes

If Yes, activities affected:
Grooming Dressing Ambulating Toileting Feeding oneself

_____

INVESTIGATOR SIGNATURE: _____ **DATE:** ____ / _____ / _____
DD MMM YYYY

13.6. SLICC/ACR Damage Index

Subject Number: ___ ___ ___-___ ___ ___ ___	Protocol BOS161721-02

SLICC / ACR Damage Index

Date of Assessment: ____ - ____ - 20____ (Day Month Year) **Please circle Study Visit: Day 1 or Day 180**

Item	Score
Ocular (either eye, by clinical assessment)	
Any cataract ever	☐ 1
Retinal change or optic atrophy	☐ 1
Neuropsychiatric	
Cognitive Impairment (eg memory deficit, difficulty with calculation, poor concentration, difficulty in spoken or written language, impaired performance levels) or major psychosis	☐ 1
Seizures requiring therapy for 6 months	☐ 1
Cerebrovascular accident ever (score 2 if > 1)	☐ 1 ☐ 2
Cranial or peripheral neuropathy (excluding optic)	☐ 1
Transverse myelitis	☐ 1
Renal	
Estimated or measured glomerular filtration rate < 50%	☐ 1
Proteinuria > 3.5gm/24 hours	☐ 1
OR	
End-stage renal disease (regardless of dialysis or transplantation)	☐ 3
Pulmonary	
Pulmonary hypertension (right ventricular prominence, or loud P2)	☐ 1
Pulmonary fibrosis (physical and radiographic)	☐ 1
Shrinking lung	☐ 1
Pleural fibrosis (radiographic)	☐ 1
Pulmonary Infarction (radiographic)	☐ 1
Cardiovascular	
Angina or coronary artery bypass	☐ 1
Myocardial Infarction ever (score 2 if > 1)	☐ 1 ☐ 2
Cardiomyopathy (ventricular dysfunction)	☐ 1
Valvular disease (diastolic, murmur, or systolic murmur > 3/6)	☐ 1
Pericarditis for 6 months, or pericardiectomy	☐ 1
Peripheral Vascular	
Claudication for 6 months	☐ 1
Minor tissue loss (pulp space)	☐ 1
Significant tissue loss ever (eg loss of digit or limb) (score 2 if > 1 site)	☐ 1 ☐ 2
Venous thrombosis with swelling, ulceration, or venous stasis	☐ 1
Gastrointestinal	
Infarction or resection or bowel below duodenum, spleen, liver, or gall bladder ever, for any cause (score 2 if > 1 site)	☐ 1 ☐ 2
Mesenteric insufficiency	☐ 1
Chronic peritonitis	☐ 1
Stricture or upper gastrointestinal tract surgery ever	☐ 1
Musculoskeletal	
Muscle atrophy or weakness	☐ 1
Deforming or erosive arthritis (including reducible deformities, excluding avascular necrosis	☐ 1
Osteoporosis with fracture or vertebral collapse (excluding avascular necrosis)	☐ 1
Avascular necrosis (score 2 is > 1)	☐ 1 ☐ 2
Osteomyelitis	☐ 1
Skin	
Scarring chronic alopecia	☐ 1
Extensive scarring or panniculum other than scalp and pulp space	☐ 1
Skin ulceration (excluding thrombosis) for > 6 months	☐ 1
Premature gonadal failure	☐ 1
Diabetes (regardless of treatment)	☐ 1
Malignancy (excluding dysplasia) (score 2 if > 1 site)	☐ 1 ☐ 2

13.7. CLASI

BOS161721-02 Subject: _ _ _ _ _ _ _ _ Visit: _____ Date: _ _-_ _ _-20_ _
Day Month Year

Cutaneous LE Disease Area and Severity Index (CLASI)

Select the score in each anatomical location that describes the most severely affected cutaneous lupus-associated lesion

	activity		damage		
Anatomical Location	Erythema	Scale/ Hypertrophy	Dyspigmentation	Scarring/ Atrophy/ Panniculitis	Anatomical Location
	0-absent 1-pink; faint erythema 2- red; 3-dark red; purple/violaceous/ crusted/ hemorrhagic	0-absent; 1-scale 2-verrucous/ hypertrophic	0-absent, 1-dyspigmentaton	0 – absent 1 – scarring 2 – severely atrophic scarring or panniculitis	
Scalp				See below	Scalp
Ears					Ears
Nose (incl. malar area)					Nose (incl. malar area)
Rest of the face					Rest of the face
V-area neck (frontal)					V-area neck (frontal)
Post. Neck &/or shoulders					Post. Neck &/or shoulders
Chest					Chest
Abdomen					Abdomen
Back, buttocks					Back, buttocks
Arms					Arms
Hands					Hands
Legs					Legs
Feet					Feet

Mucous membrane

Mucous membrane lesions (examine if patient confirms involvement)	
0-absent; 1-lesion or ulceration	

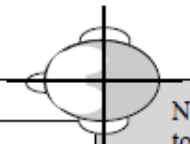
Dyspigmentation

Report duration of dyspigmentation after active lesions have resolved (verbal report by patient – tick appropriate box)

☐ Dyspigmentation usually lasts less than 12 months (dyspigmentation score above remains)

☐ Dyspigmentation usually lasts at least 12 months (dyspigmentation score is doubled)

Alopecia



Recent Hair loss (within the last 30 days / as reported by patient)	
1-Yes 0-No	

NB: if scarring and non-scarring aspects seem to coexist in one lesion, please score both

Divide the scalp into four quadrants as shown. The dividing line between right and left is the midline. The dividing line between frontal and occipital is the line connecting the highest points of the ear lobe. A quadrant is considered affected if there is a lesion within the quadrant.

Alopecia (clinically not obviously scarred)		Scarring of the scalp (judged clinically)	
0-absent 1-diffuse; non-inflammatory 2-focal or patchy in one quadrant; 3-focal or patchy in more than one quadrant		0- absent 3- in one quadrant 4- two quadrants 5- three quadrants 6- affects the whole skull	

Total Activity Score
(For the activity score please add up the scores of the left side i.e. for Erythema, Scale/Hypertrophy, Mucous membrane involvement and Alopecia)

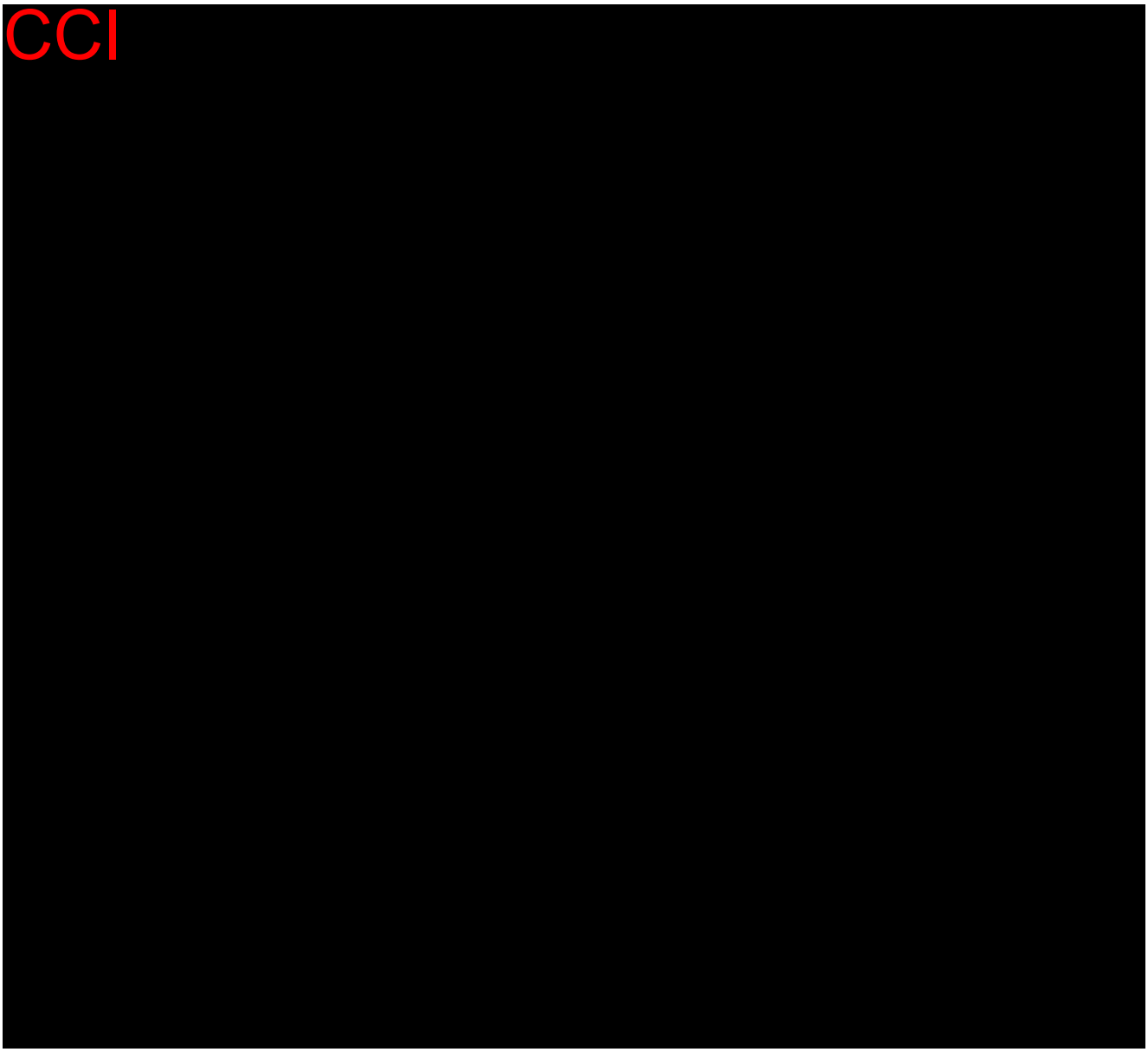
Total Damage Score
(For the damage score, please add up the scores of the right side, i.e. for Dyspigmentation, Scarring/Atrophy/Panniculitis and Scarring of the Scalp)

13.8. CCI

13.8.1. CCI

CCI


CCI


CCI

CCI

[redacted]

CCI

CCI

CCI

13.9. C-SSRS

13.9.1. C-SSRS (Baseline/Screening)

SUICIDAL IDEATION		
Ask questions 1 and 2. If both are negative, proceed to "Suicidal Behavior" section. If the answer to question 2 is "yes", ask questions 3, 4 and 5. If the answer to question 1 and/or 2 is "yes", complete "Intensity of Ideation" section below.	**Lifetime: Time He/She Felt Most Suicidal**	**Past 12 Months**
1. Wish to be Dead Subject endorses thoughts about a wish to be dead or not alive anymore, or wish to fall asleep and not wake up. ***Have you wished you were dead or wished you could go to sleep and not wake up?*** If yes, describe:	**Yes** ☐ **No** ☐	**Yes** ☐ **No** ☐
2. Non-Specific Active Suicidal Thoughts General non-specific thoughts of wanting to end one's life/commit suicide (e.g., *"I've thought about killing myself"*) without thoughts of ways to kill oneself/associated methods, intent, or plan during the assessment period. ***Have you actually had any thoughts of killing yourself?*** If yes, describe:	**Yes** ☐ **No** ☐	**Yes** ☐ **No** ☐
3. Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act Subject endorses thoughts of suicide and has thought of at least one method during the assessment period. This is different than a specific plan with time, place or method details worked out (e.g., thought of method to kill self but not a specific plan). Includes person who would say, *"I thought about taking an overdose but I never made a specific plan as to when, where or how I would actually do it… and I would never go through with it."* ***Have you been thinking about how you might do this?*** If yes, describe:	**Yes** ☐ **No** ☐	**Yes** ☐ **No** ☐
4. Active Suicidal Ideation with Some Intent to Act, without Specific Plan Active suicidal thoughts of killing oneself and subject reports having some intent to act on such thoughts, as opposed to *"I have the thoughts but I definitely will not do anything about them."* ***Have you had these thoughts and had some intention of acting on them?*** If yes, describe:	**Yes** ☐ **No** ☐	**Yes** ☐ **No** ☐
5. Active Suicidal Ideation with Specific Plan and Intent Thoughts of killing oneself with details of plan fully or partially worked out and subject has some intent to carry it out. ***Have you started to work out or worked out the details of how to kill yourself? Do you intend to carry out this plan?*** If yes, describe:	**Yes** ☐ **No** ☐	**Yes** ☐ **No** ☐
INTENSITY OF IDEATION		
The following features should be rated with respect to the most severe type of ideation (i.e., 1-5 from above, with 1 being the least severe and 5 being the most severe). Ask about time he/she was feeling the most suicidal. Lifetime - ***Most Severe Ideation:*** ____ (*Type # (1-5)*) ____ (*Description of Ideation*) Past 12 Months - ***Most Severe Ideation:*** ____ (*Type # (1-5)*) ____ (*Description of Ideation*)	Most Severe	Most Severe
Frequency ***How many times have you had these thoughts?*** (1) Less than once a week (2) Once a week (3) 2-5 times in week (4) Daily or almost daily (5) Many times each day	____	____
Duration ***When you have the thoughts how long do they last?*** (1) Fleeting - few seconds or minutes (2) Less than 1 hour/some of the time (3) 1-4 hours/a lot of time (4) 4-8 hours/most of day (5) More than 8 hours/persistent or continuous	____	____
Controllability ***Could/can you stop thinking about killing yourself or wanting to die if you want to?*** (1) Easily able to control thoughts (2) Can control thoughts with little difficulty (3) Can control thoughts with some difficulty (4) Can control thoughts with a lot of difficulty (5) Unable to control thoughts (0) Does not attempt to control thoughts	____	____
Deterrents ***Are there things - anyone or anything (e.g., family, religion, pain of death) - that stopped you from wanting to die or acting on thoughts of committing suicide?*** (1) Deterrents definitely stopped you from attempting suicide (2) Deterrents probably stopped you (3) Uncertain that deterrents stopped you (4) Deterrents most likely did not stop you (5) Deterrents definitely did not stop you (0) Does not apply	____	____
Reasons for Ideation ***What sort of reasons did you have for thinking about wanting to die or killing yourself? Was it to end the pain or stop the way you were feeling (in other words you couldn't go on living with this pain or how you were feeling) or was it to get attention, revenge or a reaction from others? Or both?*** (1) Completely to get attention, revenge or a reaction from others (2) Mostly to get attention, revenge or a reaction from others (3) Equally to get attention, revenge or a reaction from others and to end/stop the pain (4) Mostly to end or stop the pain (you couldn't go on living with the pain or how you were feeling) (5) Completely to end or stop the pain (you couldn't go on living with the pain or how you were feeling) (0) Does not apply	____	____

SUICIDAL BEHAVIOR *(Check all that apply, so long as these are separate events; must ask about all types)*		**Lifetime**	**Past 1 Year**
Actual Attempt: A potentially self-injurious act committed with at least some wish to die, *as a result of act*. Behavior was in part thought of as method to kill oneself. Intent does not have to be 100%. If there is ***any*** intent/desire to die associated with the act, then it can be considered an actual suicide attempt. ***There does not have to be any injury or harm***, just the potential for injury or harm. If person pulls trigger while gun is in mouth but gun is broken so no injury results, this is considered an attempt. Inferring Intent: Even if an individual denies intent/wish to die, it may be inferred clinically from the behavior or circumstances. For example, a highly lethal act that is clearly not an accident so no other intent but suicide can be inferred (e.g., gunshot to head, jumping from window of a high floor/story). Also, if someone denies intent to die, but they thought that what they did could be lethal, intent may be inferred. ***Have you made a suicide attempt?*** ***Have you done anything to harm yourself?*** ***Have you done anything dangerous where you could have died?*** ***What did you do?*** ***Did you_____as a way to end your life?*** ***Did you want to die (even a little) when you_____?*** ***Were you trying to end your life when you_____?*** ***Or did you think it was possible you could have died from_____?*** ***Or did you do it purely for other reasons / without ANY intention of killing yourself (like to relieve stress, feel better, get sympathy, or get something else to happen)?*** (Self-Injurious Behavior without suicidal intent) If yes, describe:		Yes ☐ No ☐ Total # of Attempts _____	Yes ☐ No ☐ Total # of Attempts _____
Has subject engaged in Non-Suicidal Self-Injurious Behavior?		Yes ☐ No ☐	Yes ☐ No ☐
Interrupted Attempt: When the person is interrupted (by an outside circumstance) from starting the potentially self-injurious act *(if not for that, actual attempt would have occurred)*. Overdose: Person has pills in hand but is stopped from ingesting. Once they ingest any pills, this becomes an attempt rather than an interrupted attempt. Shooting: Person has gun pointed toward self, gun is taken away by someone else, or is somehow prevented from pulling trigger. Once they pull the trigger, even if the gun fails to fire, it is an attempt. Jumping: Person is poised to jump, is grabbed and taken down from ledge. Hanging: Person has noose around neck but has not yet started to hang - is stopped from doing so. ***Has there been a time when you started to do something to end your life but someone or something stopped you before you actually did anything?*** If yes, describe:		Yes ☐ No ☐ Total # of interrupted _____	Yes ☐ No ☐ Total # of interrupted _____
Aborted Attempt: When person begins to take steps toward making a suicide attempt, but stops themselves before they actually have engaged in any self-destructive behavior. Examples are similar to interrupted attempts, except that the individual stops him/herself, instead of being stopped by something else. ***Has there been a time when you started to do something to try to end your life but you stopped yourself before you actually did anything?*** If yes, describe:		Yes ☐ No ☐ Total # of aborted _____	Yes ☐ No ☐ Total # of aborted _____
Preparatory Acts or Behavior: Acts or preparation towards imminently making a suicide attempt. This can include anything beyond a verbalization or thought, such as assembling a specific method (e.g., buying pills, purchasing a gun) or preparing for one's death by suicide (e.g., giving things away, writing a suicide note). ***Have you taken any steps towards making a suicide attempt or preparing to kill yourself (such as collecting pills, getting a gun, giving valuables away or writing a suicide note)?*** If yes, describe:		Yes ☐ No ☐	Yes ☐ No ☐
Suicidal Behavior: Suicidal behavior was present during the assessment period?		Yes ☐ No ☐	Yes ☐ No ☐
Answer for Actual Attempts Only	Most Recent Attempt Date:	Most Lethal Attempt Date:	Initial/First Attempt Date:
Actual Lethality/Medical Damage: 0. No physical damage or very minor physical damage (e.g., surface scratches). 1. Minor physical damage (e.g., lethargic speech; first-degree burns; mild bleeding; sprains). 2. Moderate physical damage; medical attention needed (e.g., conscious but sleepy, somewhat responsive; second-degree burns; bleeding of major vessel). 3. Moderately severe physical damage; *medical* hospitalization and likely intensive care required (e.g., comatose with reflexes intact; third-degree burns less than 20% of body; extensive blood loss but can recover; major fractures). 4. Severe physical damage; *medical* hospitalization with intensive care required (e.g., comatose without reflexes; third-degree burns over 20% of body; extensive blood loss with unstable vital signs; major damage to a vital area). 5. Death	*Enter Code* _____	*Enter Code* _____	*Enter Code* _____
Potential Lethality: Only Answer if Actual Lethality=0 Likely lethality of actual attempt if no medical damage (the following examples, while having no actual medical damage, had potential for very serious lethality: put gun in mouth and pulled the trigger but gun fails to fire so no medical damage; laying on train tracks with oncoming train but pulled away before run over). 0 = Behavior not likely to result in injury 1 = Behavior likely to result in injury but not likely to cause death 2 = Behavior likely to result in death despite available medical care	*Enter Code* _____	*Enter Code* _____	*Enter Code* _____

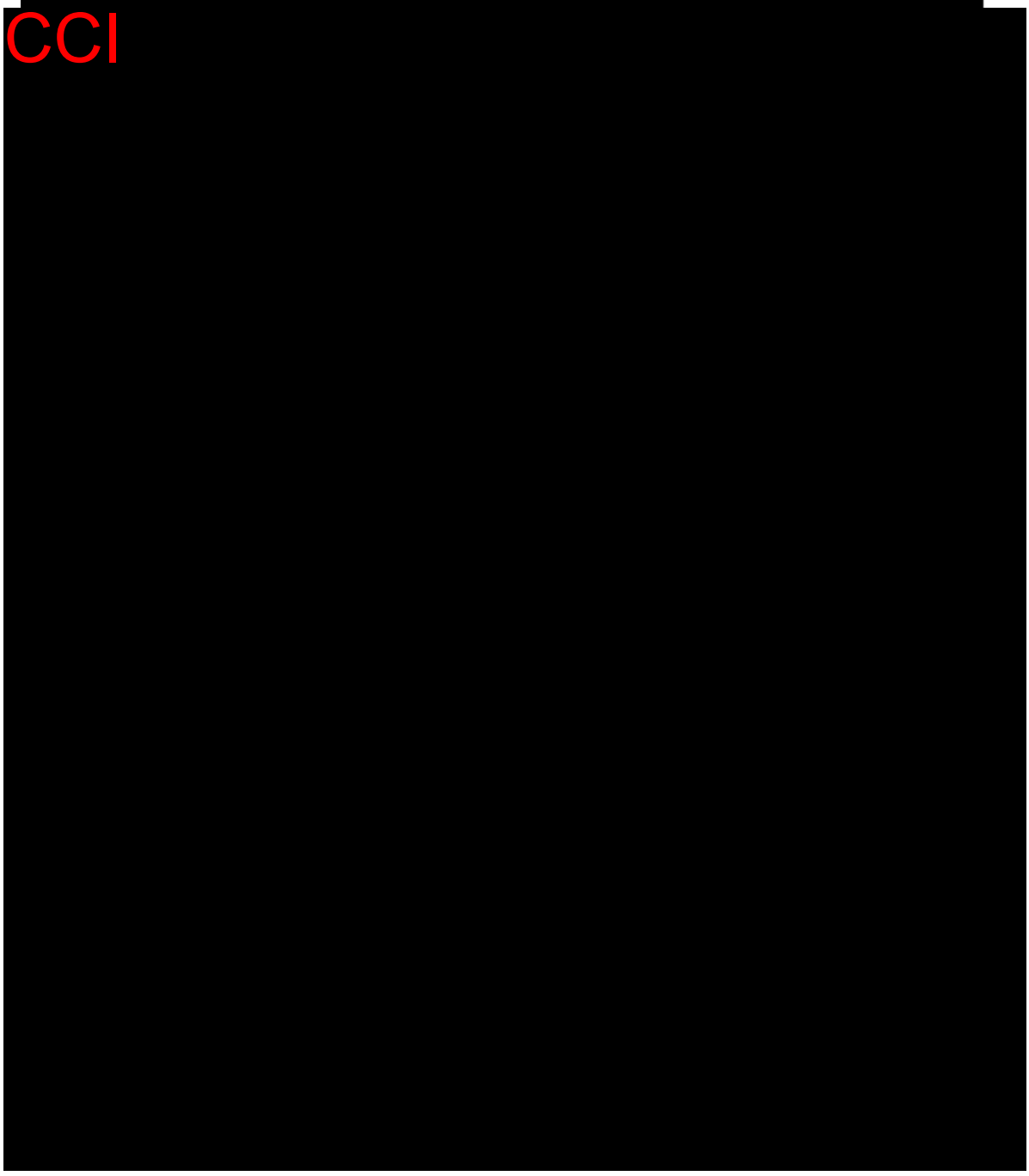
13.9.2. C-SSRS (Since Last Visit)

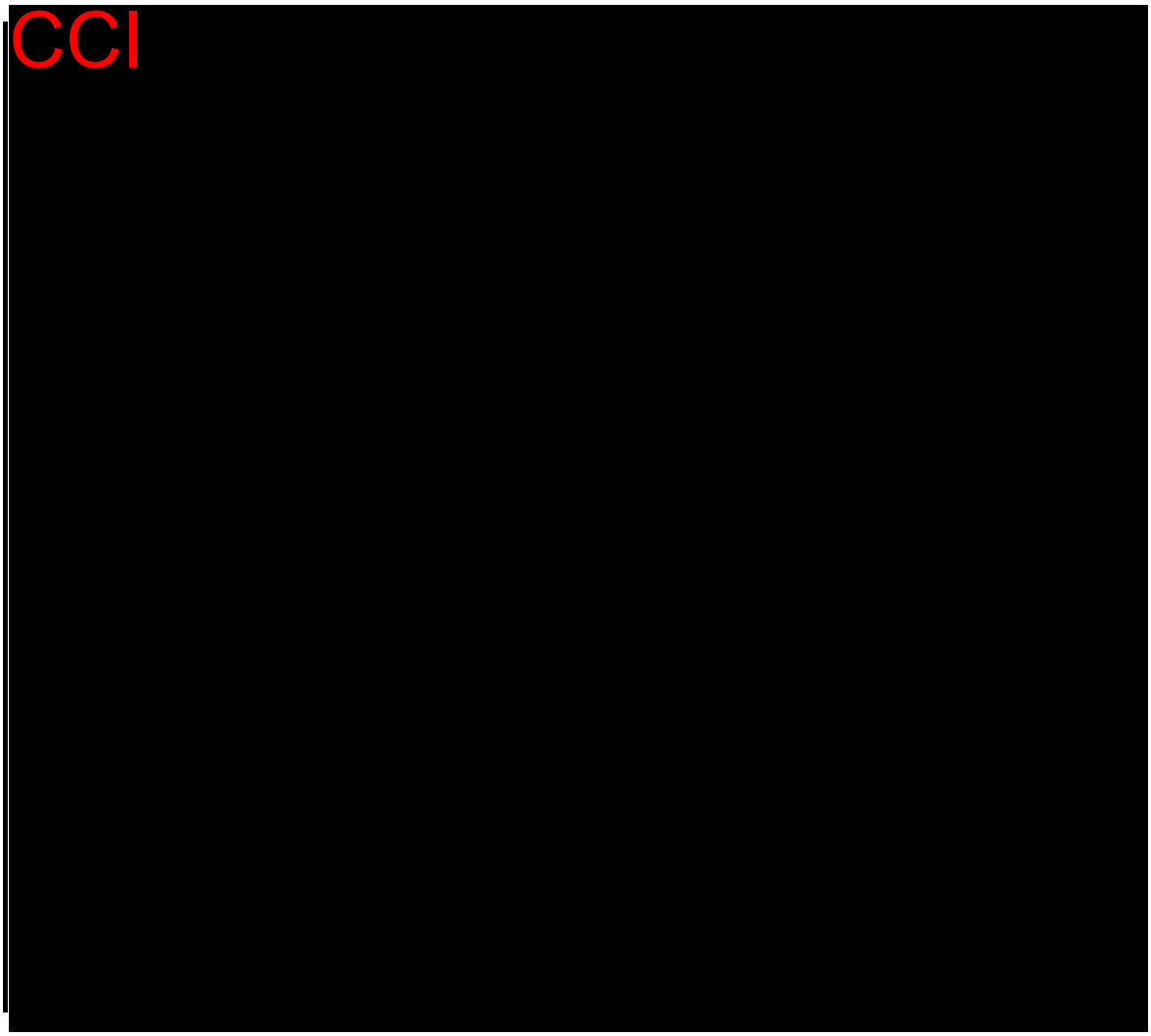
SUICIDAL IDEATION	
Ask questions 1 and 2. If both are negative, proceed to "Suicidal Behavior" section. If the answer to question 2 is "yes", ask questions 3, 4 and 5. If the answer to question 1 and/or 2 is "yes", complete "Intensity of Ideation" section below.	**Since Last Visit**
1. Wish to be Dead Subject endorses thoughts about a wish to be dead or not alive anymore, or wish to fall asleep and not wake up. ***Have you wished you were dead or wished you could go to sleep and not wake up?*** If yes, describe:	**Yes** ☐ **No** ☐
2. Non-Specific Active Suicidal Thoughts General non-specific thoughts of wanting to end one's life/commit suicide (e.g., *"I've thought about killing myself"*) without thoughts of ways to kill oneself/associated methods, intent, or plan during the assessment period. ***Have you actually had any thoughts of killing yourself?*** If yes, describe:	**Yes** ☐ **No** ☐
3. Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act Subject endorses thoughts of suicide and has thought of at least one method during the assessment period. This is different than a specific plan with time, place or method details worked out (e.g., thought of method to kill self but not a specific plan). Includes person who would say, *"I thought about taking an overdose but I never made a specific plan as to when, where or how I would actually do it.....and I would never go through with it"*. ***Have you been thinking about how you might do this?*** If yes, describe:	**Yes** ☐ **No** ☐
4. Active Suicidal Ideation with Some Intent to Act, without Specific Plan Active suicidal thoughts of killing oneself and subject reports having some intent to act on such thoughts, as opposed to *"I have the thoughts but I definitely will not do anything about them"*. ***Have you had these thoughts and had some intention of acting on them?*** If yes, describe:	**Yes** ☐ **No** ☐
5. Active Suicidal Ideation with Specific Plan and Intent Thoughts of killing oneself with details of plan fully or partially worked out and subject has some intent to carry it out. ***Have you started to work out or worked out the details of how to kill yourself? Do you intend to carry out this plan?*** If yes, describe:	**Yes** ☐ **No** ☐
INTENSITY OF IDEATION	
The following features should be rated with respect to the most severe type of ideation (i.e.,1-5 from above, with 1 being the least severe and 5 being the most severe). ***Most Severe Ideation:*** _____ _____ ***Type # (1-5)*** ***Description of Ideation***	Most Severe
Frequency ***How many times have you had these thoughts?*** (1) Less than once a week (2) Once a week (3) 2-5 times in week (4) Daily or almost daily (5) Many times each day	___
Duration ***When you have the thoughts how long do they last?*** (1) Fleeting - few seconds or minutes (2) Less than 1 hour/some of the time (3) 1-4 hours/a lot of time (4) 4-8 hours/most of day (5) More than 8 hours/persistent or continuous	___
Controllability ***Could/can you stop thinking about killing yourself or wanting to die if you want to?*** (1) Easily able to control thoughts (2) Can control thoughts with little difficulty (3) Can control thoughts with some difficulty (4) Can control thoughts with a lot of difficulty (5) Unable to control thoughts (0) Does not attempt to control thoughts	___
Deterrents ***Are there things - anyone or anything (e.g., family, religion, pain of death) - that stopped you from wanting to die or acting on thoughts of committing suicide?*** (1) Deterrents definitely stopped you from attempting suicide (2) Deterrents probably stopped you (3) Uncertain that deterrents stopped you (4) Deterrents most likely did not stop you (5) Deterrents definitely did not stop you (0) Does not apply	___
Reasons for Ideation ***What sort of reasons did you have for thinking about wanting to die or killing yourself? Was it to end the pain or stop the way you were feeling (in other words you couldn't go on living with this pain or how you were feeling) or was it to get attention, revenge or a reaction from others? Or both?*** (1) Completely to get attention, revenge or a reaction from others (2) Mostly to get attention, revenge or a reaction from others (3) Equally to get attention, revenge or a reaction from others and to end/stop the pain (4) Mostly to end or stop the pain (you couldn't go on living with the pain or how you were feeling) (5) Completely to end or stop the pain (you couldn't go on living with the pain or how you were feeling) (0) Does not apply	___

SUICIDAL BEHAVIOR *(Check all that apply, so long as these are separate events; must ask about all types)*	**Since Last Visit**
Actual Attempt: A potentially self-injurious act committed with at least some wish to die, *as a result of act*. Behavior was in part thought of as method to kill oneself. Intent does not have to be 100%. If there is ***any*** intent/desire to die associated with the act, then it can be considered an actual suicide attempt. ***There does not have to be any injury or harm***, just the potential for injury or harm. If person pulls trigger while gun is in mouth but gun is broken so no injury results, this is considered an attempt. Inferring Intent: Even if an individual denies intent/wish to die, it may be inferred clinically from the behavior or circumstances. For example, a highly lethal act that is clearly not an accident so no other intent but suicide can be inferred (e.g., gunshot to head, jumping from window of a high floor/story). Also, if someone denies intent to die, but they thought that what they did could be lethal, intent may be inferred. ***Have you made a suicide attempt?*** ***Have you done anything to harm yourself?*** ***Have you done anything dangerous where you could have died?*** ***What did you do?*** ***Did you_____as a way to end your life?*** ***Did you want to die (even a little) when you_____?*** ***Were you trying to end your life when you_____?*** ***Or Did you think it was possible you could have died from_____?*** ***Or did you do it purely for other reasons / without ANY intention of killing yourself (like to relieve stress, feel better, get sympathy, or get something else to happen)?*** (Self-Injurious Behavior without suicidal intent) If yes, describe: **Has subject engaged in Non-Suicidal Self-Injurious Behavior?**	**Yes** ☐ **No** ☐ Total # of Attempts _____ **Yes** ☐ **No** ☐
Interrupted Attempt: When the person is interrupted (by an outside circumstance) from starting the potentially self-injurious act *(if not for that, actual attempt would have occurred)*. Overdose: Person has pills in hand but is stopped from ingesting. Once they ingest any pills, this becomes an attempt rather than an interrupted attempt. Shooting: Person has gun pointed toward self, gun is taken away by someone else, or is somehow prevented from pulling trigger. Once they pull the trigger, even if the gun fails to fire, it is an attempt. Jumping: Person is poised to jump, is grabbed and taken down from ledge. Hanging: Person has noose around neck but has not yet started to hang - is stopped from doing so. ***Has there been a time when you started to do something to end your life but someone or something stopped you before you actually did anything?*** If yes, describe:	**Yes** ☐ **No** ☐ Total # of interrupted _____
Aborted Attempt: When person begins to take steps toward making a suicide attempt, but stops themselves before they actually have engaged in any self-destructive behavior. Examples are similar to interrupted attempts, except that the individual stops him/herself, instead of being stopped by something else. ***Has there been a time when you started to do something to try to end your life but you stopped yourself before you actually did anything?*** If yes, describe:	**Yes** ☐ **No** ☐ Total # of aborted _____
Preparatory Acts or Behavior: Acts or preparation towards imminently making a suicide attempt. This can include anything beyond a verbalization or thought, such as assembling a specific method (e.g., buying pills, purchasing a gun) or preparing for one's death by suicide (e.g., giving things away, writing a suicide note). ***Have you taken any steps towards making a suicide attempt or preparing to kill yourself (such as collecting pills, getting a gun, giving valuables away or writing a suicide note)?*** If yes, describe:	**Yes** ☐ **No** ☐
Suicidal Behavior: Suicidal behavior was present during the assessment period?	**Yes** ☐ **No** ☐
Suicide:	**Yes** ☐ **No** ☐
Answer for Actual Attempts Only	Most Lethal Attempt Date:
Actual Lethality/Medical Damage: 0. No physical damage or very minor physical damage (e.g., surface scratches). 1. Minor physical damage (e.g., lethargic speech; first-degree burns; mild bleeding; sprains). 2. Moderate physical damage; medical attention needed (e.g., conscious but sleepy, somewhat responsive; second-degree burns; bleeding of major vessel). 3. Moderately severe physical damage; *medical* hospitalization and likely intensive care required (e.g., comatose with reflexes intact; third-degree burns less than 20% of body; extensive blood loss but can recover; major fractures). 4. Severe physical damage; *medical* hospitalization with intensive care required (e.g., comatose without reflexes; third-degree burns over 20% of body; extensive blood loss with unstable vital signs; major damage to a vital area). 5. Death	*Enter Code* _____
Potential Lethality: Only Answer if Actual Lethality=0 Likely lethality of actual attempt if no medical damage (the following examples, while having no actual medical damage, had potential for very serious lethality: put gun in mouth and pulled the trigger but gun fails to fire so no medical damage; laying on train tracks with oncoming train but pulled away before run over). 0 = Behavior not likely to result in injury 1 = Behavior likely to result in injury but not likely to cause death 2 = Behavior likely to result in death despite available medical care	*Enter Code* _____

13.10. CCI

CCI

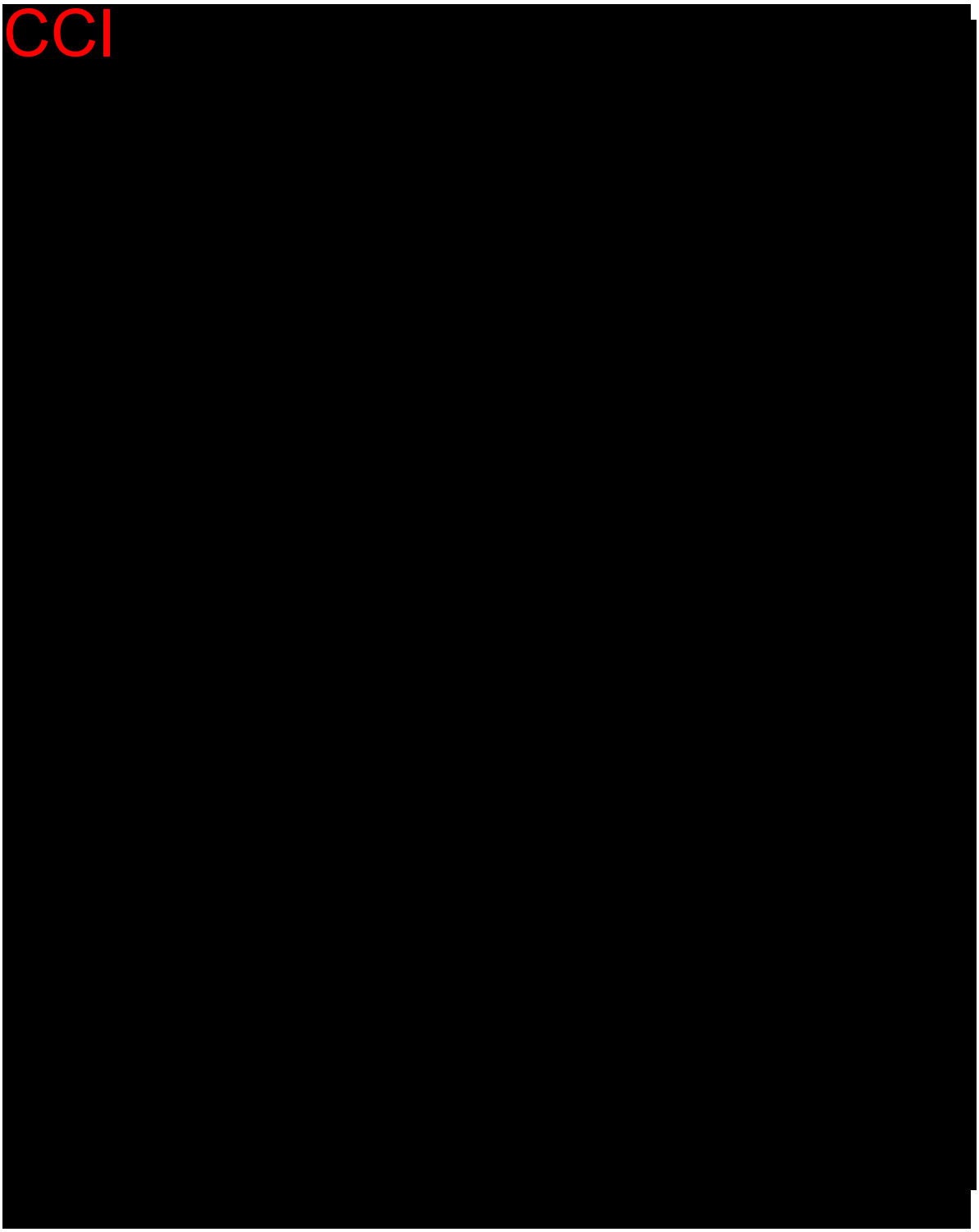



CCI

CCI

[redacted]


CCI


CCI


CCI

13.11. CTCAE Grading Details from the Central Laboratory

CTCAE CHECKS

NCI Common Terminology Criteria (CTC) for Adverse Events (AE)
Laboratory Tests

The NCI CTC for AE is a descriptive terminology which can be utilized for AE reporting. A grading (severity) scale is provided for each AE term. (CTCAE, Version 4.03)

Grade I: Mild AE
Grade II: Moderate AE
Grade III: Severe AE
Grade IV: Life-threatening or disabling AE

LLN = Lower Limit of Normal range
ULN = Upper Limit of Normal range

NOTES: The ranges in this table are adjusted to the precision and units used by BARC. In case the grading ranges are in absolute values, the ranges are listed in the different units in which the parameter can be reported. In case the grading ranges are expressed as a factor of the lower and/or upper limit of the normal ranges, no units are shown. It is possible that in some cases, the lower or upper limit of normal range is at the same time a Grade I AE.
Please note CTC checks only apply on parameters taken within the scope of the same visit.
More information about NCI CTC can be accessed on the CTEP website "http://ctep.cancer.gov/reporting/ctc.html"

Parameter	Grade I	Grade II	Grade III	Grade IV	Units
		HEMATOLOGY			
CD4 count	500 - < LLN[4]	200 - < 500	50 - < 200	< 50	/mm³
Hemoglobin	10.0 - < LLN	8.0 - < 10.0	< 8.0	-	g/dL
	100 - < LLN	80 - < 100	< 80	-	g/L
	ULN + > 0.0-2.0	ULN + > 2.0-4.0	ULN + > 4.0	-	g/dL
	ULN + > 0-20	ULN + > 20-40	ULN + > 40	-	g/L
Leukocytes	3.0 - < LLN	2.0 - < 3.0	1.0 - < 2.0	< 1.0	10^3/µL or 10^9/L
	-	-	> 100.0	-	10^3/µL or 10^9/L
Lymphocytes	0.80 - < LLN	0.50 - < 0.80	0.20 - < 0.50	< 0.20	10^3/µL or 10^9/L
	-	> 4.00 – 20.00	> 20.00	-	10^3/µL or 10^9/L
Neutrophils	1.50 - < LLN	1.00 - < 1.50	0.50 - < 1.00	< 0.50	10^3/µL or 10^9/L
Platelets	75 - < LLN	50 - < 75	25 - < 50	< 25	10^3/µL or 10^9/L
		COAGULATION			
Fibrinogen	0.75xLLN - < LLN	0.5xLLN - < 0.75xLLN	0.25xLLN - < 0.5xLLN	< 0.25xLLN	
PT (INR)	> ULN - 1.5xULN	> 1.5xULN – 2.5xULN	> 2.5xULN	-	
APTT	> ULN - 1.5xULN	> 1.5xULN – 2.5xULN	> 2.5xULN	-	
		BIOCHEMISTRY			
Albumin	3.0 - < LLN	2.0 - < 3.0	< 2.0	-	g/dL
	30 - < LLN	20 - < 30	< 20	-	g/L
Alkaline phosphatase	> ULN - 2.5xULN	> 2.5xULN - 5xULN	> 5xULN - 20xULN	> 20xULN	
ALT	> ULN – 3xULN	> 3xULN - 5xULN	> 5xULN - 20xULN	> 20xULN	
Amylase	> ULN - 1.5xULN	> 1.5xULN - 2xULN	> 2xULN - 5xULN	> 5xULN	
AST	> ULN – 3xULN	> 3xULN - 5xULN	> 5xULN - 20xULN	> 20xULN	
Bilirubin, total	> ULN - 1.5xULN	> 1.5xULN - 3xULN	> 3xULN - 10xULN	> 10xULN	

Parameter	Grade I	Grade II	Grade III	Grade IV	Units
Corrected calcium	8.0 - < LLN	7.0 - < 8.0	6.0 - < 7.0	< 6.0	mg/dL
	2.00 - < LLN	1.75 - < 2.00	1.50 - < 1.75	< 1.50	mmol/L
	> ULN - 11.5	> 11.5 - 12.5	> 12.5 - 13.5	> 13.5	mg/dL
	> ULN - 2.90	> 2.90 - 3.10	> 3.10 - 3.40	> 3.40	mmol/L
Cholesterol, total	> ULN - 300	> 300 - 400	> 400 - 500	> 500	mg/dL
	> ULN - 7.75	> 7.75 - 10.34	> 10.34 - 12.92	> 12.92	mmol/L
Creatine kinase	> ULN - 2.5xULN	> 2.5xULN - 5xULN	> 5xULN - 10xULN	> 10xULN	
Creatinine (serum)	> ULN - 1.5xULN	> 1.5xULN - 3xULN	> 3xULN - 6xULN	> 6xULN	
eGFR	60 - < LLN[4]	30 - 59	15 - 29	< 15	mL/min (/1.73m^2)
GGT	> ULN - 2.5xULN	> 2.5xULN - 5xULN	> 5xULN - 20xULN	> 20xULN	
Glucose (2h PP; random)	55 - < LLN[1]	40 - < 55	30 - < 40	< 30	mg/dL
	3.00 < LLN[1]	2.20- < 3.00	1.70- < 2.20	< 1.70	mmol/L
Glucose (fasting)	55 - < LLN	40 - < 55	30 - < 40	< 30	mg/dL
	3.00 < LLN	2.20- < 3.00	1.70- < 2.20	< 1.70	mmol/L
	> ULN - 160	> 160 - 250	> 250 - 500	> 500	mg/dL
	> ULN - 8.90	> 8.90 - 13.90	> 13.90 - 27.80	> 27.80	mmol/L
Haptoglobin	< LLN	-	-	-	
Lipase	> ULN - 1.5xULN	> 1.5xULN - 2xULN	> 2xULN - 5xULN	> 5xULN	
Magnesium	1.2 - < LLN	0.9 - < 1.2	0.7 - < 0.9	< 0.7	mg/dL
	0.50 - < LLN	0.40 - < 0.50	0.30 - < 0.40	< 0.30	mmol/L
	> ULN – 3.0	-	> 3.0 – 8.0	> 8.0	mg/dL
	> ULN - 1.23	-	> 1.23 - 3.30	> 3.30	mmol/L
Phosphorus	2.5 - < LLN	2.0 - < 2.5	1.0 - < 2.0	< 1.00	mg/dL
	0.80 - < LLN	0.60 - < 0.80	0.30 - < 0.60	< 0.30	mmol/L
Potassium	3.0 - < LLN[2]	3.0 - < LLN[2]	2.5 - < 3.0	< 2.5	mmol/L
	> ULN - 5.5	> 5.5 - 6.0	> 6.0 - 7.0	> 7.0	mmol/L
Sodium	130 - < LLN	-	120 - < 130	< 120	mmol/L
	> ULN - 150	> 150 - 155	> 155 - 160	> 160	mmol/L
Triglycerides	150 - 300	> 300 - 500	> 500 - 1000	> 1000	mg/dL
	1.71 - 3.42	> 3.42 - 5.70	> 5.70 - 11.40	> 11.40	mmol/L
Troponin I	> 0.01 - < 0.30 (♀)		≥ 0.30 (♀ & ♂)		ng/mL or µg/L
	> 0.03 - < 0.30 (♂)				ng/mL or µg/L
Troponin T	> 0.01 - <0.10	-	≥ 0.10	-	ng/mL or mg/L
Uric acid	> ULN - 10.0[3]	-	> ULN - 10.0[3]	> 10.0	mg/dL
	> ULN - 590[3]	-	> ULN - 590[3]	> 590	µmol/L
URINALYSIS					
Creatinine clearance	60 - < LLN	30 - 59	15 - 29	< 15	mL/min
Blood	+(++)	-	-	-	
Protein	+	++ or +++ (≥ 18y)	-	-	
	< 1000	1000 – 3499 (≥18y)	≥ 3500 (≥ 18y)	-	mg/24h
Protein/creatinine ratio (< 18y)		0.5 – 1.9	> 1.9		

[1] Glucose: BARC reference range has no LLN, therefore 74 mg/dL (4.11 mmol/L) is applied as LLN in Grade I flagging (reference Tietz Clinical guide to laboratory tests 4th ed.)
[2] Potassium: Grade II when intervention is indicated, else Grade I
[3] Uric acid: Grade III, when in combination with physiologic consequences, else Grade I
[4] CD4 count & eGFR: LLN is lower or equal to lower threshold for Grade I flagging, therefore only the exact lower threshold value is flagged as Grade I

13.12. Table of Contents for Tables, Listings, and Figures

The table of contents for tables, listings, and figures will be provided in a corresponding mock shell document.

13.13. Other SAP(s) Describing Analyses not Covered in the SAP

SAPs describing analyses (pharmacokinetics [PK], PK/pharmacodynamics[PD], etc.) not covered within this SAP will be included within this Appendix.

BOS161721-02: A RANDOMIZED DOUBLE-BLIND PHASE 1b/2 COMBINED STAGGERED MULTIPLE DOSE ESCALATION STUDY OF BOS161721 IN SYSTEMIC LUPUS ERYTHEMATOSUS (SLE) PATIENTS ON A BACKGROUND OF LIMITED STANDARD OF CARE

Statistical Analysis Plan

VERSION 2.0
DATE OF PLAN:

29Nov2018

STUDY DRUG:
BOS161721

PREPARED FOR:

Boston Pharmaceuticals, Inc.

Approval Signature Page: Array Biostatistics, LLC

Document Author: PPD, MS Senior Biostatistician	Date
Document Reviewer: PPD, MS Vice President, Biostatistics & Statistical Programming	Date

Approval Signature Page: Boston Pharmaceuticals, Inc.

PPD

PPD

Date

Clinical Operations Lead

PPD

PPD

Date

Vice President, Clinical Development

Contents

ABBREVIATIONS

ACR	American College of Rheumatology
ADA	Anti-drug antibody
AE	Adverse event
AESI	Adverse event of special interest
ANA	Antinuclear antibody
ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
APL	Antiphospholipid
AUC	Area under the curve
BICLA	BILAG-based Composite Lupus Assessment
BILAG	British Isles Lupus Assessment Group
BP	Blood pressure
C3	Complement C3
C4	Complement C4
C_L	Systematic clearance
CLASI	Cutaneous Lupus Area and Severity Index
C_{max}	Maximum plasma concentration
CPAP	Clinical Pharmacology Analysis Plan
CS	Corticosteroid
CSR	Clinical study report
C-SSRS	Columbia Suicide Severity Rating Scale
CTCAE	Common Terminology Criteria for Adverse Events
CV	Coefficient of variation
DLT	Dose-limiting toxicity
DMC	Data Monitoring Committee
dsDNA	Double-stranded deoxyribonucleic acid
ECG	Electrocardiogram
CCI	[redacted]
FAS	Full analysis set
HR	Hazard ratio
IA	Interim Analysis
IEC	Independent Ethics Committee
IL-21	interleukin 21
IRB	Institutional Review Board
ISR	Injection site reaction
IWRS	Interactive web response system
LLN	Lower limit of normal
LLOQ	Lower limit of quantification
MAD	Multiple Ascending Dose
MedDRA	Medical Dictionary for Medical Affairs
nAb	Neutralizing antibody
NCI	National Cancer Institute
PD	Pharmacodynamics

PGA	Physician's Global Assessment
PK	Pharmacokinetics
POC	Proof of concept
PP	Per protocol set
pSTAT3	Phosphorylated signal transducer and activator of transcription 3
PT	Preferred Term
SAD	Single ascending dose
SAE	Serious adverse events
SAP	Statistical analysis plan
SAS	Statistical analysis system
SC	Subcutaneously
CCI	[redacted]
SLE	Systemic Lupus Erythematosus
SLEDAI-2K	SLE Disease Activity Index 2000
SLICC	Systemic Lupus International Collaborating Clinics
Sm	Smith
SOC	System Organ Class
SRI-4	SLE Responder Index 4
SRI-5	SLE Responder Index 5
SRI-6	SLE Responder Index 6
SS	Safety analysis set
SSA	Sjögren syndrome-A (Ro)
SSB	Sjögren syndrome-B (La)
$t_{1/2}$	Terminal elimination half-life
TEAE	Treatment emergent adverse event
T_{max}	First time to maximum concentration
TTMF	Time to medication failure
ULN	Upper limit of normal
V_d	Volume of distribution
WHO	World Health Organization

1. INTRODUCTION

The statistical analysis plan (SAP) details the planned analyses that will be included in the Clinical Study Report (CSR) of study number BOS161721-02: A Randomized Double-Blind Phase 1b/2 Combined Staggered Multiple Dose Escalation Study of BOS161721 in Systemic Lupus Erythematosus (SLE) Patients on a Background of Limited Standard of Care. The content of this SAP is based on the protocol Version 4.0 (Amendment 3) dated 27July2018.

Revision Chronology:

V1.0	05SEP2018	Original

2. STUDY OBJECTIVES AND ENDPOINTS

2.1. Multiple Ascending Dose (MAD) Phase 1b Study

Objectives	Endpoints
Primary	
• To assess safety, tolerability, and immunogenicity of repeat doses of BOS161721 (20, 60, and 120 mg) administered subcutaneously (SC) in adult patients with moderately to severely active SLE on limited background standard of care treatment, in order to estimate the optimal dose.	Safety Endpoints • Incidence and severity of adverse events (AEs) and serious adverse events (SAEs), related AEs, AEs leading to study drug discontinuation, AEs by severity and relatedness • Injection site reactions • Columbia Suicide Severity Rating Scale (C-SSRS) • 12-lead electrocardiograms (ECGs) parameter results at each visit and change from baseline • Vital signs (blood pressure [BP], heart rate, and temperature) parameter results at each visit and change from baseline • Clinical laboratory results and change from baseline • Physical examinations changes from baseline • Anti-drug antibodies (ADAs) • Study drug exposure/compliance

Secondary	
• To characterize the PK of BOS161721 and select the optimal dose of BOS161721 based on safety, PK, and PD effects in patients with mild to moderate SLE.	Pharmacokinetic (PK) Endpoints • BOS161721 concentration by visit and time point • Maximum observed concentration (C_{max}), first time to maximum concentration (T_{max}), area under the concentration-time curve (AUC), $t_{½}$, systematic clearance (C_L), volume of distribution (V_d) Pharmacodynamic (PD) Endpoints • Results and changes (or shifts) from baseline to each visit in phosphorylated signal transducer and activator of transcription 3 (pSTAT3), C3 and C4 levels, and leukocyte immunophenotype • Results and changes (or shifts) from baseline in anti-double-stranded DNA (dsDNA), antinuclear antibodies (ANA), anti-Sjögren syndrome A and B (SSA, SSB), Smith (Sm), and antiphospholipid (APL) autoantibodies at each visit • Results and changes (or shifts) from baseline in abrogation of IL-21 gene signature at each indicated visit
Exploratory	
• CCI [redacted]	CCI [redacted]

2.2. Proof of Concept (POC) Phase 2 Study

Objectives	Endpoints
Primary	
• To demonstrate a superior effect of BOS161721 at the chosen dose compared with placebo for response on the SLE Responder Index 4 (SRI-4)	Primary Efficacy Endpoint • The proportion of patients with a SRI-4 response at Day 210

Secondary	
• To demonstrate a superior effect of BOS161721 at the chosen dose compared with placebo for response on clinical indicators of SLE activity, in adult patients with moderately to severely active SLE on limited background standard of care treatment	Secondary Efficacy Endpoints • The proportion of patients with: - SRI-4 response at each visit - SRI-5 and SRI-6 response at each visit - a sustained reduction of oral corticosteroid (CS) (≤ 10 mg/day and ≤ Day 0 dose) between Day 120 and Day 210 - new BILAG A flare or > 1 BILAG B flares relative to baseline through Day 210 - PGA worsening - a BICLA response - a CLASI response - medication failures • Results and changes from baseline in: - CLASI - swollen and tender joints ACR-28 - SLEDAI-2K - SLICC/ACR damage index • Time to medication failure • Duration of longest SRI-4 response • Time to first SRI-4 response • Time to BILAG A flare or > 1 BILAG B flare compared to baseline through Day 210
Safety	
• To assess safety and tolerability of repeat doses of BOS161721 (20, 60, and 120 mg) administered SC in adult patients with moderately to severely active SLE on limited background standard of care treatment	Safety Endpoints • Incidence and severity of adverse events (AEs) and serious adverse events (SAEs), related AEs, AEs leading to study drug discontinuation, AEs by severity and relatedness • Injection site reactions • C-SSRS • 12-lead ECGs parameter results at each visit and change from baseline • Vital signs (blood pressure [BP], heart rate, and temperature) parameter results at each visit and change from baseline

	• Clinical laboratory results and change from baseline • Physical examinations changes from baseline • ADAs • Study drug exposure/compliance
Exploratory	
• CCI	CCI
• CCI	• CCI
• CCI	• CCI

The order in which the endpoints for this trial will be discussed in subsequent sections is based on the phase 2 POC portion of the study. The phase 1b MAD endpoints will also be discussed within the context of the phase 2 POC portion of the study but will appear out of the specified order in the above table in an effort to maintain common concepts across both phases in one location within the SAP.

3. STUDY DESIGN

3.1. Study Design and Population

This is a Phase 1b/2 combined, randomized, multicenter, double-blind, placebo-controlled trial to study the clinical efficacy, safety, and pharmacokinetics (PK) of subcutaneously (SC) doses of BOS161721 in adult patients with moderately to severely active SLE. After successfully

completing a screening phase, eligible patients will be randomized to a specified dose of BOS161721 or placebo. The dosing schedule will be monthly.

The trial will consist of 2 double blinded- portions: MAD Phase 1b and POC Phase 2. Patients may receive a total of 7 SC monthly doses of study drug on Days 0, 30, 60, 90, 120, 150, and 180, followed by safety/efficacy follow-up visit at Days 210, and safety follow up visits at Days 240, and 270.

SLE disease activity assessment data will be centrally reviewed by medical monitor and sponsor to ensure the scores are clinically meaningful and compliant with specific definition. The scope of responsibility includes, but is not limited to, review and confirmation of "A" and "B" British Isles Lupus Assessment Group (BILAG) system organ disease, confirmation of clinical components of the SLE Disease Activity Index 2000 (SLEDAI-2K) score at screening and during the study, and cross validation- of the instruments used in this study to assess the disease activity. Further details on the content and methods of data reports by the medical monitor and sponsor will be outlined in the Medical Data Review Plan along with the processes and procedures that will be followed.

3.1.1. Multiple Ascending Dose Phase 1b

The MAD portion will consist of 3 cohorts:

- Cohort 1 (20 mg SC) will include 6 patients
    - 5 patients will receive BOS161721 (active group) and 1 patient will receive placebo (placebo group)
- Cohorts 2 (60 mg SC) and 3 (120 mg SC) will include 12 patients each
    - 9 patients in the active group and 3 in the placebo group

Doses selected for each of the 3 cohorts is based on a 90-day safety, tolerability, PK and PD data review from the Phase 1 single ascending dose (SAD) study (BOS161721-01) in healthy subjects. All doses selected for the MAD part of the study are projected not to exceed the mean exposure of that achieved in the SAD study.

The MAD portion of the study design is staggered, where after the 6 patients in Cohort 1 have received 2 doses and have completed 2 weeks of follow up- after the second dose, Cohort 2 begins dosing (after the data monitoring committee (DMC) evaluation of the safety and tolerability data from Cohort 1). Similarly, after 8 of the 12 patients in Cohort 2 have received 2 doses and have completed 2 weeks of follow up- after the second dose, Cohort 3 begins dosing after the DMC evaluation of the safety and tolerability data from Cohorts 1 and 2. Each cohort will continue at their assigned dose level through their respective 6-month treatment periods (See Study Schematic in the study protocol). If patients discontinue the study in a cohort prior to adequate safety follow-up, he/she may be replaced.

Criteria for dose escalation are further described in the DMC Charter.

3.1.2. Proof of Concept Phase 2

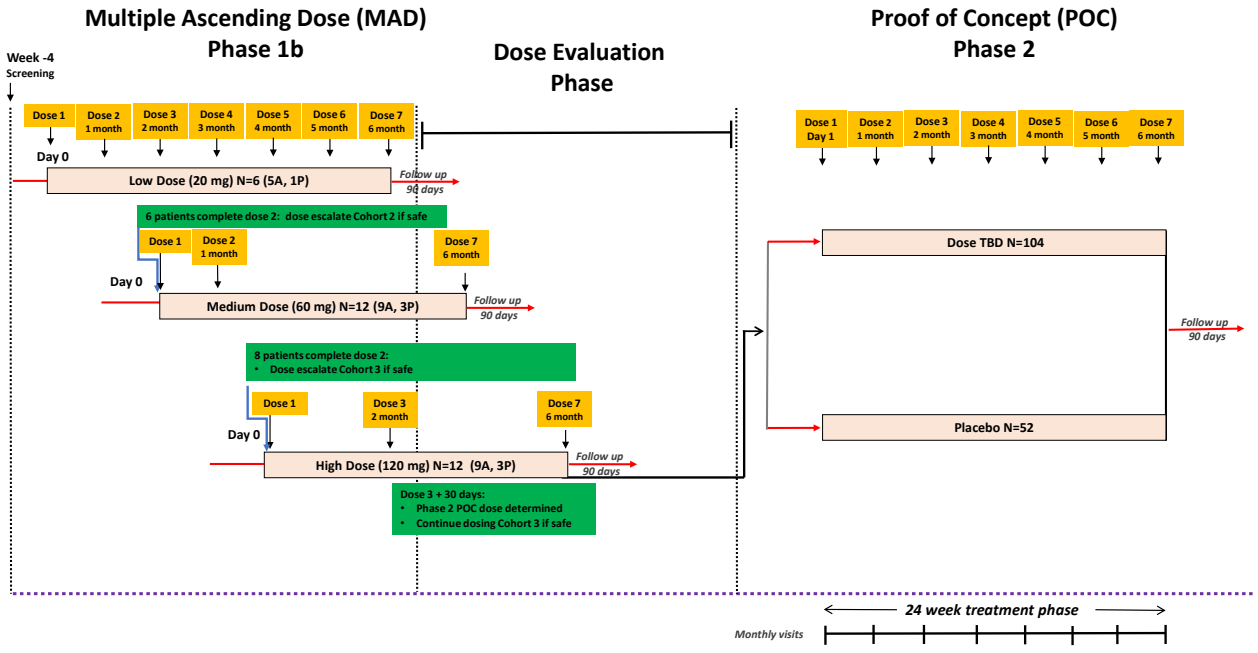
The optimal dose is chosen based on MAD Phase 1b safety, tolerability, immunogenicity, and PK and PD data. The dose will be communicated by a letter to site investigators participating in the POC Phase 2 portion, and to the Institutional Review Board / Independent Ethics Committee (IRB/IEC).

For the POC part of the study, approximately 156 additional patients will be randomized to active or placebo groups in a 2:1 ratio.

As in the MAD part of the study, each patient in the POC portion may receive a total of 7 SC monthly doses of study drug on Days 0, 30, 60, 90, 120, 150, and 180. Assessments will be completed according to the Schedule of Assessments detailed in the study protocol. Dose selection will be based on the DMC and sponsor's assessment of safety, tolerability, immunogenicity, and PK and PD data from the MAD portion of 30 patients treated with BOS161721/placebo.

DMC safety reviews will be conducted periodically throughout the study as described in the DMC charter.

Study Diagram for MAD Phase 1b and POC Phase 2



A = active drug (BOS161721); P = placebo

3.2. Randomization and Blinding

This is a randomized, double-blind study.

Patients meeting all inclusion and exclusion criteria will be centrally randomized to either placebo or BOS161721 using an IWRS according to a randomization list generated by an independent, non-study statistician. Randomization will be performed separately for each study portion and separately for each cohort in the Phase 1b and 2 portions as follows:

Phase/Cohort	Number of Patients	Randomization Ratio BOS161721:Placebo
Phase 1b/Cohort 1	6	5:1
Phase 1b/Cohort 2	12	3:1
Phase 1b/Cohort 3	12	3:1
Phase 2	156*	2:1

*Additional patients may be enrolled to ensure sufficient numbers of patients are in the full analysis set (FAS).

Eligible patients will be assigned to the study portion which is active at time of enrollment. Similarly, patients in the Phase 1 MAD will be assigned to the cohort which is active. Each patient will be assigned a unique randomization number which will not be reused.

All patients, investigators, and study team participants will be blinded to treatment assignment. An independent biostatistician not otherwise involved on the study will be unblinded and prepare materials for the interim analysis (IA) and the DMC safety reviews. The DMC will review unblinded data during safety reviews and the IA. A limited team at Boston Pharmaceuticals will review unblinded results from the Phase 1b MAD portion during the IA to determine the dose that will be used for the Phase 2 POC portion of the study. Details regarding maintenance of the blinding and content of data reviews will be described in the DMC charter or related study documentation.

3.3. Sample Size Considerations

Sample size in the Phase 1b MAD part of the study is based on operational consideration.

The sample size in the Phase 2 POC study is based on the primary endpoint, SRI-4 Response at Day 210. Approximately 156 additional patients will be randomized in a 2:1 ratio to BOS161721 versus placebo to achieve 132 evaluable patients in the FAS. This assumes 24 patients will drop prior to having received treatment and an evaluable post-baseline efficacy measure. Enrollment and randomization will be monitored in a blinded fashion and increased if needed to ensure at least 132 patients are randomized into the FAS.

A total of 132 patients randomized provides more than 80% power to detect a treatment difference of 25% in SRI-4 response rates at Day 210, based on a targeted 2-sided significance level of 10% and using a 2-sided Pearson's chi-squared test. This assumes a response rate of 65% for BOS161721 and 40% for placebo, where missing data at Day 210 is carried forward

from most recent non-missing result and patients are treated as a non-responder if a prohibited medication defined as a medication failure occurs.

3.4. Data Monitoring Committee (DMC)

This study will use an external DMC. The DMC is an independent committee established to provide oversight of safety considerations in the study and to provide advice to the sponsor regarding actions the committee deems necessary for the continuing protection of enrolled patients and those yet to be recruited to the trial as well as for the continuing validity and scientific merit of the study results. The DMC is charged with assessing such actions in light of an acceptable benefit/risk profile for BOS161721. The recommendations made by the DMC (i.e., dose escalation, etc.) will be forwarded to the sponsor for final decision. The sponsor will forward such decisions, which may include summaries of safety data which are not endpoints, to regulatory authorities, as appropriate.

The sponsor will appoint a DMC for the periodic review of available study data. The DMC is an independent group of experts that advises the sponsor and the study investigators. The members of the DMC serve in an individual capacity and provide their expertise and recommendations. The primary responsibilities of the DMC are to (1) periodically review and evaluate the accumulated study data for participant safety, study conduct, and progress, (2) make recommendations to the sponsor concerning the continuation, modification, or termination of the study and (3) suggest dose for POC portion of the study as described in the study protocol.

The DMC considers study-specific data as well as relevant background knowledge about the disease, test agent, or patient population under study. The DMC is responsible for defining its deliberative processes, including event triggers that would call for an unscheduled review, stopping guidelines, unmasking (unblinding), and voting procedures prior to initiating any data review. The DMC is also responsible for maintaining the confidentiality of its internal discussions and activities as well as the contents of reports provided to it.

The DMC will have access to unblinded treatment information during the clinical trial. Details regarding management and process of this committee are found in the DMC Charter.

The DMC may recommend termination of BOS161721 treatment arm or the entire BOS161721 MAD/POC trial for any safety concern that is felt to outweigh potential benefits. The recommendation must be supported by the sponsor as indicated in the DMC Charter.

3.5. Interim Analysis (IA)

One interim analysis is planned for this study and will be conducted at the time of the POC decision for dose selection. The POC dose will be selected based on the MAD data. Since there is no IA during the POC part of the study, there is no impact on the type 1 error.

Interim analysis outputs will be a subset of those to be performed for the final analysis.

3.6. Timing of Analyses

The following analyses are planned:

- An IA will be performed during the last cohort of the MAD portion to determine dose selection for the POC portion.
- The final analysis will be performed when all patients have either completed the POC safety follow up or withdrawn from the study.
- Safety analyses will be performed for DMC reviews throughout the study to evaluate dose escalation decisions and accumulating safety data. The frequency and details of the content and format of the safety review meetings are described in the DMC charter and SAP/mock shells.

4. DATA ANALYSIS CONSIDERATIONS

All analyses will be conducted based on SAS 9.3 or higher.

All data in the database will be presented in by-patient data listings.

Unless otherwise stated, all listings will be sorted by study phase, treatment group, patient number, and assessment date (and time, if available). If assessment date is missing, chronological visit will be used for sorting.

Unless stated otherwise, continuous data will be summarized by treatment group based on number of patients (n), mean, median, standard deviation (SD), 25^{th} percentile/first quartile (Q1), 75^{th} percentile/third quartile (Q3), minimum value, and maximum value.

The geometric % coefficient of variation is calculated as 100*sqrt(exp(SDlog^2)-1) where SDlog is standard deviation of log-transformed values.

Unless stated otherwise, categorical data will be summarized by treatment group using n and percentage based on the number of non-missing values.

- The number of missing values will be presented as a separate category with no percentage, but only if one or more patients are missing data.
- Counts of zero will be presented without percentages.

Statistics will be presented in the summary tables based on the following:

- Minimum and Maximum: same number of significant digits as the raw value
- Mean, Median, Q1, and Q3: one additional decimal place to that reported for Minimum and Maximum
- SD: two additional decimal places than presented for the Minimum and Maximum

- Percentages: reported to one decimal place with the exception of 100% which will be displayed without decimals.
- P-values will be reported to four decimal places. If the value is below 0.0001 it will be noted as < 0.0001; if the value above 0.9999 it will be noted as > 0.9999.

Unless otherwise noted, statistical inference will be based on a 10% significance level (i.e. 90% confidence intervals will be produced).

Statistical testing will be performed on data from the Phase 2 study only.

All data up to the time of study completion/withdrawal from the study will be included in the analysis, regardless of duration of treatment.

Numbering for data displays will be based on ICH E3.

4.1. Stratification and Covariates

The effects of noncompliance, background therapy use, treatment discontinuations, premature withdrawal from study and covariates will be assessed to determine the impact on the general applicability of results from this study. Exploratory analyses of the data may be conducted as deemed appropriate to evaluate factors for which analyses are not controlled.

Continuous efficacy endpoints will be assessed via analysis of variance (ANOVA) or analysis of covariance (ANCOVA) when applicable, adjusting for the baseline value.

4.2. Evaluation of Subgroups

There are no formal plans for examining subgroups.

4.3. Multiple Comparisons and Multiplicity

The Type 1 error rate will not be adjusted for multiplicity.

5. GENERAL DATA HANDLING CONVENTIONS

5.1. Assigned and Actual Treatment

Patients are assigned to a study treatment and cohort based on the study schematic in Section 3.

Full analysis set (FAS) and per protocol (PP) analyses will be conducted on the basis of the randomized (assigned) treatment. Safety analyses will be conducted on the basis of the actual treatment received based on the exposure data collected in the case report form (CRF).

5.2. Reference Dates

- Screening date is defined as the CRF provided date on which a patient was screened for trial entry.
- Randomization date is defined as the date on which the patient is randomized to study treatment.
- Treatment start date is defined as the date of first dose of study drug.
- Treatment end date is defined as the date of last dose of study drug.
- Age will not be calculated and will come directly from the CRF. The CRF uses the informed consent date as its reference date for age calculation.
- Safety data, such as AEs and laboratory assessments will use the treatment start date as a reference date.
- Efficacy data will use the randomization date as a reference date.
- Study day will be based on treatment start date as a reference date.

5.3. Study Day and Duration Variables

Reference date calculations will be defined as follows, assuming non-missing dates:

- date of interest – reference date + 1 when the date of interest is on or after the reference date;
- otherwise, date of interest – reference date.

If either date is missing or incomplete, reference date calculations will not be performed. Date imputation will be performed as identified in Section 5.6.

Study day will either have a negative value if collected before dosing or a positive value if collected on or after the day of drug dosing; there will be no study day zero.

Duration of time is dependent on reference dates and will be calculated in a manner similar to that of the reference date calculation. Duration on study is defined as the end of study date – randomization date + 1. Duration of treatment is defined as treatment end date – treatment start date + 1, where treatment end date is the date of last dose of study drug.

Time-to-event endpoints are followed until first event or censoring. As a result, event time will be calculated as: event or censoring date – reference date + 1. These are further described in Section 7.

When reporting time-to-event data (Time to Medication Failure, Time to first SRI-4 Response, Time to BILAG A Flare or >1 BILAG B flare) or duration outcomes, the results (in days) above will be converted to an appropriate unit. When reporting in months it will be divided by 30.4375; for reporting in weeks it will be divided by 7; and for reporting in years it will be divided by 365.25.

5.4. Study Time Periods

Where applicable, data reporting will be classified by the following study periods for analysis:

- Pre-treatment is defined as the period prior to a patient's treatment start date.
- On-treatment is defined as the period between a patient's treatment start date and treatment end date, inclusive.
- Post-treatment is defined as the period of time following the on-treatment period.

5.5. Baseline, Post-Baseline Changes, and Last Observation Carried Forward (LOCF)

Unless stated otherwise, baseline and post-baseline change values will be based on the following:

- Baseline will be based on the last non-missing value collected prior to or on the treatment start date (and time, if applicable).
- Post-baseline values will be those collected after the date/time of first dose of study drug.
- Change from baseline is defined as: value – baseline value.
- Most extreme change: The maximum most extreme change will be the maximum post-baseline value; the minimum most extreme change will be the smallest post-baseline value. This calculation will consider all assessments collected after the first dose of study drug, scheduled or unscheduled.

LOCF will be based on the last non-missing value collected. This calculation will consider all assessments, scheduled or unscheduled, including baseline. If a patient withdraws from the study, their most recent non-missing assessment prior to withdraw will be carried forward. Additional details regarding LOCF for specific endpoints are provided in Section 7.

Observed values do not incorporate any imputations and are presented as collected.

5.6. Imputation of Partial Dates

Adverse Events

- If the AE start date is completely missing, or if the patient was not treated, no imputation will be conducted.
- If the AE start date is missing day and month, do the following:
    - If the treatment start date is missing or the AE start year does not fall in the same year as that of the treatment start date or if the AE record contains information to indicate that the event ended before the treatment start date (e.g. the AE end date month and year are earlier than the treatment start date or the full AE end date is

known and occurs earlier than the treatment start date), then set the AE start month and day to January 1st.
    - Otherwise, set the AE start date to the treatment start date.
- If only the AE start day is missing, do the following:
    - If the study treatment start date is missing or the AE start month and year does not fall in the same month/year as that of the treatment start date or if the AE contains information to indicate that the event ended before the treatment start date, then set the AE start month and day to the 1st day of the month of the treatment start date.
    - Otherwise, set the AE start date to the treatment start date.
- AE end dates will not be imputed.

Prior, Concomitant, and Subsequent Medications

- The imputation rules for AE start dates will be used for medication start dates.
- Medication stop dates will be imputed as follows:
    - If the stop date is only missing the day, then the stop day is the last day of the month
    - If the stop date is missing both the day and month and the year matches the last study date, then the stop month and day is the earlier of the last study date and December 31
    - If the stop date is missing both the day and month and the year is earlier than the last study date, then the stop month and day is December 31
    - If the stop date is completely missing, no imputation is performed, and the medication will be classified as a concomitant medication for subjects who were treated unless the medication began after the end of study treatment.

5.7. Multiple Assessments and Visit Windows

Nominal visits (e.g. those identified by the study CRF) will be the basis of summarization and statistical analysis; no visit date windowing will be conducted. Unscheduled data may be included in summaries of most extreme, baseline, and endpoint values; summaries of specific abnormalities any time post-baseline; and patient data listings

5.8. Treatment Group Display

Treatment groups will be displayed with the following columns. See the mock shells for additional details and a visual representation of the treatment group display.

- Phase 1b MAD: Cohort 1 BOS 20mg
- Phase 1b MAD: Cohort 2 BOS 60mg
- Phase 1b MAD: Cohort 3 BOS 120mg
- Phase 1b MAD: Placebo

- Phase 1b MAD: Total BOS (All patients treated with BOS161721, excludes placebo)
- Phase 2 POC: BOS <dose> where <dose> is the value determined during the IA
- Phase 2 POC: Placebo
- Overall: Total BOS <dose>mg where <dose> is the common dose between the Phase 1b MAD and Phase 2 POC portions of the study (to be determined during the IA). If a common dose is not determined, this column will be omitted.
- Overall: Total BOS (Patients treated with BOS161721 in MAD and POC, excludes placebo)
- Overall: Total Placebo (Placebo patients in MAD and POC, excludes patients treated with BOS161721)

5.9. Missing Data

Missing data imputation for the efficacy endpoints are discussed in the corresponding efficacy sections. AE and concomitant medication date imputations are described in Section 5.6. Missing data for noncompartmental analysis of pharmacokinetics will follow standard imputation methods, as described in a separate Clinical Pharmacology Analysis Plan (CPAP). Otherwise, missing data will not be imputed.

6. STUDY PATIENT DATA

6.1. Analysis Populations/Sets

Full Analysis Set (FAS): Defined as all patients who receive at least one dose of study treatment and have at least one evaluable post-baseline efficacy evaluation (BILAG, PGA, SLEDAI-2K, CLASI, ACR-28, SLICC/ACR, CCI , or CCI). FAS analyses will be conducted on the basis of the randomized treatment. The FAS will be used as the basis for the primary efficacy analysis.

Safety Analysis Set (SS): Defined as all patients who receive at least one dose of study treatment. Safety analyses will be conducted on the basis of actual treatment received. The SS will be the basis of all safety reporting.

Per Protocol Set (PP): Defined to exclude major protocol violations from the efficacy analysis. The PP will include all patients from the FAS except those with major violations to the protocol deemed to impact the analysis of the primary endpoint. These violations will be identified based on blinded data prior to study unblinding. PP analyses will be conducted on the basis of the randomized treatment. Additional information on inclusion into this population can be found in Section 6.3.

Pharmacokinetic Analysis Set (PK): Defined as the FAS patients who received at least one dose of BOS161721 with sufficient concentration data for the calculation of at least one PK parameter.

6.2. Patient Disposition

Disposition data will be summarized for all randomized patients. The number of patients in each analysis population, the number of patients who discontinued study treatment including reasons, and the number of patients who discontinued the study including reasons will be summarized. The number of patients who completed the study will also be summarized. Data will be presented by study phase and treatment group.

A by-patient listing of patient disposition data including reason for discontinuation, if applicable, will be presented for all randomized patients. A by-patient listing of randomization details will also be provided for all randomized patients.

A by-patient listing of inclusion/exclusion criteria and screen failures will be produced for all enrolled patients. Screen failures will otherwise not be included in any analyses.

6.3. Protocol Deviations

Protocol deviations will be identified and classified as major violations before database lock and unblinding. Major protocol deviations may include but are not limited to:

- Violation of Inclusion/Exclusion Criteria
- Use of prohibited therapies
- Incorrect treatment

Protocol deviations will be summarized by protocol version, deviation category, and major/minor designation. Protocol deviations will be summarized by study phase and treatment group for all randomized patients.

A listing of protocol deviations will be provided for all randomized patients.

6.4. Demographic and Baseline Characteristics

Patient demographics and baseline characteristics will be summarized in the SS. These will include age (years), gender (Male / Female), ethnicity (Hispanic or Latino / Not Hispanic or Latino), race (American Indian or Alaska Native / Asian / Black or African American / Native Hawaiian or Pacific Islander / White / Other), baseline height (cm), baseline weight (kg), and baseline BMI (kg/m^2). Age is reported as collected in the clinical database and will also be categorized as a categorical variable ($\leq$45, >45-<65, $\geq$65) for reporting. Patient eligibility for central review will be summarized. Childbearing potential and birth control methods for female and male patients will be summarized for the DMC meetings and interim analysis only. Demographics and baseline characteristics will be summarized by study phase and treatment group.

Demographics and baseline characteristics will be listed for all randomized patients. A separate by-patient listing will be created for childbearing potential and birth control for all randomized patients.

The following conversions and equations will be used as applicable:

Height (in cm) = height (in inches) * 2.54

Weight (in kg) = weight (in lbs) * 0.4536

BMI (kg/m^2) = weight(kg)/[height$(m)^2$]

6.5. Medical History

Medical history will be presented in data listings for all randomized patients. Medical history coded terms will be provided, including the system organ class and preferred term. The data management plan specifies the version of MedDRA used for medical history coding.

6.6. Prior and Concomitant Medication and Procedures

The incidence of medication use will be summarized by WHO Drug Dictionary anatomic therapeutic class (ATC) Level 2 classification (i.e. therapeutic main group) and preferred name. A patient will be counted only once at each level of reporting. Prior medications are those which have been identified to have been discontinued prior to the first dose of study medication. Concomitant medications are those which have been identified to have been taken at any point during the period of time between the first and last dose of study medication, including medications which started prior to first dose of study medication but are ongoing at first dose. Subsequent medications are those which have been started after the last dose date. The data management plan specifies the version of WHO Drug used.

Partial dates will be imputed according to Section 5.6 for the determination of prior, concomitant, and subsequent medications.

Concomitant medication use will be summarized and presented by study phase and treatment group in the SS.

All prior and concomitant medication data will be provided together in a by-patient listing including the verbatim and preferred drug name and WHO ATC Class (ATC Level 2) for all randomized patients. Past biological, immunosuppressant, antimalarial and corticosteroid SLE medication history data will be listed separately for all randomized patients.

Concomitant procedures will be presented in a data listing for all randomized patients.

6.7. Study Drug Exposure and Compliance

The total dose administered (mL) and the number of doses received will be summarized as continuous variables. The number and percentage of patients dosed and the number of patients who received the total planned dose at each nominal, protocol specified, visit (Day 0 / Day 30 / Day 60 / Day 90 / Day 120 / Day 150 / Day 180) will be summarized. The denominator for the number of patients dosed will be the number of patients in the SS. The denominator for the number of patients who received the total planned dose will be the number of patients dosed at the corresponding nominal visit.

The percent compliance will be calculated descriptively and is defined as 100 * (total actual dose summed across all visits)/total planned dose summed across all visits).

The duration of exposure in days will also be summarized descriptively and is defined as the last dose date – first dose date +1.

Study drug exposure and compliance will be presented by study phase and treatment group in the SS.

By-patient listings of BOS161721 and placebo dosing data will be produced for all patients in the SS.

7. EFFICACY

7.1. Primary Efficacy Endpoint and Analyses

Safety is the primary endpoint for the Phase 1b MAD portion of the study and is discussed in Section 9. Efficacy will be evaluated for the Phase 1b MAD portion of the study as an exploratory endpoint and details will be included in relation to the Phase 2 POC portion of the study within this section.

7.1.1. The proportion of patients with a SRI-4 response at Day 210

The primary efficacy endpoint is the proportion of patients who achieve a SRI-4 response at Day 210.

The SRI-4 is a composite index of SLE disease improvement that consists of scores derived from the SLEDAI-2K and the BILAG 2004 Index. Response based on the SRI-4 is defined by:

1) ≥ 4-point reduction from baseline in SLEDAI-2K global score

AND

2) No new severe disease activity (BILAG A organ score) or more than 1 new moderate organ score (BILAG B) compared with baseline

AND

3) No deterioration from baseline in the PGA by ≥ 30mm.

The proportion of patients who achieve a SRI-4 response at Day 210 will be presented and assessed using Pearson's chi-square test. The number of SRI-4 responders, treatment difference versus placebo, odds ratio, and 90% confidence interval will be displayed. The 90% confidence interval will be calculated with Wilson (score) confidence limits for the binomial proportion by treatment group. If any of the expected cell counts are less than 5, Fisher's exact test will be used and exact confidence intervals will be constructed. The analyses will be repeated for each of the three components contributing to the SRI-4 response. The primary analysis of the primary efficacy endpoint will be performed in the FAS and presented by treatment group. The analysis will be repeated in the PP analysis set.

The superiority of BOS161721 relative to placebo will be evaluated. If the proportion of SRI-4 responders for the selected POC dose in BOS161721 arm is higher than that of the placebo arm, then BOS161721 will be considered superior to placebo if the 2-sided p-value is less than 0.10.

Evaluability/Non-Responders:

Patients who received prohibited medications or unallowable corticosteroid (CS) usage will be considered "medication failures" and will be treated as non-responders for the primary efficacy analysis. Prohibited medications are detailed in Appendix 4 of the study protocol. Prohibited medications and unallowable CS usage will be recorded as protocol deviations in the clinical database.

A patient is considered to have unallowable CS usage if any of the following is not met:

- After Day 0 (after initiation of study therapy), no up-titration above 10 mg/day is allowed except for up to 1 CS burst for increased disease activity.

- Tapering is allowed after randomization except within 60 days of the primary (Day 210) and secondary (Day 120) endpoint assessments. Between Day 60 and Day 120, and between Day 150 and Day 210, oral CS doses must be held constant.

- A maximum of 1 oral CS "burst" for increased SLE disease activity will be allowed during the study between Day 0 and Day 60, according to the following:

    - An oral CS "burst" between Day 0 and Day 60; (an increase of ≤ 40 mg/day of prednisone or equivalent), which must be tapered down to a maximum of 10 mg/day within 2 weeks of initiation of the "burst"

- Alternatively, a single intramuscular (IM) dose of methylprednisolone (40 mg or equivalent) is permitted

- The course of the oral CS "burst" is not permitted to extend beyond Day 60

Patients who received an allowable CS burst but having missing data at Day 210 will be considered a medication failure and will be treated as non-responders for the primary efficacy analysis. The determination of medication failures will be reviewed using blinded data and finalized prior to unblinding. Once a patient is determined to be a non-responder, they will not be considered for response at a later visit.

If a patient has a BILAG, SLEDAI-2K, or PGA assessment outside of a 30 day +/-4 day window since the previous assessment (0-25 or >=35 days since the prior assessment), the corresponding result will be considered not evaluable and will be excluded from all analyses. The excluded result will not contribute to LOCF analyses. In these cases, the individual assessment (i.e. BILAG, SLEDAI-2K, or PGA) will be considered as if it was missed and analyses will be handled per the below details.

Missing Data:

- If a patient is missing data during the pre-treatment period for SLEDAI-2K, PGA, or BILAG, the patient will be excluded from the primary efficacy analysis.

- Missing data at Day 210 for overall components for SLEDAI-2K, BILAG, and PGA will be addressed by employing a last observation carried forward (LOCF) analysis within each overall component as described in Section 5.5. For example, if a patient has data for SLEDAI-2K and BILAG at Day 210 but is missing PGA at Day 210, the PGA component would employ LOCF. Data collection and scoring will be performed in the clinical database for BILAG, SLEDAI-2K, and PGA. As such, missing individual items contributing to any of the overall components will not be imputed/carried forward.

- Missing data for visits other than Day 210 will be handled in the same way as Day 210.

- If one or more consecutive visits are missed, LOCF analyses will be carried forward from the most recent prior assessment. For example, if the Day 150, Day 180 and Day 210 assessments are all missed but the Day 120 assessment was performed, the Day 120 assessment would be carried forward to Days 150, 180, and 210 for analysis.

Several sensitivity analyses will be performed:

Analysis	**Criteria**
Sensitivity Analysis 1	A sensitivity analysis of the primary efficacy endpoint will be performed (including LOCF methods) for patients in the full analysis set excluding patients with:

	(1) Positive anti-drug antibodies at baseline through Day 210. AND (2) Confirmed nAb at Day 90 and Day 180 OR missing Day 90 and missing Day 180 nAb OR missing Day 90 nAb and confirmed nAb at Day 180. Missing ADA and missing nAb results will be handled with a LOCF approach as discussed in Section 5.5. Additional analyses exploring impact of baseline factors such as baseline disease severity, race, and other characteristics may be performed.
Sensitivity Analysis 2	A sensitivity analysis of the primary efficacy endpoint will be performed (including LOCF methods) for patients in the full analysis set; however, if a patient has a BILAG, SLEDAI-2K, or PGA assessment outside of a 30 day +/-4 day window since the previous assessment (0-25 or >=35 days since the prior assessment), the corresponding result will be considered evaluable and will be included in the analysis and will contribute to LOCF.
Sensitivity Analysis 3	A sensitivity analysis of the primary efficacy endpoint will be performed (excluding LOCF methods) to consider the effect of missing data by excluding patients with either or both of the following criteria: (1) Missing assessments at Day 210 (patients with a missed assessment of any of BILAG, SLEDAI-2K, or PGA at Day 210 would be excluded). AND/OR (2) Patients with any Day 210 assessment which occurred outside of the 30 day +/- 4 day window (patients with an out of window assessment of any of BILAG, SLEDAI-2K, or PGA at Day 210 would be excluded).

Additional sensitivity analyses may be performed to consider the effect of missing data.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

A summary of disposition of SRI-4 response at day 210 will be presented by study phase and treatment group in the FAS and PP. Sensitivity analysis 1 will also be performed for disposition of SRI-4 response at day 210. These summaries will include the number of SRI-4 responders and the number of SRI-4 non-responders with a breakdown of reasons for non-response (early withdrawal, medication failure, <4-point reduction from baseline in SLEDAI-2K global score, new BILAG 1A/2B, and PGA worsening). Denominators for percentages will be the number of patients eligible for SRI-4 response analysis.

A listing of SRI Response data at Day 150, Day 180, and Day 210 will be presented in the FAS. A listing of patients excluded from the efficacy analysis will also be provided for all randomized patients.

7.2. Secondary Efficacy Endpoints and Analyses

Pharmacokinetics and pharmacodynamics are the secondary endpoints for the Phase 1b MAD part of the study and are described in Section 8. Efficacy will be evaluated for the Phase 1b MAD part of the study as an exploratory endpoint and details will be included in relation to the Phase 2 POC part of the study within this section.

7.2.1. Proportion of Patients with SRI-4 Response at Each Visit

The analyses described in Section 7.1.1, including sensitivity analyses allowing out of window assessment results (Sensitivity Analysis 2) and excluding patients with missing or out of window assessments (Sensitivity Analysis 3), will be repeated for each nominal visit and will be presented by study phase and treatment group in the FAS. The 3 components will not be summarized separately by visit.

A graphical display of SRI-4 response by study phase, treatment group, and visit in the FAS will be presented.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.2. Proportion of Patients with SRI-5 Response at Each Visit

The SRI-5 is a composite index of SLE disease improvement that consists of scores derived from the SLEDAI-2K and the BILAG 2004 Index. Response based on the SRI-5 is defined by:

1) ≥ 5-point reduction from baseline in SLEDAI-2K global score

AND

2) No new severe disease activity (BILAG A organ score) or more than 1 new moderate organ score (BILAG B) compared with baseline

AND

3) No deterioration from baseline in the PGA by ≥ 30mm.

The analyses described in Section 7.1.1, including sensitivity analyses excluding patients with missing or out of window assessments (Sensitivity Analysis 3), will be repeated for each nominal visit based on the SRI-5 response definition and will be presented by study phase and treatment group in the FAS. The 3 components will not be summarized separately by visit for SRI-5 response.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.3. Proportion of Patients with SRI-6 Response at Each Visit

The SRI-6 is a composite index of SLE disease improvement that consists of scores derived from the SLEDAI-2K and the BILAG 2004 Index. Response based on the SRI-6 is defined by:

1) ≥ 6-point reduction from baseline in SLEDAI-2K global score

AND

2) No new severe disease activity (BILAG A organ score) or more than 1 new moderate organ score (BILAG B) compared with baseline

AND

3) No deterioration from baseline in the PGA by ≥ 30mm.

The analyses described in Section 7.1.1, including sensitivity analyses excluding patients with missing or out of window assessments (Sensitivity Analysis 3), will be repeated for each nominal visit based on the SRI-6 response definition and will be presented by study phase and treatment group in the FAS. The 3 components will not be summarized separately by visit for SRI-6 response.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.4. Proportion of Patients with a sustained reduction of oral corticosteroid (CS) between Day 120 and Day 210

The proportion of patients with a sustained reduction of oral corticosteroid, defined as ≤ 10 mg/day and ≤ Day 0 dose, between Day 120 and Day 210 will be summarized by study phase and treatment group in the FAS. LOCF and imputation for missing values will not be performed. Otherwise, statistical methods described in Section 7.1.1 will be implemented.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.5. Proportion of new BILAG A flare or >1 BILAG B flare relative to baseline through Day 210

The proportion of new BILAG A flare or >1 BILAG B flare relative to baseline through Day 210 will be summarized by study phase and treatment group in the FAS. Statistical methods described in Section 7.1.1 will be implemented, including LOCF methods within a body system. The analyses will also be repeated in observed data only (no LOCF). A BILAG A flare is defined as a new A score in at least one body system when compared to baseline. A BILAG B flare is defined as a new B score in at least two body systems at the same assessment when compared to baseline. BILAG grading is collected at each protocol specified visit as an assessment of "the past 4 weeks". As such, the BILAG grade will be compared to baseline and the most recent prior assessment, as appropriate, to determine a new flare. The proportion of new BILAG A flare or >1 BILAG B flare relative to baseline will be summarized overall and by nominal visit.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

Observed results will be presented in separate data listings for BILAG SLE history, BILAG-2004 index, and BILAG grading for all randomized patients.

7.2.6. Proportion of patients with PGA worsening

The PGA is used to assess investigator's general impression on the patient's overall status of SLE disease activity via visual analogue scale (100 mm) with 0 being "very good, asymptomatic and no limitation of normal activities" with 100 mm being "most severe possible disease ever seen in all SLE patients." The proportion of patients with PGA worsening, defined as an increase of >30mm from baseline, will be summarized by study phase and treatment group in the FAS overall and by nominal visit. Statistical methods described in Section 7.1.1 will be implemented, including LOCF methods. The analyses will also be repeated in observed data only (no LOCF).

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

PGA observed results will be presented in a data listing for all randomized patients.

7.2.7. Proportion of patients with BICLA response

The BICLA is a responder index developed to measure response to therapy, and it includes scores from the BILAG, SLEDAI-2K, and PGA. BICLA response is defined as:

1) at least 1 gradation of improvement in baseline BILAG 2004 scores in all body systems with moderate disease activity at entry (e.g., all B [moderate disease] scores falling to C [mild], or D [no activity])

AND

2) no new BILAG A or more than 1 new BILAG B scores

AND

3) no worsening of total SLEDAI-2K score from baseline

AND

4) ≤ 10% deterioration in PGA score

AND

5) no medication failure.

LOCF and imputation for missing values will not be performed. Otherwise, statistical methods described in Section 7.1.1 will be implemented. The analyses will be repeated overall and for each nominal visit based on the BICLA response definition and will be presented by study phase and treatment group in the FAS.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.8. Cutaneous Lupus Erythematosus Area and Severity Index (CLASI)

The CLASI is a comprehensive tool for assessment of disease activity and damage in cutaneous lupus, shown to be valid, reliable, and sensitive to changes in disease activity.

The total activity score will be calculated by summing the scores on the left side of the form (erythema, scale/hypertrophy, mucous membrane, alopecia – recent hair loss and clinically not obviously scarred).

The total damage score will be calculated by:

(1) Sum the responses for “Dyspigmentation” in all anatomical locations

(2) If a response for “Report duration of dyspigmentation after active lesions have resolved” is “Dyspigmentation usually lasts at least 12 months”, then multiply the sum in #1 by 2.

(3) If a response for “Report duration of dyspigmentation after active lesions have resolved” is “Dyspigmentation usually lasts less than 12 months”, then the sum in #1 will not change.

(4) If the sum in #1 is 0 and a response for “Report duration of dyspigmentation after active lesions have resolved” is missing, then the sum in #1 will not change.

(5) Determine the Dyspigmentation Score based on the results from 2, 3, or 4.

(6) Sum the remaining scores on the right side of the form (scarring/atrophy/panniculitis, alopecia - scarring of the scalp judged clinically).

(7) Sum the scores from #5 and #6 to obtain the total damage score.

If the total activity score or total damage score cannot be calculated, the total activity score or total damage score will be imputed using the LOCF method described in Section 5.5. If one of the total scores can be calculated and the other cannot, LOCF will be implemented for only the total score which cannot be calculated. If any individual scores which contribute to the total activity score or the total damage score, the total score in question will not be calculated and LOCF methods for total score will be implemented. LOCF methods will be implemented for CLASI response analyses only.

A sensitivity analysis for CLASI response will also be performed in the FAS excluding patients with missing CLASI results within a visit. If one or both of the total activity score or total damage score are missing at a visit, the patient will not be included in the response analysis at that visit. LOCF methods will not be implemented for the sensitivity analysis.

The mean and standard deviation of the change from baseline in CLASI total activity and total damage scores will be presented graphically over time by study phase and treatment group.

7.2.8.1. Proportion of patients with CLASI response

CLASI response is defined as 50% improvement from baseline in "A" (total activity) or "B" (total damage) scores. This assessment will be applied to all patients as all are required to have cutaneous disease activity.

The analyses described in Section 7.1.1 will be repeated overall and for each nominal visit based on the CLASI response definition and will be presented by study phase and treatment group in the FAS.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.8.2. Results and Changes from Baseline in CLASI

The analysis of CLASI total activity score and CLASI total damage score will be performed separately using an analysis of covariance (ANCOVA) model with change from baseline to each nominal visit as the dependent variable, treatment group as the factor, and the corresponding baseline CLASI score as a covariate. The p-value for the difference between the BOS161721 and placebo groups at each nominal visit is based on the F-test.

Descriptive statistics for the activity score and damage score will be summarized separately at baseline and each visit. Change from baseline will be summarized at each post-baseline visit. The LS Mean estimate, its standard error and 90% CI will be tabulated by treatment group and for the difference between treatment groups (BOS161721 versus placebo) for change at each visit. Missing CLASI results will not be imputed. Observed results only will be presented for the change from baseline analysis.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

CLASI observed results will be presented in data listings for all randomized patients.

7.2.9. Results and Changes from Baseline in Swollen and Tender Joints ACR-28

The ACR-28 joint count evaluates the number of tender and swollen joints in the shoulder, elbow, wrist, hand, knee joints. Joints of the feet are excluded.

The analysis of Swollen and Tender Joints ACR-28 for the sum of tenderness (left and right) and the sum of swelling (left and right) will be performed separately using an analysis of covariance (ANCOVA) model with change from baseline to each nominal visit as the dependent variable, treatment group as the factor, and the baseline ACR-28 score as a covariate. The derived sum for tenderness and swelling is calculated and provided in the clinical database. A missing total sum of swelling or sum of tenderness will not be imputed. Observed results only will be presented for the change from baseline analysis. The p-value for the difference between the BOS161721 and placebo groups at each nominal visit is based on the F-test.

Descriptive statistics for the tenderness and swelling derived sums will be summarized separately at baseline and each visit. Change from baseline will be summarized at each post-baseline visit. The LS Mean estimate, its standard error and 90% CI will be tabulated by treatment group and for the difference between treatment groups (BOS161721 versus placebo) for change at each visit.

The mean and standard deviation of the change from baseline in ACR-28 sum of tenderness and sum of swelling will be presented graphically over time by study phase and treatment group.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

Swollen and Tender Joints ACR-28 observed results will be presented in data listings for all randomized patients.

7.2.10. Results and Changes from Baseline in SLEDAI-2K

The SLEDAI-2K is a validated instrument that measures disease activity in SLE patients at the time of the visit and in the previous 30 days. It is a global index and includes 24 clinical and laboratory variables that are weighted by the type of manifestation, but not by severity. The total score falls between 0 and 105, with higher scores representing increased disease activity. The SLEDAI-2K has been shown to be a valid and reliable disease activity measure in multiple patient groups. A SLEDAI-2K of 6 or more generally represents moderately to severely active disease.

The analysis of SLEDAI-2K will be performed using an analysis of covariance (ANCOVA) model with change from baseline to each nominal visit as the dependent variable, treatment group as the factor, and the baseline SLEDAI-2K score as a covariate. The SLEDAI-2K total score is calculated and provided in the clinical database. A missing total SLEDAI-2K score will not be imputed. Observed results only will be presented for the change from baseline analysis.

The p-value for the difference between the BOS161721 and placebo groups at each nominal visit is based on the F-test.

Descriptive statistics for the SLEDAI-2K total score will be summarized separately at baseline and each visit. Change from baseline will be summarized at each post-baseline visit. The LS Mean estimate, its standard error and 90% CI will be tabulated by treatment group and for the difference between treatment groups (BOS161721 versus placebo) for change at each visit.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

The mean and standard deviation of the change from baseline in observed SLEDAI-2K total score will be presented graphically over time by study phase and treatment group.

SLEDAI-2K observed results will be presented in a data listing for all randomized patients.

7.2.11. Results and Changes from Baseline in SLICC/ACR Damage Index

The SLICC/ACR damage index is a validated instrument to assess damage, defined as irreversible impairment, continuously persistent for 6 months (ascertained by clinical assessment), occurring since the onset of lupus, and it is based on a weighted scoring system. This index records damage occurring in patients with SLE regardless of cause, with demonstrated content, face, criterion, and discriminant validity. It will be performed on Days 0 and 180.

The analysis of SLICC/ACR Damage Index total score will be performed using an analysis of covariance (ANCOVA) model with change from baseline to each nominal visit as the dependent variable, treatment group as the factor, and the baseline SLICC/ACR Damage Index total score as a covariate. The SLICC/ACR Damage Index total score is calculated and provided in the clinical database. A missing total SLICC/ACR Damage Index score will not be imputed. Observed results only will be presented for the change from baseline analysis. The p-value for the difference between the BOS161721 and placebo groups at each nominal visit is based on the F-test.

Descriptive statistics for the SLICC/ACR Damage Index total score will be summarized separately at baseline and each visit. Change from baseline will be summarized at each post-baseline visit. The LS Mean estimate, its standard error and 90% CI will be tabulated by treatment group and for the difference between treatment groups (BOS161721 versus placebo) for change at each visit.

These analyses will also be performed for the MAD Phase 1b with the exception of the statistical testing.

Observed SLICC/ACR Damage Index data will be presented in a by-patient listing for all randomized patients. SLICC Criteria for SLE at screening will also be presented in a by-patient listing for all randomized patients.

7.2.12. Medication Failure

7.2.12.1. Proportion of patients with medication failures

Patients who received prohibited medications or unallowable CS usage as described in Section 4.6 of the study protocol will be considered "medication failures" and will be treated as non-responders for the primary efficacy analysis and selected sensitivity and secondary efficacy analyses. Patients who received an allowable CS burst but having missing data at Day 210 will considered a medication failure. The determination of medication failures will be reviewed using blinded data and finalized prior to unblinding.

The proportion of patients with medication failures will be summarized by study phase and treatment group in the FAS overall and by visit. A patient will be counted as a medication failure at each visit occurring on or after the date in which a patient is considered a medication failure. Statistical methods described in Section 7.1.1 will be implemented for observed data only. LOCF and missing data imputation will not be performed.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.12.2. Time to Medication Failure (TTMF)

Medication failure is defined in Section 4.6 of the study protocol and discussed in Section 7.1.1. Time to medication failure will be computed as the event date (as described below) – randomization date + 1.

The date of medication failure will be reviewed by Boston Pharmaceuticals and agreed upon prior to database lock and unblinding.

Kaplan-Meier methods will be used to estimate TTMF for each treatment group. Estimates of median TTMF will be provided along with 90% confidence intervals. Q1 and Q3 will also be tabulated. Comparisons between treatment groups will use the one-sided log-rank test. In addition, hazard ratio (HR) estimates with 90% CIs will be calculated from a Cox proportional hazards regression analysis that includes treatment as a factor. Censoring rules are defined as follows.

Situation	Date of Event or Censoring	Outcome
Patient is designated a medication failure (received prohibited meds or unallowable CS burst)	Date of first dose of medication leading to medication failure designation. These dates will be obtained from the date of deviation on the protocol deviation CRF form.	Event
Patient withdraws from study (no medication failure)	Censored at date of withdrawal	Censored
Patient completes the study (no medication failure)	Censored at last study visit	Censored
Patient dies during study (no medication failure)	Censored at the date of death	Censored

Kaplan-Meier curves of TTMF will be plotted over time.

These analyses will also be performed for the MAD Phase 1b with the exception of the statistical testing.

Time to medication failure data will be provided in a by-patient listing in the full analysis set.

7.2.13. Duration of Longest SRI-4 Response

Duration of longest SRI-4 response will be computed for patients who have been identified as a responder at least once based on criteria defined in Section 7.1.1. The duration of longest response will be defined as the longest period a patient meets the SRI-4 responder criteria at consecutive visits. This will be computed as the date of last consecutive SRI-4 response – date of first consecutive response + 1.

The duration of longest SRI-4 response will be analyzed using an analysis of variance (ANOVA) model with treatment group as the effect. The p-value for the difference between the BOS161721 and placebo group is based on the F-test.

The mean, standard deviation, median, minimum, maximum, Q3, Q4, LS Mean estimate, its standard error and 90% CI will be tabulated by treatment group and for the difference between treatment groups (BOS161721 versus placebo).

If a patient has missed a visit between two scheduled visits with a response at both, the patient will be counted as having a response at the missed visit.

Duration of longest SRI-4 response will be displayed in the SRI response listing.

These analyses will also be performed for the MAD Phase 1b with the exception of the statistical testing.

7.2.14. Time to First SRI-4 Response

SRI-4 response is defined in Section 7.1.1. Time to first SRI-4 response will be computed as the event date (as described below) – randomization date + 1.

Time to first SRI-4 response will be analyzed with Kaplan-Meier methods described for TTMF and censoring as described below. Kaplan-Meier curves will be plotted over time.

Situation	**Date of Event or Censoring**	**Outcome**
Patient has SRI-4 Response (prior to medication failure, withdrawal, or death)	First date where a patient has been documented to have achieved a SRI-4 responder designation	Event
Patient is designated a medication failure (received prohibited meds or unallowable CS burst) prior to SRI-4 response/non-responder	Date of first dose of medication leading to medication failure designation. These dates will be obtained from the date of deviation on the protocol deviation CRF form.	Censored
Patient withdraws from study (no SRI-4 response, no medication failure)	Censored at date of last visit with non-missing SRI-4 assessments (BILAG, SLEDAI-2K, and PGA) prior to withdrawal	Censored
Patient completes the study (no SRI-4 Response, no medication failure)	Censored at date of last visit with non-missing SRI-4 assessments (BILAG, SLEDAI-2K, and PGA)	Censored
Patient dies during study (no SRI-4 response, no medication failure)	Censored at the date of death	Censored

These analyses will also be performed for the MAD Phase 1b with the exception of the statistical testing.

Time to first SRI-4 response data will be provided in a by-patient listing in the full analysis set.

7.2.15. Time to First BILAG A Flare or >1 BILAG B Flare Compared to Baseline through Day 210

Time to first BILAG A Flare or >1 BILAG B Flare will be will be computed as the event date (as described below) – randomization date + 1. If a patient is designated as a medication failure prior to or on their day 210 visit, they will be considered as having a BILAG flare (event) as of the date of first dose of the medication leading to the medication failure designation due to prohibited medication or unallowable CS burst.

Time to first BILAG A Flare or >1 BILAG B Flare will be analyzed with Kaplan-Meier methods described for TTMF and censoring as described below. Kaplan-Meier curves will be plotted over time.

Situation	**Date of Event or Censoring**	**Outcome**
Patient has a BILAG A Flare or >1 BILAG B Flare compared to baseline (Noted as BILAG 1A/2B Flare) prior to or on their day 210 visit	Date of first BILAG 1A/2B Flare	Event
Patient is designated a medication failure (received prohibited meds or unallowable CS burst) prior to or on their day 210 visit	Date of first dose of medication leading to medication failure designation. These dates will be obtained from the date of deviation on the protocol deviation CRF form.	Censored
Patient withdraws from study (no BILAG 1A/2B Flare, no medication failure) prior to or on their day 210 visit	Censored at date of last non-missing BILAG assessment (prior to withdrawal) through day 210	Censored
Patient completes the study (no BILAG 1A/2B Flare, no medication failure)	Censored at date of last non-missing BILAG assessment through day 210	Censored
Patient dies during study (no BILAG 1A/2B Flare, no medication failure) prior to or on their day 210 visit	Censored at the day of death	Censored

These analyses will also be performed for the MAD Phase 1b with the exception of the statistical testing.

Time to first BILAG A Flare or >1 BILAG B Flare data will be provided in a by-patient listing in the full analysis set.

8. PHARMACOKINETICS/PHARMACODYNAMICS

Pharmacokinetics and pharmacodynamics are secondary endpoints for the Phase 1b MAD study and exploratory endpoints for the Phase 2 POC study.

PK concentrations and PD results reported as less than the lower limit of quantification (LLOQ) will be counted as 0 to calculate summary linear statistics, 0.5*LLOQ for geometric statistics, and will be reported as <LLOQ in listings.

PK blood sampling as collected on the CRF will be provided in by-patient data listings. All observed PK concentrations and PD data will also be presented in by-patient listings, including CCI [redacted], nAb, and ADA. PK parameters, as provided by Human Predictions, LLC, will be displayed in by-patient data listings.

Additional details surrounding the PK and PD analysis, modeling, and evidence of PK/PD relationships will be provided in a separate CPAP.

8.1. Pharmacokinetics

BOS161721 PK concentrations will be summarized and listed in the PK set by study phase, treatment group, visit, and timepoint. Summary statistics will include the geometric mean, geometric standard deviation, geometric % coefficient of variation, n, mean, standard deviation, median, minimum, and maximum. BOS161721 concentrations will be plotted over time by treatment group and visit.

PK parameters for the PK analysis set will be calculated by Human Predictions, LLC and will be provided to Array Biostatistics to prepare outputs for the MAD Phase 1b. Details surrounding the PK parameter calculations will be provided in a separate CPAP. Summary statistics will be presented for C_{max}, T_{max}, AUC, $t_{½}$, CL, and V_d and will include the same statistics as PK concentrations except Tmax which will include n, median, minimum, and maximum, only.

8.2. Pharmacodynamics

Descriptive statistics for the observed results of the following PD endpoints/results will be summarized separately at baseline and each visit. Change from baseline will be summarized at each post-baseline visit. In cases where both numeric and qualitative results are provided, the summaries will be based on the numeric results. Graphical presentations of mean and standard deviation of the change from baseline over time will also be presented for the below parameters.

PD Endpoint	**Applicable Study Portion(s)**	**Parameters to be Included in Summary Tables and Figures ('NA (all)' is displayed if all collected parameters are to be included)**
Phosphorylated signal transducer and activator of transcription 3 (pSTAT3)	MAD only	% pSTAT3+ Lymphocytes –Stimulated; Ratio MEF (Stimulated / Unstimulated)
CCI [redacted]	MAD and POC	NA (all)
Leukocyte immunophenotype (B and NK Cells)	MAD and POC	CD19+% of CD45+ Lymphocytes; IgD+C27-% of CD19+; IgD+CD27-CD38++CD24++% of CD19+; IgD-CD27-% of CD19+; IgD+CD27+% of CD19+; IgD-CD27+% of CD19+; IgD-CD27+CD38++CD138-% of CD19+; IgD-CD27+CD38++CD138+% of CD19+; CD56+% OF CD45+ Lymphocytes
Leukocyte imunophenotype (T-Cells)	MAD and POC	CD4+CD8-% of CD3+; CXCR5+% of CD4+CD8-; PD-1+% of CD4+CD8-; ICOS+% of CD4+CD8-; CD25+CD127-% of CD4+CD8-; CXCR5+% of CD25+CD127-; PD-1+% of CD25+CD127-; ICOS+% of CD25+CD127-; CD8+CD4-% of CD3+
CCI [redacted]	MAD and POC	CCI [redacted]
CCI [redacted]	MAD and POC	NA (all)

PD Endpoint	Applicable Study Portion(s)	Parameters to be Included in Summary Tables and Figures ('NA (all)' is displayed if all collected parameters are to be included)
CCI	MAD and POC	NA (all)
CCI	MAD and POC	NA (all)
CCI	MAD and POC	NA (all)
CCI	MAD and POC	All parameters for 'average target CP' values

CCI

Additional immunogenicity and PD results may be plotted over time by treatment group and visit.

9. QUALITY OF LIFE

9.1. CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

9.2. CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

10. SAFETY

All safety analysis reporting will be based on the Safety Analysis Set.

10.1. Adverse Events

Adverse events (AEs) will be collected and recorded for each patient from the date of the first dose of study drug until the end of their participation in the study, including the safety follow up period. Severity of adverse events will be graded according to National Cancer Institute

Common Terminology Criteria for Adverse Events (NCI CTCAE) version 4.03. AEs will also be assessed for relationship to study drug, and seriousness. Adverse events will be marked as adverse events of special interest (AESIs), dose-limiting toxicities (DLTs), or injection site reactions (ISRs), where applicable. AEs will be considered treatment-emergent adverse events (TEAEs) if they started any time after first dose or if the start date was missing. Any missing severity assessments will be assumed to be Grade 3, missing relationship assessments will be assumed to be related, and missing seriousness assessments will be assumed as serious.

An overview of TEAEs will be produced, including counts and percentages of patients with any incidences of TEAE, TEAEs related to study treatment, Serious adverse events (SAEs), SAEs related to study treatment, CTCAE grade 2 or higher TEAEs, CTCAE grade 2 or higher TEAEs related to study treatment, CTCAE grade 3 TEAEs, CTCAE grade 3 or higher TEAEs related to study treatment, CTCAE grade 4 TEAEs, CTCAE grade 4 or higher TEAEs related to study treatment, TEAEs leading to study treatment discontinuation, Related TEAEs leading to study treatment discontinuation, TEAEs of special interest, Dose-limiting toxicities, TEAEs resulting in death, and injection site reactions.

Adverse events will be coded based on the Medical Dictionary for Regulatory Affairs (MedDRA) for reporting by system organ class (SOC) and preferred term (PT) in descending order of overall incidence. The MedDRA version used is defined in the study data management plan.

Summaries of adverse events by SOC and PT will include the following types:

- TEAEs;
- TEAEs related to study treatment;
- SAEs;
- CTCAE Grade 3 or higher TEAEs related to study treatment;
- TEAEs leading to treatment discontinuation; and
- Dose-Limiting Toxicities
- TEAEs of Special Interest

A summary of TEAEs by SOC, PT, and maximum CTCAE grade will also be prepared. For these summaries, TEAEs will be sorted for each patient by SOC/PT and CTCAE grade; patients will be counted once within a SOC/PT based on their maximum CTCAE grade. Missing CTCAE grade will be counted in a missing row.

A summary of TEAEs by PT and maximum CTCAE grade will also be prepared with sections for any TEAE and related TEAEs. For these summaries, patients will be counted once within a PT based on their maximum CTCAE grade. Missing CTCAE grade will be counted as grade 3. Preferred terms will be included if at least one patient has a CTCAE grade 3 or higher for the PT. PTs will also be included if 10% or more of patients experience a CTCAE grade 1 or 2 TEAE. The summary will be sorted by the total incidence of the PT in the any TEAE section.

A comprehensive listing of all AEs will be provided in a by-patient data listing for the safety analysis set. In addition, the following listings will be provided for the safety analysis set:

- TEAEs related to study treatment;
- SAEs;
- TEAEs leading to treatment discontinuation; and
- Adverse Events of Special Interest

10.2. Clinical Laboratory Evaluations

Laboratory tests will be performed at times defined in the protocol Schedule of Assessments. Clinical chemistry and hematology parameters will be reported based on the International System of Units (SI). The following laboratory evaluations will be reported in data summaries, with asterisks indicating those that will be graded using CTCAE. Grading will be performed by the Central Lab based on the details provided in Appendix 13.2 for relevant laboratory parameters. If a grade is not present in the central lab database for a graded lab parameter, the grade will be analyzed as grade 0.

Hematology: Hemoglobin*, Hematocrit, RBC count, RDW, MCV, MCH, MCHC, Platelet count*, WBC count*, CD4+ count*, Total neutrophils (Abs)*, Eosinophils (Abs), Monocytes (Abs), Basophils (Abs), Lymphocytes (Abs)*

Clinical chemistry: Alanine Aminotransferase*, Albumin*, Alkaline Phosphatase*, Aspartate Aminotransferase*, Bicarbonate, Bilirubin*, Blood urea nitrogen, C Reactive Protein, Calcium, Chloride, Creatine Kinase*, Creatinine*, Glucose (fasting)*, Glucose (random)*, Potassium*, Protein, Sodium*, Uric acid*, eGFR (Cockcroft-Gault)*, eGFR (MDRD)*, Total cholesterol (fasting)*, LDL-C (fasting), HDL-C (fasting), Triglycerides (fasting)*, Gamma Glutamyl Transferase*.

Coagulation: Prothrombin Intl. Normalized Ratio*, Activated Partial Thromboplastin Time*

Urinalysis: pH, Glucose (qual), Protein (qual)*, Blood (qual)*, Ketones, Nitrites, Leukocyte esterase, Urobilinogen, Urine bilirubin, Microscopy

Observed values and changes from baseline for hematology, clinical chemistry, and coagulation laboratory evaluations will be summarized at each visit and most extreme change. Endpoint analyses will be based on changes from baseline assessments at Days 120 and 210.

The number and percent of patients with a NCI-CTCAE toxicity grade of 3 or higher will be tabulated by laboratory evaluation with defined NCI-CTCAE grading at each visit for hematology, clinical chemistry, and coagulation. Patients with missing values post-baseline will be excluded from the summary. The denominator for percentages will be the number of patients with at least one post-baseline assessment for the laboratory parameter in question.

Hematology, clinical chemistry, and coagulation shift tables displaying the shift from baseline to the worst value of NCI-CTCAE grade will be presented based on the most extreme change as it relates to the relevant NCI-CTCAE definition. NCI-CTCAE relating to “high/hyper” conditions will depend on the maximum post-baseline value while NCI-CTCAE “low/hypo” will be

reported based on the minimum post-baseline value. Separate shift tables will be prepared for parameters with bi-directional toxicity grading. Lab parameters which are optional, such as gamma glutamyl transferase, will not be included in the shift tables.

Hematology and clinical chemistry laboratory data which is not graded will also be summarized in shift tables of baseline to each visit and most extreme change based on range categories of low (below lower limit of normal [LLN], normal, and high [above upper limit of normal [ULN]]).

Selected hematology and clinical chemistry laboratory results will be plotted over time for treatment group for MAD phase 1b portion of the study in the safety analysis set. The mean and standard deviation of the change from baseline for selected hematology and clinical chemistry laboratory results will be plotted over time for the POC Phase 2 portion of the study in the safety analysis set.

All laboratory parameters as well as a summary of central laboratory tests and chest x-ray will be provided in patient data listings for all randomized patients. By patient listings of clinical chemistry and hematology data will be presented in the safety analysis set for the DMC and interim analysis only.

10.3. Other Safety Evaluations

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

The C-SSRS is a low-burden measure of the spectrum of suicidal ideation and behavior that was developed by Columbia University researchers for the National Institute of Mental Health Treatment of Adolescent Suicide Attempters Study to assess severity and track suicidal events through any treatment. It is a clinical interview providing a summary of both ideation and behavior that can be administered during any evaluation or risk assessment to identify the level and type of suicidality present. The C-SSRS can also be used during treatment to monitor for clinical worsening.[i] The C-SSRS evaluation will be performed as specified in the protocol Schedule of Assessments.

The C-SSRS has 11 binary (yes/no) outcomes corresponding to five categories of suicidal ideation, five categories of suicidal behavior, and the category of self-injurious behavior without suicidal intent:

Suicidal ideation is present if there is a “yes” response to any of the 5 suicidal ideation category questions. Suicidal behavior is present if there is a “yes” response to any of the 5 suicidal behavior category questions. Suicidal ideation or behavior is present if there is a “yes” response to any of the 10 suicidal or ideation or behavior category questions.

The number and percentage of patients with any post-baseline occurrence of the following ideations and behaviors will be presented by study phase and treatment group:

- Suicidal ideation (overall and by individual question),
- Suicidal behavior (overall and by individual question),

- Suicidal ideation or behavior,

A similar tabulation of the occurrence of any of the above ideations and behaviors at baseline or as part of their lifetime history/past 12-month history (collected at the screening visit) will be presented by study phase and treatment group.

A shift table displaying the shift from worst pre-treatment category to the worst post-baseline category will be presented by study phase and treatment group. Best to worst C-SSRS category is defined in the following order: no suicidal ideation or behavior, suicidal ideation, and suicidal behavior. The worst post-baseline shift is calculated based on all post-baseline visits, including unscheduled visits. Worst pre-treatment will include all C-SSRS assessments prior to first dose date/time (screening or baseline). The denominator will be the number of patients with at least one non-missing baseline and post-baseline C-SSRS assessment.

Missing data will not be imputed.

All C-SSRS individual items will be presented in data listings for screening and post-screening (since last visit), separately. A separate listing will be presented for patients with suicidal ideation, suicidal behavior, or non-suicidal self-injurious behavior for all randomized patients.

10.3.2. Vital Signs

Vital signs include: heart rate (beats/min); temperature (^{o}C); systolic and diastolic blood pressure (mmHg); and weight (kg). Observed values and changes from baseline for vital signs will be summarized at each visit and time point, as well as for most extreme change by study phase and treatment group.

All vital signs data will be presented in patient data listings for all randomized patients.

10.3.3. 12-Lead Electrocardiogram (ECG)

ECG will be assessed as specified in the protocol Schedule of Assessments. The following ECG parameters will be collected: PR interval (msec), QRS interval (msec), RR interval (msec), QT interval (msec), and QT interval corrected for heart rate using Fridericia's formula (QTcF) (msec). ECG parameters and ECG evaluation will come directly from the database and will not be calculated during analysis.

Observed values and changes from baseline for ECG parameters will be summarized at each visit, as well as for most extreme change by study phase and treatment group.

The number and percent of patients with the following parameters at any post-baseline visit and at each visit will be summarized:

- Abnormal, not clinically significant ECGs
- Abnormal, clinically significant ECGs

The number and percent of patients with each ECG evaluation result will be summarized by visit. The worst ECG evaluation at any post-baseline visit will also be summarized in the order of: Normal, Abnormal, not clinically significant, Abnormal, clinically significant.

For the summary of incidence at any post-baseline visit, the denominators for percentages will include patients with one or more non-missing value at any post-baseline visit. For the by-visit summaries, the denominators for percentages will include patients with a non-missing value at that visit.

Investigator reported ECG result shifts from baseline to each visit and worst case post-baseline will be summarized. Worst case post-baseline will be based on the most abnormal observed value on or after the randomization date.

All ECG data will be presented in a by patient data listing for all randomized patients. A separate by patient listing for patients with abnormal, clinically significant 12-lead ECGs will also be provided in the safety analysis set for DMC meetings and the interim analysis only.

10.3.4. Physical Examinations

Targeted physical examination data and full physical examination data will be presented in patient data listings. A separate by patient listing for patients with abnormal, clinically significant physical examination findings will also be provided in the safety analysis set for DMC meetings and the interim analysis only.

10.3.5. Anti-Drug Antibodies (ADAs)

The serum samples to measure the presence of ADA will be collected as specified in the protocol Schedule of Assessments. ADAs will be summarized and listed according to details provided in Section 8.

11. CHANGES TO THE PLANNED ANALYSIS

No changes have been made to the planned analysis.

12. REFERENCES

i Posner K, Melvin G a, Stanley B, Oquendo M a, Gould M. Factors in the assessment of suicidality in youth. *CNS Spectr*. 2007;12(2):156-162.

CCI [redacted]

International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH E3 Guideline: Structure and content of clinical study reports questions & answers (R1). 2012. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_QAs_R1_Step4.pdf. Accessed 13 Jan 2016.

International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonised Tripartite Guideline: Structure and content of clinical study reports E3. Step 4.
1995. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf. Accessed 13 Jan 2016.

International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH E9: Statistical principles for clinical trials.
1998. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf . Accessed 15 June 2016.

FDA Guidance for Industry Systemic Lupus Erythematosus – Developing Medical Products for Treatment.
2010. https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm072063.pdf. Accessed 23February2018.

Committee for Medicinal Products for Human use (CHMP). EMA Guideline on clinical investigation of medicinal products for the treatment of systemic lupus erythematosus and lupus nephritis.
2015. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/03/WC500184889.pdf. Accessed 04March2018.

CCI [redacted]

Nilsson M, Suryawanshi S, Gassmann-Mayer C, Dubrava S, McSorley P, Jiang K. Columbia-Suicide Severity Rating Scale Scoring and Data Analysis Guide. V2
2013. https://cssrs.columbia.edu/wp-content/uploads/ScoringandDataAnalysisGuide-for-Clinical-Trials-1.pdf. Accessed 18May2018.

CCI [redacted]

CCI [redacted]

CCI [redacted]

13. APPENDICES

CCI

CCI

[redacted]

CCI

CCI

CCI

CCI

13.2. CTCAE Grading Details from the Central Laboratory

CTCAE CHECKS

NCI Common Terminology Criteria (CTC) for Adverse Events (AE)
Laboratory Tests

The NCI CTC for AE is a descriptive terminology which can be utilized for AE reporting. A grading (severity) scale is provided for each AE term. (CTCAE, Version 4.03)

Grade I: Mild AE
Grade II: Moderate AE
Grade III: Severe AE
Grade IV: Life-threatening or disabling AE

LLN = Lower Limit of Normal range
ULN = Upper Limit of Normal range

NOTES: The ranges in this table are adjusted to the precision and units used by BARC. In case the grading ranges are in absolute values, the ranges are listed in the different units in which the parameter can be reported. In case the grading ranges are expressed as a factor of the lower and/or upper limit of the normal ranges, no units are shown. It is possible that in some cases, the lower or upper limit of normal range is at the same time a Grade I AE.
Please note CTC checks only apply on parameters taken within the scope of the same visit.
More information about NCI CTC can be accessed on the CTEP website "http://ctep.cancer.gov/reporting/ctc.html"

Parameter	Grade I	Grade II	Grade III	Grade IV	Units
		HEMATOLOGY			
CD4 count	500 - < LLN[4]	200 - < 500	50 - < 200	< 50	/mm³
Hemoglobin	10.0 - < LLN	8.0 - < 10.0	< 8.0	-	g/dL
	100 - < LLN	80 - < 100	< 80	-	g/L
	ULN + > 0.0-2.0	ULN + > 2.0-4.0	ULN + > 4.0	-	g/dL
	ULN + > 0-20	ULN + > 20-40	ULN + > 40	-	g/L
Leukocytes	3.0 - < LLN	2.0 - < 3.0	1.0 - < 2.0	< 1.0	10^3/µL or 10^9/L
	-	-	> 100.0	-	10^3/µL or 10^9/L
Lymphocytes	0.80 - < LLN	0.50 - < 0.80	0.20 - < 0.50	< 0.20	10^3/µL or 10^9/L
	-	> 4.00 – 20.00	> 20.00	-	10^3/µL or 10^9/L
Neutrophils	1.50 - < LLN	1.00 - < 1.50	0.50 - < 1.00	< 0.50	10^3/µL or 10^9/L
Platelets	75 - < LLN	50 - < 75	25 - < 50	< 25	10^3/µL or 10^9/L
		COAGULATION			
Fibrinogen	0.75xLLN - < LLN	0.5xLLN - < 0.75xLLN	0.25xLLN - < 0.5xLLN	< 0.25xLLN	
PT (INR)	> ULN - 1.5xULN	> 1.5xULN – 2.5xULN	> 2.5xULN	-	
APTT	> ULN - 1.5xULN	> 1.5xULN – 2.5xULN	> 2.5xULN	-	
		BIOCHEMISTRY			
Albumin	3.0 - < LLN	2.0 - < 3.0	< 2.0	-	g/dL
	30 - < LLN	20 - < 30	< 20	-	g/L
Alkaline phosphatase	> ULN - 2.5xULN	> 2.5xULN - 5xULN	> 5xULN - 20xULN	> 20xULN	
ALT	> ULN – 3xULN	> 3xULN - 5xULN	> 5xULN - 20xULN	> 20xULN	
Amylase	> ULN - 1.5xULN	> 1.5xULN - 2xULN	> 2xULN - 5xULN	> 5xULN	
AST	> ULN – 3xULN	> 3xULN - 5xULN	> 5xULN - 20xULN	> 20xULN	
Bilirubin, total	> ULN - 1.5xULN	> 1.5xULN - 3xULN	> 3xULN - 10xULN	> 10xULN	

Parameter	Grade I	Grade II	Grade III	Grade IV	Units
Corrected calcium	8.0 - < LLN	7.0 - < 8.0	6.0 - < 7.0	< 6.0	mg/dL
	2.00 - < LLN	1.75 - < 2.00	1.50 - < 1.75	< 1.50	mmol/L
	> ULN - 11.5	> 11.5 - 12.5	> 12.5 - 13.5	> 13.5	mg/dL
	> ULN - 2.90	> 2.90 - 3.10	> 3.10 - 3.40	> 3.40	mmol/L
Cholesterol, total	> ULN - 300	> 300 - 400	> 400 - 500	> 500	mg/dL
	> ULN - 7.75	> 7.75 - 10.34	> 10.34 - 12.92	> 12.92	mmol/L
Creatine kinase	> ULN - 2.5xULN	> 2.5xULN - 5xULN	> 5xULN - 10xULN	> 10xULN	
Creatinine (serum)	> ULN - 1.5xULN	> 1.5xULN - 3xULN	> 3xULN - 6xULN	> 6xULN	
eGFR	60 - < LLN[4]	30 - 59	15 - 29	< 15	mL/min (/1.73m^2)
GGT	> ULN - 2.5xULN	> 2.5xULN - 5xULN	> 5xULN - 20xULN	> 20xULN	
Glucose (2h PP; random)	55 - < LLN[1]	40 - < 55	30 - < 40	< 30	mg/dL
	3.00 < LLN[1]	2.20- < 3.00	1.70- < 2.20	< 1.70	mmol/L
Glucose (fasting)	55 - < LLN	40 - < 55	30 - < 40	< 30	mg/dL
	3.00 < LLN	2.20- < 3.00	1.70- < 2.20	< 1.70	mmol/L
	> ULN - 160	> 160 - 250	> 250 - 500	> 500	mg/dL
	> ULN - 8.90	> 8.90 - 13.90	> 13.90 - 27.80	> 27.80	mmol/L
Haptoglobin	< LLN	-	-	-	
Lipase	> ULN - 1.5xULN	> 1.5xULN - 2xULN	> 2xULN - 5xULN	> 5xULN	
Magnesium	1.2 - < LLN	0.9 - < 1.2	0.7 - < 0.9	< 0.7	mg/dL
	0.50 - < LLN	0.40 - < 0.50	0.30 - < 0.40	< 0.30	mmol/L
	> ULN – 3.0	-	> 3.0 – 8.0	> 8.0	mg/dL
	> ULN - 1.23	-	> 1.23 - 3.30	> 3.30	mmol/L
Phosphorus	2.5 - < LLN	2.0 - < 2.5	1.0 - < 2.0	< 1.00	mg/dL
	0.80 - < LLN	0.60 - < 0.80	0.30 - < 0.60	< 0.30	mmol/L
Potassium	3.0 - < LLN[2]	3.0 - < LLN[2]	2.5 - < 3.0	< 2.5	mmol/L
	> ULN - 5.5	> 5.5 - 6.0	> 6.0 - 7.0	> 7.0	mmol/L
Sodium	130 - < LLN	-	120 - < 130	< 120	mmol/L
	> ULN - 150	> 150 - 155	> 155 - 160	> 160	mmol/L
Triglycerides	150 - 300	> 300 - 500	> 500 - 1000	> 1000	mg/dL
	1.71 - 3.42	> 3.42 - 5.70	> 5.70 - 11.40	> 11.40	mmol/L
Troponin I	> 0.01 - < 0.30 (♀)		≥ 0.30 (♀ & ♂)		ng/mL or µg/L
	> 0.03 - < 0.30 (♂)				ng/mL or µg/L
Troponin T	> 0.01 - <0.10	-	≥ 0.10	-	ng/mL or mg/L
Uric acid	> ULN - 10.0[3]	-	> ULN - 10.0[3]	> 10.0	mg/dL
	> ULN - 590[3]	-	> ULN - 590[3]	> 590	µmol/L
URINALYSIS					
Creatinine clearance	60 - < LLN	30 - 59	15 - 29	< 15	mL/min
Blood	+(++)	-	-	-	
Protein	+	++ or +++ (≥ 18y)	-	-	
	< 1000	1000 – 3499 (≥18y)	≥ 3500 (≥ 18y)	-	mg/24h
Protein/creatinine ratio (< 18y)		0.5 – 1.9	> 1.9		

[1] Glucose: BARC reference range has no LLN, therefore 74 mg/dL (4.11 mmol/L) is applied as LLN in Grade I flagging (reference Tietz Clinical guide to laboratory tests 4th ed.)
[2] Potassium: Grade II when intervention is indicated, else Grade I
[3] Uric acid: Grade III, when in combination with physiologic consequences, else Grade I
[4] CD4 count & eGFR: LLN is lower or equal to lower threshold for Grade I flagging, therefore only the exact lower threshold value is flagged as Grade I

13.3. Table of Contents for Tables, Listings, and Figures

The table of contents for tables, listings, and figures will be provided in a corresponding mock shell document.

BOS161721-02: A RANDOMIZED DOUBLE-BLIND PHASE 1b/2 COMBINED STAGGERED MULTIPLE DOSE ESCALATION STUDY OF BOS161721 IN SYSTEMIC LUPUS ERYTHEMATOSUS (SLE) PATIENTS ON A BACKGROUND OF LIMITED STANDARD OF CARE

Statistical Analysis Plan

VERSION 1.0
DATE OF PLAN:

05Sep2018

STUDY DRUG:
BOS161721

PREPARED FOR:

Boston Pharmaceuticals, Inc.

Sponsor: Boston Pharmaceuticals, Inc.
Protocol Number: BOS161721-02
SAP Version and Date: Version 1.0, 05Sep2018

Approval Signature Page: Array Biostatistics, LLC

PPD

PPD

Date

Document Author:
PPD , MS
Senior Biostatistician

PPD

PPD

Date

PPD , MS
Vice President, Biostatistics & Statistical Programming

Sponsor: Boston Pharmaceuticals, Inc.
Protocol Number: BOS161721-02
SAP Version and Date: Version 1.0, 05Sep2018

Approval Signature Page: Boston Pharmaceuticals, Inc.

PPD

PPD

Date

Clinical Operations Lead

PPD

PPD

Date

Vice President, Clinical Development

Contents

ABBREVIATIONS

ACR	American College of Rheumatology
ADA	Anti-drug antibody
AE	Adverse event
AESI	Adverse event of special interest
ANA	Antinuclear antibody
ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
APL	Antiphospholipid
AUC	Area under the curve
BICLA	BILAG-based Composite Lupus Assessment
BILAG	British Isles Lupus Assessment Group
BP	Blood pressure
C3	Complement C3
C4	Complement C4
C_L	Systematic clearance
CLASI	Cutaneous Lupus Area and Severity Index
C_{max}	Maximum plasma concentration
CPAP	Clinical Pharmacology Analysis Plan
CS	Corticosteroid
CSR	Clinical study report
C-SSRS	Columbia Suicide Severity Rating Scale
CTCAE	Common Terminology Criteria for Adverse Events
CV	Coefficient of variation
DLT	Dose-limiting toxicity
DMC	Data Monitoring Committee
dsDNA	Double-stranded deoxyribonucleic acid
ECG	Electrocardiogram
CCI	[redacted]
FAS	Full analysis set
HR	Hazard ratio
IA	Interim Analysis
IEC	Independent Ethics Committee
IL-21	interleukin 21
IRB	Institutional Review Board
ISR	Injection site reaction
IWRS	Interactive web response system
LLN	Lower limit of normal
LLOQ	Lower limit of quantification
MAD	Multiple Ascending Dose
MedDRA	Medical Dictionary for Medical Affairs
nAb	Neutralizing antibody
NCI	National Cancer Institute
PD	Pharmacodynamics

PGA	Physician's Global Assessment
PK	Pharmacokinetics
POC	Proof of concept
PP	Per protocol set
pSTAT3	Phosphorylated signal transducer and activator of transcription 3
PT	Preferred Term
SAD	Single ascending dose
SAE	Serious adverse events
SAP	Statistical analysis plan
SAS	Statistical analysis system
SC	Subcutaneously
CCI	[redacted]
SLE	Systemic Lupus Erythematosus
SLEDAI-2K	SLE Disease Activity Index 2000
SLICC	Systemic Lupus International Collaborating Clinics
Sm	Smith
SOC	System Organ Class
SRI-4	SLE Responder Index 4
SRI-5	SLE Responder Index 5
SRI-6	SLE Responder Index 6
SS	Safety analysis set
SSA	Sjögren syndrome-A (Ro)
SSB	Sjögren syndrome-B (La)
$t_{1/2}$	Terminal elimination half-life
TEAE	Treatment emergent adverse event
T_{max}	First time to maximum concentration
TTMF	Time to medication failure
ULN	Upper limit of normal
V_d	Volume of distribution
WHO	World Health Organization

1. INTRODUCTION

The statistical analysis plan (SAP) details the planned analyses that will be included in the Clinical Study Report (CSR) of study number BOS161721-02: A Randomized Double-Blind Phase 1b/2 Combined Staggered Multiple Dose Escalation Study of BOS161721 in Systemic Lupus Erythematosus (SLE) Patients on a Background of Limited Standard of Care. The content of this SAP is based on the protocol Version 4.0 (Amendment 3) dated 27July2018.

Revision Chronology:

V1.0	05SEP2018	Original

2. STUDY OBJECTIVES AND ENDPOINTS

2.1. Multiple Ascending Dose (MAD) Phase 1b Study

Objectives	Endpoints
Primary	
• To assess safety, tolerability, and immunogenicity of repeat doses of BOS161721 (20, 60, and 120 mg) administered subcutaneously (SC) in adult patients with moderately to severely active SLE on limited background standard of care treatment, in order to estimate the optimal dose.	Safety Endpoints • Incidence and severity of adverse events (AEs) and serious adverse events (SAEs), related AEs, AEs leading to study drug discontinuation, AEs by severity and relatedness • Injection site reactions • Columbia Suicide Severity Rating Scale (C-SSRS) • 12-lead electrocardiograms (ECGs) parameter results at each visit and change from baseline • Vital signs (blood pressure [BP], heart rate, and temperature) parameter results at each visit and change from baseline • Clinical laboratory results and change from baseline • Physical examinations changes from baseline • Anti-drug antibodies (ADAs) • Study drug exposure/compliance

Secondary	
• To characterize the PK of BOS161721 and select the optimal dose of BOS161721 based on safety, PK, and PD effects in patients with mild to moderate SLE.	Pharmacokinetic (PK) Endpoints • BOS161721 concentration by visit and time point • Maximum observed concentration (C_{max}), first time to maximum concentration (T_{max}), area under the concentration-time curve (AUC), $t_{½}$, systematic clearance (C_L), volume of distribution (V_d) Pharmacodynamic (PD) Endpoints • Results and changes (or shifts) from baseline to each visit in phosphorylated signal transducer and activator of transcription 3 (pSTAT3), C3 and C4 levels, and leukocyte immunophenotype • Results and changes (or shifts) from baseline in anti-double-stranded DNA (dsDNA), antinuclear antibodies (ANA), anti-Sjögren syndrome A and B (SSA, SSB), Smith (Sm), and antiphospholipid (APL) autoantibodies at each visit • Results and changes (or shifts) from baseline in abrogation of IL-21 gene signature at each indicated visit
Exploratory	
• CCI [redacted]	CCI [redacted]

2.2. Proof of Concept (POC) Phase 2 Study

Objectives	Endpoints
Primary	
• To demonstrate a superior effect of BOS161721 at the chosen dose compared with placebo for response on the SLE Responder Index 4 (SRI-4)	Primary Efficacy Endpoint • The proportion of patients with a SRI-4 response at Day 210

Secondary	
• To demonstrate a superior effect of BOS161721 at the chosen dose compared with placebo for response on clinical indicators of SLE activity, in adult patients with moderately to severely active SLE on limited background standard of care treatment	Secondary Efficacy Endpoints • The proportion of patients with: - SRI-4 response at each visit - SRI-5 and SRI-6 response at each visit - a sustained reduction of oral corticosteroid (CS) (≤ 10 mg/day and ≤ Day 0 dose) between Day 120 and Day 210 - new BILAG A flare or > 1 BILAG B flares relative to baseline through Day 210 - PGA worsening - a BICLA response - a CLASI response - medication failures • Results and changes from baseline in: - CLASI - swollen and tender joints ACR-28 - SLEDAI-2K - SLICC/ACR damage index • Time to medication failure • Duration of longest SRI-4 response • Time to first SRI-4 response • Time to BILAG A flare or > 1 BILAG B flare compared to baseline through Day 210
Safety	
• To assess safety and tolerability of repeat doses of BOS161721 (20, 60, and 120 mg) administered SC in adult patients with moderately to severely active SLE on limited background standard of care treatment	Safety Endpoints • Incidence and severity of adverse events (AEs) and serious adverse events (SAEs), related AEs, AEs leading to study drug discontinuation, AEs by severity and relatedness • Injection site reactions • C-SSRS • 12-lead ECGs parameter results at each visit and change from baseline • Vital signs (blood pressure [BP], heart rate, and temperature) parameter results at each visit and change from baseline

	• Clinical laboratory results and change from baseline • Physical examinations changes from baseline • ADAs • Study drug exposure/compliance
Exploratory	
• CCI	CCI
• CCI	• CCI
• CCI	• CCI

The order in which the endpoints for this trial will be discussed in subsequent sections is based on the phase 2 POC portion of the study. The phase 1b MAD endpoints will also be discussed within the context of the phase 2 POC portion of the study but will appear out of the specified order in the above table in an effort to maintain common concepts across both phases in one location within the SAP.

3. STUDY DESIGN

3.1. Study Design and Population

This is a Phase 1b/2 combined, randomized, multicenter, double-blind, placebo-controlled trial to study the clinical efficacy, safety, and pharmacokinetics (PK) of subcutaneously (SC) doses of BOS161721 in adult patients with moderately to severely active SLE. After successfully

completing a screening phase, eligible patients will be randomized to a specified dose of BOS161721 or placebo. The dosing schedule will be monthly.

The trial will consist of 2 double blinded- portions: MAD Phase 1b and POC Phase 2. Patients may receive a total of 7 SC monthly doses of study drug on Days 0, 30, 60, 90, 120, 150, and 180, followed by safety/efficacy follow-up visit at Days 210, and safety follow up visits at Days 240, and 270.

SLE disease activity assessment data will be centrally reviewed by medical monitor and sponsor to ensure the scores are clinically meaningful and compliant with specific definition. The scope of responsibility includes, but is not limited to, review and confirmation of "A" and "B" British Isles Lupus Assessment Group (BILAG) system organ disease, confirmation of clinical components of the SLE Disease Activity Index 2000 (SLEDAI-2K) score at screening and during the study, and cross validation- of the instruments used in this study to assess the disease activity. Further details on the content and methods of data reports by the medical monitor and sponsor will be outlined in the Medical Data Review Plan along with the processes and procedures that will be followed.

3.1.1. Multiple Ascending Dose Phase 1b

The MAD portion will consist of 3 cohorts:

- Cohort 1 (20 mg SC) will include 6 patients
    - 5 patients will receive BOS161721 (active group) and 1 patient will receive placebo (placebo group)
- Cohorts 2 (60 mg SC) and 3 (120 mg SC) will include 12 patients each
    - 9 patients in the active group and 3 in the placebo group

Doses selected for each of the 3 cohorts is based on a 90-day safety, tolerability, PK and PD data review from the Phase 1 single ascending dose (SAD) study (BOS161721-01) in healthy subjects. All doses selected for the MAD part of the study are projected not to exceed the mean exposure of that achieved in the SAD study.

The MAD portion of the study design is staggered, where after the 6 patients in Cohort 1 have received 2 doses and have completed 2 weeks of follow up- after the second dose, Cohort 2 begins dosing (after the data monitoring committee (DMC) evaluation of the safety and tolerability data from Cohort 1). Similarly, after 8 of the 12 patients in Cohort 2 have received 2 doses and have completed 2 weeks of follow up- after the second dose, Cohort 3 begins dosing after the DMC evaluation of the safety and tolerability data from Cohorts 1 and 2. Each cohort will continue at their assigned dose level through their respective 6-month treatment periods (See Study Schematic in the study protocol). If patients discontinue the study in a cohort prior to adequate safety follow-up, he/she may be replaced.

Criteria for dose escalation are further described in the DMC Charter.

3.1.2. Proof of Concept Phase 2

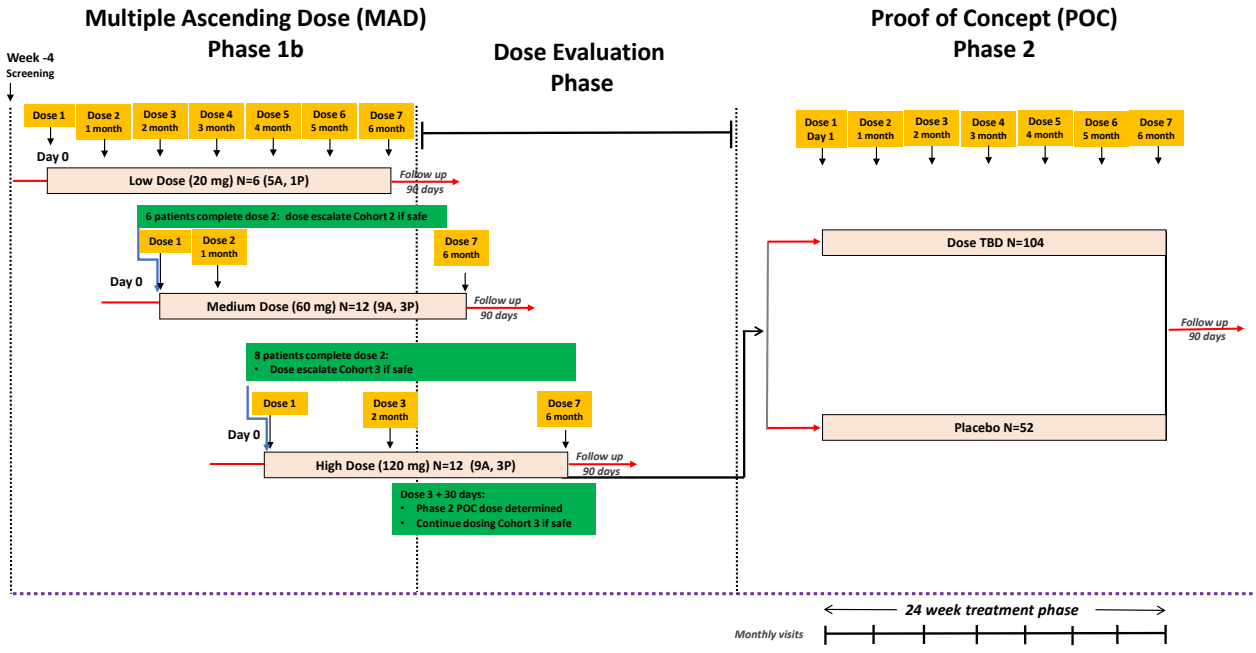
The optimal dose is chosen based on MAD Phase 1b safety, tolerability, immunogenicity, and PK and PD data. The dose will be communicated by a letter to site investigators participating in the POC Phase 2 portion, and to the Institutional Review Board / Independent Ethics Committee (IRB/IEC).

For the POC part of the study, approximately 156 additional patients will be randomized to active or placebo groups in a 2:1 ratio.

As in the MAD part of the study, each patient in the POC portion may receive a total of 7 SC monthly doses of study drug on Days 0, 30, 60, 90, 120, 150, and 180. Assessments will be completed according to the Schedule of Assessments detailed in the study protocol. Dose selection will be based on the DMC and sponsor's assessment of safety, tolerability, immunogenicity, and PK and PD data from the MAD portion of 30 patients treated with BOS161721/placebo.

DMC safety reviews will be conducted periodically throughout the study as described in the DMC charter.

Study Diagram for MAD Phase 1b and POC Phase 2



A = active drug (BOS161721); P = placebo

3.2. Randomization and Blinding

This is a randomized, double-blind study.

Patients meeting all inclusion and exclusion criteria will be centrally randomized to either placebo or BOS161721 using an IWRS according to a randomization list generated by an independent, non-study statistician. Randomization will be performed separately for each study portion and separately for each cohort in the Phase 1b and 2 portions as follows:

Phase/Cohort	Number of Patients	Randomization Ratio BOS161721:Placebo
Phase 1b/Cohort 1	6	5:1
Phase 1b/Cohort 2	12	3:1
Phase 1b/Cohort 3	12	3:1
Phase 2	156*	2:1

*Additional patients may be enrolled to ensure sufficient numbers of patients are in the full analysis set (FAS).

Eligible patients will be assigned to the study portion which is active at time of enrollment. Similarly, patients in the Phase 1 MAD will be assigned to the cohort which is active. Each patient will be assigned a unique randomization number which will not be reused.

All patients, investigators, and study team participants will be blinded to treatment assignment. An independent biostatistician not otherwise involved on the study will be unblinded and prepare materials for the interim analysis (IA) and the DMC safety reviews. The DMC will review unblinded data during safety reviews and the IA. A limited team at Boston Pharmaceuticals will review unblinded results from the Phase 1b MAD portion during the IA to determine the dose that will be used for the Phase 2 POC portion of the study. Details regarding maintenance of the blinding and content of data reviews will be described in the DMC charter or related study documentation.

3.3. Sample Size Considerations

Sample size in the Phase 1b MAD part of the study is based on operational consideration.

The sample size in the Phase 2 POC study is based on the primary endpoint, SRI-4 Response at Day 210. Approximately 156 additional patients will be randomized in a 2:1 ratio to BOS161721 versus placebo to achieve 132 evaluable patients in the FAS. This assumes 24 patients will drop prior to having received treatment and an evaluable post-baseline efficacy measure. Enrollment and randomization will be monitored in a blinded fashion and increased if needed to ensure at least 132 patients are randomized into the FAS.

A total of 132 patients randomized provides more than 80% power to detect a treatment difference of 25% in SRI-4 response rates at Day 210, based on a targeted 2-sided significance level of 10% and using a 2-sided Pearson's chi-squared test. This assumes a response rate of 65% for BOS161721 and 40% for placebo, where missing data at Day 210 is carried forward

from most recent non-missing result and patients are treated as a non-responder if a prohibited medication defined as a medication failure occurs.

3.4. Data Monitoring Committee (DMC)

This study will use an external DMC. The DMC is an independent committee established to provide oversight of safety considerations in the study and to provide advice to the sponsor regarding actions the committee deems necessary for the continuing protection of enrolled patients and those yet to be recruited to the trial as well as for the continuing validity and scientific merit of the study results. The DMC is charged with assessing such actions in light of an acceptable benefit/risk profile for BOS161721. The recommendations made by the DMC (i.e., dose escalation, etc.) will be forwarded to the sponsor for final decision. The sponsor will forward such decisions, which may include summaries of safety data which are not endpoints, to regulatory authorities, as appropriate.

The sponsor will appoint a DMC for the periodic review of available study data. The DMC is an independent group of experts that advises the sponsor and the study investigators. The members of the DMC serve in an individual capacity and provide their expertise and recommendations. The primary responsibilities of the DMC are to (1) periodically review and evaluate the accumulated study data for participant safety, study conduct, and progress, (2) make recommendations to the sponsor concerning the continuation, modification, or termination of the study and (3) suggest dose for POC portion of the study as described in the study protocol.

The DMC considers study-specific data as well as relevant background knowledge about the disease, test agent, or patient population under study. The DMC is responsible for defining its deliberative processes, including event triggers that would call for an unscheduled review, stopping guidelines, unmasking (unblinding), and voting procedures prior to initiating any data review. The DMC is also responsible for maintaining the confidentiality of its internal discussions and activities as well as the contents of reports provided to it.

The DMC will have access to unblinded treatment information during the clinical trial. Details regarding management and process of this committee are found in the DMC Charter.

The DMC may recommend termination of BOS161721 treatment arm or the entire BOS161721 MAD/POC trial for any safety concern that is felt to outweigh potential benefits. The recommendation must be supported by the sponsor as indicated in the DMC Charter.

3.5. Interim Analysis (IA)

One interim analysis is planned for this study and will be conducted at the time of the POC decision for dose selection. The POC dose will be selected based on the MAD data. Since there is no IA during the POC part of the study, there is no impact on the type 1 error.

Interim analysis outputs will be a subset of those to be performed for the final analysis.

3.6. Timing of Analyses

The following analyses are planned:

- An IA will be performed during the last cohort of the MAD portion to determine dose selection for the POC portion.
- The final analysis will be performed when all patients have either completed the POC safety follow up or withdrawn from the study.
- Safety analyses will be performed for DMC reviews throughout the study to evaluate dose escalation decisions and accumulating safety data. The frequency and details of the content and format of the safety review meetings are described in the DMC charter and SAP/mock shells.

4. DATA ANALYSIS CONSIDERATIONS

All analyses will be conducted based on SAS 9.3 or higher.

All data in the database will be presented in by-patient data listings.

Unless otherwise stated, all listings will be sorted by study phase, treatment group, patient number, and assessment date (and time, if available). If assessment date is missing, chronological visit will be used for sorting.

Unless stated otherwise, continuous data will be summarized by treatment group based on number of patients (n), mean, median, standard deviation (SD), 25^{th} percentile/first quartile (Q1), 75^{th} percentile/third quartile (Q3), minimum value, and maximum value.

The geometric % coefficient of variation is calculated as 100*sqrt(exp(SDlog^2)-1) where SDlog is standard deviation of log-transformed values.

Unless stated otherwise, categorical data will be summarized by treatment group using n and percentage based on the number of non-missing values.

- The number of missing values will be presented as a separate category with no percentage, but only if one or more patients are missing data.
- Counts of zero will be presented without percentages.

Statistics will be presented in the summary tables based on the following:

- Minimum and Maximum: same number of significant digits as the raw value
- Mean, Median, Q1, and Q3: one additional decimal place to that reported for Minimum and Maximum
- SD: two additional decimal places than presented for the Minimum and Maximum

- Percentages: reported to one decimal place with the exception of 100% which will be displayed without decimals.
- P-values will be reported to four decimal places. If the value is below 0.0001 it will be noted as < 0.0001; if the value above 0.9999 it will be noted as > 0.9999.

Unless otherwise noted, statistical inference will be based on a 10% significance level (i.e. 90% confidence intervals will be produced).

Statistical testing will be performed on data from the Phase 2 study only.

All data up to the time of study completion/withdrawal from the study will be included in the analysis, regardless of duration of treatment.

Numbering for data displays will be based on ICH E3.

4.1. Stratification and Covariates

The effects of noncompliance, background therapy use, treatment discontinuations, premature withdrawal from study and covariates will be assessed to determine the impact on the general applicability of results from this study. Exploratory analyses of the data may be conducted as deemed appropriate to evaluate factors for which analyses are not controlled.

Continuous efficacy endpoints will be assessed via analysis of variance (ANOVA) or analysis of covariance (ANCOVA) when applicable, adjusting for the baseline value.

4.2. Evaluation of Subgroups

There are no formal plans for examining subgroups.

4.3. Multiple Comparisons and Multiplicity

The Type 1 error rate will not be adjusted for multiplicity.

5. GENERAL DATA HANDLING CONVENTIONS

5.1. Assigned and Actual Treatment

Patients are assigned to a study treatment and cohort based on the study schematic in Section 3.

Full analysis set (FAS) and per protocol (PP) analyses will be conducted on the basis of the randomized (assigned) treatment. Safety analyses will be conducted on the basis of the actual treatment received based on the exposure data collected in the case report form (CRF).

5.2. Reference Dates

- Screening date is defined as the CRF provided date on which a patient was screened for trial entry.
- Randomization date is defined as the date on which the patient is randomized to study treatment.
- Treatment start date is defined as the date of first dose of study drug.
- Treatment end date is defined as the date of last dose of study drug.
- Age will not be calculated and will come directly from the CRF. The CRF uses the informed consent date as its reference date for age calculation.
- Safety data, such as AEs and laboratory assessments will use the treatment start date as a reference date.
- Efficacy data will use the randomization date as a reference date.
- Study day will be based on treatment start date as a reference date.

5.3. Study Day and Duration Variables

Reference date calculations will be defined as follows, assuming non-missing dates:

- date of interest – reference date + 1 when the date of interest is on or after the reference date;
- otherwise, date of interest – reference date.

If either date is missing or incomplete, reference date calculations will not be performed. Date imputation will be performed as identified in Section 5.6.

Study day will either have a negative value if collected before dosing or a positive value if collected on or after the day of drug dosing; there will be no study day zero.

Duration of time is dependent on reference dates and will be calculated in a manner similar to that of the reference date calculation. Duration on study is defined as the end of study date – randomization date + 1. Duration of treatment is defined as treatment end date – treatment start date + 1, where treatment end date is the date of last dose of study drug.

Time-to-event endpoints are followed until first event or censoring. As a result, event time will be calculated as: event or censoring date – reference date + 1. These are further described in Section 7.

When reporting time-to-event data (Time to Medication Failure, Time to first SRI-4 Response, Time to BILAG A Flare or >1 BILAG B flare) or duration outcomes, the results (in days) above will be converted to an appropriate unit. When reporting in months it will be divided by 30.4375; for reporting in weeks it will be divided by 7; and for reporting in years it will be divided by 365.25.

5.4. Study Time Periods

Where applicable, data reporting will be classified by the following study periods for analysis:

- Pre-treatment is defined as the period prior to a patient's treatment start date.
- On-treatment is defined as the period between a patient's treatment start date and treatment end date, inclusive.
- Post-treatment is defined as the period of time following the on-treatment period.

5.5. Baseline, Post-Baseline Changes, and Last Observation Carried Forward (LOCF)

Unless stated otherwise, baseline and post-baseline change values will be based on the following:

- Baseline will be based on the last non-missing value collected prior to or on the treatment start date (and time, if applicable).
- Post-baseline values will be those collected after the date/time of first dose of study drug.
- Change from baseline is defined as: value – baseline value.
- Most extreme change: The maximum most extreme change will be the maximum post-baseline value; the minimum most extreme change will be the smallest post-baseline value. This calculation will consider all assessments collected after the first dose of study drug, scheduled or unscheduled.

LOCF will be based on the last non-missing value collected. This calculation will consider all assessments, scheduled or unscheduled, including baseline. If a patient withdraws from the study, their most recent non-missing assessment prior to withdraw will be carried forward. Additional details regarding LOCF for specific endpoints are provided in Section 7.

Observed values do not incorporate any imputations and are presented as collected.

5.6. Imputation of Partial Dates

Adverse Events

- If the AE start date is completely missing, or if the patient was not treated, no imputation will be conducted.
- If the AE start date is missing day and month, do the following:
    - If the treatment start date is missing or the AE start year does not fall in the same year as that of the treatment start date or if the AE record contains information to indicate that the event ended before the treatment start date (e.g. the AE end date month and year are earlier than the treatment start date or the full AE end date is

known and occurs earlier than the treatment start date), then set the AE start month and day to January 1st.
    - Otherwise, set the AE start date to the treatment start date.
- If only the AE start day is missing, do the following:
    - If the study treatment start date is missing or the AE start month and year does not fall in the same month/year as that of the treatment start date or if the AE contains information to indicate that the event ended before the treatment start date, then set the AE start month and day to the 1st day of the month of the treatment start date.
    - Otherwise, set the AE start date to the treatment start date.
- AE end dates will not be imputed.

Prior, Concomitant, and Subsequent Medications

- The imputation rules for AE start dates will be used for medication start dates.
- Medication stop dates will be imputed as follows:
    - If the stop date is only missing the day, then the stop day is the last day of the month
    - If the stop date is missing both the day and month and the year matches the last study date, then the stop month and day is the earlier of the last study date and December 31
    - If the stop date is missing both the day and month and the year is earlier than the last study date, then the stop month and day is December 31
    - If the stop date is completely missing, no imputation is performed, and the medication will be classified as a concomitant medication for subjects who were treated unless the medication began after the end of study treatment.

5.7. Multiple Assessments and Visit Windows

Nominal visits (e.g. those identified by the study CRF) will be the basis of summarization and statistical analysis; no visit date windowing will be conducted. Unscheduled data may be included in summaries of most extreme, baseline, and endpoint values; summaries of specific abnormalities any time post-baseline; and patient data listings

5.8. Treatment Group Display

Treatment groups will be displayed with the following columns. See the mock shells for additional details and a visual representation of the treatment group display.

- Phase 1b MAD: Cohort 1 BOS 20mg
- Phase 1b MAD: Cohort 2 BOS 60mg
- Phase 1b MAD: Cohort 3 BOS 120mg
- Phase 1b MAD: Placebo

- Phase 1b MAD: Total BOS (All patients treated with BOS161721, excludes placebo)
- Phase 2 POC: BOS <dose> where <dose> is the value determined during the IA
- Phase 2 POC: Placebo
- Overall: Total BOS <dose>mg where <dose> is the common dose between the Phase 1b MAD and Phase 2 POC portions of the study (to be determined during the IA). If a common dose is not determined, this column will be omitted.
- Overall: Total BOS (Patients treated with BOS161721 in MAD and POC, excludes placebo)
- Overall: Total Placebo (Placebo patients in MAD and POC, excludes patients treated with BOS161721)

5.9. Missing Data

Missing data imputation for the efficacy endpoints are discussed in the corresponding efficacy sections. AE and concomitant medication date imputations are described in Section 5.6. Missing data for noncompartmental analysis of pharmacokinetics will follow standard imputation methods, as described in a separate Clinical Pharmacology Analysis Plan (CPAP). Otherwise, missing data will not be imputed.

6. STUDY PATIENT DATA

6.1. Analysis Populations/Sets

Full Analysis Set (FAS): Defined as all patients who receive at least one dose of study treatment and have at least one evaluable post-baseline efficacy evaluation (BILAG, PGA, SLEDAI-2K, CLASI, ACR-28, SLICC/ACR, CCI , or CCI). FAS analyses will be conducted on the basis of the randomized treatment. The FAS will be used as the basis for the primary efficacy analysis.

Safety Analysis Set (SS): Defined as all patients who receive at least one dose of study treatment. Safety analyses will be conducted on the basis of actual treatment received. The SS will be the basis of all safety reporting.

Per Protocol Set (PP): Defined to exclude major protocol violations from the efficacy analysis. The PP will include all patients from the FAS except those with major violations to the protocol deemed to impact the analysis of the primary endpoint. These violations will be identified based on blinded data prior to study unblinding. PP analyses will be conducted on the basis of the randomized treatment. Additional information on inclusion into this population can be found in Section 6.3.

Pharmacokinetic Analysis Set (PK): Defined as the FAS patients who received at least one dose of BOS161721 with sufficient concentration data for the calculation of at least one PK parameter.

6.2. Patient Disposition

Disposition data will be summarized for all randomized patients. The number of patients in each analysis population, the number of patients who discontinued study treatment including reasons, and the number of patients who discontinued the study including reasons will be summarized. The number of patients who completed the study will also be summarized. Data will be presented by study phase and treatment group.

A by-patient listing of patient disposition data including reason for discontinuation, if applicable, will be presented for all randomized patients. A by-patient listing of randomization details will also be provided for all randomized patients.

A by-patient listing of inclusion/exclusion criteria and screen failures will be produced for all enrolled patients. Screen failures will otherwise not be included in any analyses.

6.3. Protocol Deviations

Protocol deviations will be identified and classified as major violations before database lock and unblinding. Major protocol deviations may include but are not limited to:

- Violation of Inclusion/Exclusion Criteria
- Use of prohibited therapies
- Incorrect treatment

Protocol deviations will be summarized by protocol version, deviation category, and major/minor designation. Protocol deviations will be summarized by study phase and treatment group for all randomized patients.

A listing of protocol deviations will be provided for all randomized patients.

6.4. Demographic and Baseline Characteristics

Patient demographics and baseline characteristics will be summarized in the SS. These will include age (years), gender (Male / Female), ethnicity (Hispanic or Latino / Not Hispanic or Latino), race (American Indian or Alaska Native / Asian / Black or African American / Native Hawaiian or Pacific Islander / White / Other), baseline height (cm), baseline weight (kg), and baseline BMI (kg/m^2). Age is reported as collected in the clinical database and will also be categorized as a categorical variable ($\leq$45, >45-<65, $\geq$65) for reporting. Patient eligibility for central review will be summarized. Childbearing potential and birth control methods for female and male patients will be summarized for the DMC meetings and interim analysis only. Demographics and baseline characteristics will be summarized by study phase and treatment group.

Demographics and baseline characteristics will be listed for all randomized patients. A separate by-patient listing will be created for childbearing potential and birth control for all randomized patients.

The following conversions and equations will be used as applicable:

Height (in cm) = height (in inches) * 2.54

Weight (in kg) = weight (in lbs) * 0.4536

BMI (kg/m^2) = weight(kg)/[height$(m)^2$]

6.5. Medical History

Medical history will be presented in data listings for all randomized patients. Medical history coded terms will be provided, including the system organ class and preferred term. The data management plan specifies the version of MedDRA used for medical history coding.

6.6. Prior and Concomitant Medication and Procedures

The incidence of medication use will be summarized by WHO Drug Dictionary anatomic therapeutic class (ATC) Level 2 classification (i.e. therapeutic main group) and preferred name. A patient will be counted only once at each level of reporting. Prior medications are those which have been identified to have been discontinued prior to the first dose of study medication. Concomitant medications are those which have been identified to have been taken at any point during the period of time between the first and last dose of study medication, including medications which started prior to first dose of study medication but are ongoing at first dose. Subsequent medications are those which have been started after the last dose date. The data management plan specifies the version of WHO Drug used.

Partial dates will be imputed according to Section 5.6 for the determination of prior, concomitant, and subsequent medications.

Concomitant medication use will be summarized and presented by study phase and treatment group in the SS.

All prior and concomitant medication data will be provided together in a by-patient listing including the verbatim and preferred drug name and WHO ATC Class (ATC Level 2) for all randomized patients. Past biological, immunosuppressant, antimalarial and corticosteroid SLE medication history data will be listed separately for all randomized patients.

Concomitant procedures will be presented in a data listing for all randomized patients.

6.7. Study Drug Exposure and Compliance

The total dose administered (mL) and the number of doses received will be summarized as continuous variables. The number and percentage of patients dosed and the number of patients who received the total planned dose at each nominal, protocol specified, visit (Day 0 / Day 30 / Day 60 / Day 90 / Day 120 / Day 150 / Day 180) will be summarized. The denominator for the number of patients dosed will be the number of patients in the SS. The denominator for the number of patients who received the total planned dose will be the number of patients dosed at the corresponding nominal visit.

The percent compliance will be calculated descriptively and is defined as 100 * (total actual dose summed across all visits)/total planned dose summed across all visits).

The duration of exposure in days will also be summarized descriptively and is defined as the last dose date – first dose date +1.

Study drug exposure and compliance will be presented by study phase and treatment group in the SS.

By-patient listings of BOS161721 and placebo dosing data will be produced for all patients in the SS.

7. EFFICACY

7.1. Primary Efficacy Endpoint and Analyses

Safety is the primary endpoint for the Phase 1b MAD portion of the study and is discussed in Section 9. Efficacy will be evaluated for the Phase 1b MAD portion of the study as an exploratory endpoint and details will be included in relation to the Phase 2 POC portion of the study within this section.

7.1.1. The proportion of patients with a SRI-4 response at Day 210

The primary efficacy endpoint is the proportion of patients who achieve a SRI-4 response at Day 210.

The SRI-4 is a composite index of SLE disease improvement that consists of scores derived from the SLEDAI-2K and the BILAG 2004 Index. Response based on the SRI-4 is defined by:

1) ≥ 4-point reduction from baseline in SLEDAI-2K global score

AND

2) No new severe disease activity (BILAG A organ score) or more than 1 new moderate organ score (BILAG B) compared with baseline

AND

3) No deterioration from baseline in the PGA by ≥ 30mm.

The proportion of patients who achieve a SRI-4 response at Day 210 will be presented and assessed using Pearson's chi-square test. The number of SRI-4 responders, treatment difference versus placebo, odds ratio, and 90% confidence interval will be displayed. The 90% confidence interval will be calculated with Wilson (score) confidence limits for the binomial proportion by treatment group. If any of the expected cell counts are less than 5, Fisher's exact test will be used and exact confidence intervals will be constructed. The analyses will be repeated for each of the three components contributing to the SRI-4 response. The primary analysis of the primary efficacy endpoint will be performed in the FAS and presented by treatment group. The analysis will be repeated in the PP analysis set.

The superiority of BOS161721 relative to placebo will be evaluated. If the proportion of SRI-4 responders for the selected POC dose in BOS161721 arm is higher than that of the placebo arm, then BOS161721 will be considered superior to placebo if the 2-sided p-value is less than 0.10.

Evaluability/Non-Responders:

Patients who received prohibited medications or unallowable corticosteroid (CS) usage will be considered "medication failures" and will be treated as non-responders for the primary efficacy analysis. Prohibited medications are detailed in Appendix 4 of the study protocol. Prohibited medications and unallowable CS usage will be recorded as protocol deviations in the clinical database.

A patient is considered to have unallowable CS usage if any of the following is not met:

- After Day 0 (after initiation of study therapy), no up-titration above 10 mg/day is allowed except for up to 1 CS burst for increased disease activity.

- Tapering is allowed after randomization except within 60 days of the primary (Day 210) and secondary (Day 120) endpoint assessments. Between Day 60 and Day 120, and between Day 150 and Day 210, oral CS doses must be held constant.

- A maximum of 1 oral CS "burst" for increased SLE disease activity will be allowed during the study between Day 0 and Day 60, according to the following:

    - An oral CS "burst" between Day 0 and Day 60; (an increase of ≤ 40 mg/day of prednisone or equivalent), which must be tapered down to a maximum of 10 mg/day within 2 weeks of initiation of the "burst"

- Alternatively, a single intramuscular (IM) dose of methylprednisolone (40 mg or equivalent) is permitted

- The course of the oral CS "burst" is not permitted to extend beyond Day 60

Patients who received an allowable CS burst but having missing data at Day 210 will be considered a medication failure and will be treated as non-responders for the primary efficacy analysis. The determination of medication failures will be reviewed using blinded data and finalized prior to unblinding. Once a patient is determined to be a non-responder, they will not be considered for response at a later visit.

If a patient has a BILAG, SLEDAI-2K, or PGA assessment outside of a 30 day +/-4 day window since the previous assessment (0-25 or >35 days since the prior assessment), the corresponding result will be considered not evaluable and will be excluded from all analyses. The excluded result will not contribute to LOCF analyses. In these cases, the individual assessment (i.e. BILAG, SLEDAI-2K, or PGA) will be considered as if it was missed and analyses will be handled per the below details.

Missing Data:

- If a patient is missing data during the pre-treatment period for SLEDAI-2K, PGA, or BILAG, the patient will be excluded from the primary efficacy analysis.

- Missing data at Day 210 for overall components for SLEDAI-2K, BILAG, and PGA will be addressed by employing a last observation carried forward (LOCF) analysis within each overall component as described in Section 5.5. For example, if a patient has data for SLEDAI-2K and BILAG at Day 210 but is missing PGA at Day 210, the PGA component would employ LOCF. Data collection and scoring will be performed in the clinical database for BILAG, SLEDAI-2K, and PGA. As such, missing individual items contributing to any of the overall components will not be imputed/carried forward.

- Missing data for visits other than Day 210 will be handled in the same way as Day 210.

- If one or more consecutive visits are missed, LOCF analyses will be carried forward from the most recent prior assessment. For example, if the Day 150, Day 180 and Day 210 assessments are all missed but the Day 120 assessment was performed, the Day 120 assessment would be carried forward to Days 150, 180, and 210 for analysis.

Several sensitivity analyses will be performed:

Analysis	**Criteria**
Sensitivity Analysis 1	A sensitivity analysis of the primary efficacy endpoint will be performed (including LOCF methods) for patients in the full analysis set excluding patients with:

	(1) Positive anti-drug antibodies at baseline through Day 210. AND (2) Confirmed nAb at Day 90 and Day 180 OR missing Day 90 and missing Day 180 nAb OR missing Day 90 nAb and confirmed nAb at Day 180. Missing ADA and missing nAb results will be handled with a LOCF approach as discussed in Section 5.5. Additional analyses exploring impact of baseline factors such as baseline disease severity, race, and other characteristics may be performed.
Sensitivity Analysis 2	A sensitivity analysis of the primary efficacy endpoint will be performed (including LOCF methods) for patients in the full analysis set; however, if a patient has a BILAG, SLEDAI-2K, or PGA assessment outside of a 30 day +/-4 day window since the previous assessment (0-25 or >35 days since the prior assessment), the corresponding result will be considered evaluable and will be included in the analysis and will contribute to LOCF.
Sensitivity Analysis 3	A sensitivity analysis of the primary efficacy endpoint will be performed (excluding LOCF methods) to consider the effect of missing data by excluding patients with either or both of the following criteria: (1) Missing assessments at Day 210 (patients with a missed assessment of any of BILAG, SLEDAI-2K, or PGA at Day 210 would be excluded). AND/OR (2) Patients with any Day 210 assessment which occurred outside of the 30 day +/- 4 day window (patients with an out of window assessment of any of BILAG, SLEDAI-2K, or PGA at Day 210 would be excluded).

Additional sensitivity analyses may be performed to consider the effect of missing data.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

A summary of disposition of SRI-4 response at day 210 will be presented by study phase and treatment group in the FAS and PP. Sensitivity analysis 1 will also be performed for disposition of SRI-4 response at day 210. These summaries will include the number of SRI-4 responders and the number of SRI-4 non-responders with a breakdown of reasons for non-response (early withdrawal, medication failure, <4-point reduction from baseline in SLEDAI-2K global score, new BILAG 1A/2B, and PGA worsening). Denominators for percentages will be the number of patients eligible for SRI-4 response analysis.

A listing of SRI Response data at Day 150, Day 180, and Day 210 will be presented in the FAS. A listing of patients excluded from the efficacy analysis will also be provided for all randomized patients.

7.2. Secondary Efficacy Endpoints and Analyses

Pharmacokinetics and pharmacodynamics are the secondary endpoints for the Phase 1b MAD part of the study and are described in Section 8. Efficacy will be evaluated for the Phase 1b MAD part of the study as an exploratory endpoint and details will be included in relation to the Phase 2 POC part of the study within this section.

7.2.1. Proportion of Patients with SRI-4 Response at Each Visit

The analyses described in Section 7.1.1, including sensitivity analyses allowing out of window assessment results (Sensitivity Analysis 2) and excluding patients with missing or out of window assessments (Sensitivity Analysis 3), will be repeated for each nominal visit and will be presented by study phase and treatment group in the FAS. The 3 components will not be summarized separately by visit.

A graphical display of SRI-4 response by study phase, treatment group, and visit in the FAS will be presented.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.2. Proportion of Patients with SRI-5 Response at Each Visit

The SRI-5 is a composite index of SLE disease improvement that consists of scores derived from the SLEDAI-2K and the BILAG 2004 Index. Response based on the SRI-5 is defined by:

1) ≥ 5-point reduction from baseline in SLEDAI-2K global score

AND

2) No new severe disease activity (BILAG A organ score) or more than 1 new moderate organ score (BILAG B) compared with baseline

AND

3) No deterioration from baseline in the PGA by ≥ 30mm.

The analyses described in Section 7.1.1, including sensitivity analyses excluding patients with missing or out of window assessments (Sensitivity Analysis 3), will be repeated for each nominal visit based on the SRI-5 response definition and will be presented by study phase and treatment group in the FAS. The 3 components will not be summarized separately by visit for SRI-5 response.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.3. Proportion of Patients with SRI-6 Response at Each Visit

The SRI-6 is a composite index of SLE disease improvement that consists of scores derived from the SLEDAI-2K and the BILAG 2004 Index. Response based on the SRI-6 is defined by:

1) ≥ 6-point reduction from baseline in SLEDAI-2K global score

AND

2) No new severe disease activity (BILAG A organ score) or more than 1 new moderate organ score (BILAG B) compared with baseline

AND

3) No deterioration from baseline in the PGA by ≥ 30mm.

The analyses described in Section 7.1.1, including sensitivity analyses excluding patients with missing or out of window assessments (Sensitivity Analysis 3), will be repeated for each nominal visit based on the SRI-6 response definition and will be presented by study phase and treatment group in the FAS. The 3 components will not be summarized separately by visit for SRI-6 response.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.4. Proportion of Patients with a sustained reduction of oral corticosteroid (CS) between Day 120 and Day 210

The proportion of patients with a sustained reduction of oral corticosteroid, defined as ≤ 10 mg/day and ≤ Day 0 dose, between Day 120 and Day 210 will be summarized by study phase and treatment group in the FAS. LOCF and imputation for missing values will not be performed. Otherwise, statistical methods described in Section 7.1.1 will be implemented.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.5. Proportion of new BILAG A flare or >1 BILAG B flare relative to baseline through Day 210

The proportion of new BILAG A flare or >1 BILAG B flare relative to baseline through Day 210 will be summarized by study phase and treatment group in the FAS. Statistical methods described in Section 7.1.1 will be implemented, including LOCF methods within a body system. The analyses will also be repeated in observed data only (no LOCF). A BILAG A flare is defined as a new A score in at least one body system when compared to baseline. A BILAG B flare is defined as a new B score in at least two body systems at the same assessment when compared to baseline. BILAG grading is collected at each protocol specified visit as an assessment of "the past 4 weeks". As such, the BILAG grade will be compared to baseline and the most recent prior assessment, as appropriate, to determine a new flare. The proportion of new BILAG A flare or >1 BILAG B flare relative to baseline will be summarized overall and by nominal visit.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

Observed results will be presented in separate data listings for BILAG SLE history, BILAG-2004 index, and BILAG grading for all randomized patients.

7.2.6. Proportion of patients with PGA worsening

The PGA is used to assess investigator's general impression on the patient's overall status of SLE disease activity via visual analogue scale (100 mm) with 0 being "very good, asymptomatic and no limitation of normal activities" with 100 mm being "most severe possible disease ever seen in all SLE patients." The proportion of patients with PGA worsening, defined as an increase of >30mm from baseline, will be summarized by study phase and treatment group in the FAS overall and by nominal visit. Statistical methods described in Section 7.1.1 will be implemented, including LOCF methods. The analyses will also be repeated in observed data only (no LOCF).

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

PGA observed results will be presented in a data listing for all randomized patients.

7.2.7. Proportion of patients with BICLA response

The BICLA is a responder index developed to measure response to therapy, and it includes scores from the BILAG, SLEDAI-2K, and PGA. BICLA response is defined as:

1) at least 1 gradation of improvement in baseline BILAG 2004 scores in all body systems with moderate disease activity at entry (e.g., all B [moderate disease] scores falling to C [mild], or D [no activity])

AND

2) no new BILAG A or more than 1 new BILAG B scores

AND

3) no worsening of total SLEDAI-2K score from baseline

AND

4) ≤ 10% deterioration in PGA score

AND

5) no medication failure.

LOCF and imputation for missing values will not be performed. Otherwise, statistical methods described in Section 7.1.1 will be implemented. The analyses will be repeated overall and for each nominal visit based on the BICLA response definition and will be presented by study phase and treatment group in the FAS.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.8. Cutaneous Lupus Erythematosus Area and Severity Index (CLASI)

The CLASI is a comprehensive tool for assessment of disease activity and damage in cutaneous lupus, shown to be valid, reliable, and sensitive to changes in disease activity.

The total activity score will be calculated by summing the scores on the left side of the form (erythema, scale/hypertrophy, mucous membrane, alopecia – recent hair loss and clinically not obviously scarred).

The total damage score will be calculated by:

(1) Sum the responses for “Dyspigmentation” in all anatomical locations

(2) If a response for “Report duration of dyspigmentation after active lesions have resolved” is “Dyspigmentation usually lasts at least 12 months”, then multiply the sum in #1 by 2.

(3) If a response for “Report duration of dyspigmentation after active lesions have resolved” is “Dyspigmentation usually lasts less than 12 months”, then the sum in #1 will not change.

(4) If the sum in #1 is 0 and a response for “Report duration of dyspigmentation after active lesions have resolved” is missing, then the sum in #1 will not change.

(5) Determine the Dyspigmentation Score based on the results from 2, 3, or 4.

(6) Sum the remaining scores on the right side of the form (scarring/atrophy/panniculitis, alopecia - scarring of the scalp judged clinically).

(7) Sum the scores from #5 and #6 to obtain the total damage score.

If the total activity score or total damage score cannot be calculated, the total activity score or total damage score will be imputed using the LOCF method described in Section 5.5. If one of the total scores can be calculated and the other cannot, LOCF will be implemented for only the total score which cannot be calculated. If any individual scores which contribute to the total activity score or the total damage score, the total score in question will not be calculated and LOCF methods for total score will be implemented. LOCF methods will be implemented for CLASI response analyses only.

A sensitivity analysis for CLASI response will also be performed in the FAS excluding patients with missing CLASI results within a visit. If one or both of the total activity score or total damage score are missing at a visit, the patient will not be included in the response analysis at that visit. LOCF methods will not be implemented for the sensitivity analysis.

The mean and standard deviation of the change from baseline in CLASI total activity and total damage scores will be presented graphically over time by study phase and treatment group.

7.2.8.1. Proportion of patients with CLASI response

CLASI response is defined as 50% improvement from baseline in "A" (total activity) or "B" (total damage) scores. This assessment will be applied to all patients as all are required to have cutaneous disease activity.

The analyses described in Section 7.1.1 will be repeated overall and for each nominal visit based on the CLASI response definition and will be presented by study phase and treatment group in the FAS.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.8.2. Results and Changes from Baseline in CLASI

The analysis of CLASI total activity score and CLASI total damage score will be performed separately using an analysis of covariance (ANCOVA) model with change from baseline to each nominal visit as the dependent variable, treatment group as the factor, and the corresponding baseline CLASI score as a covariate. The p-value for the difference between the BOS161721 and placebo groups at each nominal visit is based on the F-test.

Descriptive statistics for the activity score and damage score will be summarized separately at baseline and each visit. Change from baseline will be summarized at each post-baseline visit. The LS Mean estimate, its standard error and 90% CI will be tabulated by treatment group and for the difference between treatment groups (BOS161721 versus placebo) for change at each visit. Missing CLASI results will not be imputed. Observed results only will be presented for the change from baseline analysis.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

CLASI observed results will be presented in data listings for all randomized patients.

7.2.9. Results and Changes from Baseline in Swollen and Tender Joints ACR-28

The ACR-28 joint count evaluates the number of tender and swollen joints in the shoulder, elbow, wrist, hand, knee joints. Joints of the feet are excluded.

The analysis of Swollen and Tender Joints ACR-28 for the sum of tenderness (left and right) and the sum of swelling (left and right) will be performed separately using an analysis of covariance (ANCOVA) model with change from baseline to each nominal visit as the dependent variable, treatment group as the factor, and the baseline ACR-28 score as a covariate. The derived sum for tenderness and swelling is calculated and provided in the clinical database. A missing total sum of swelling or sum of tenderness will not be imputed. Observed results only will be presented for the change from baseline analysis. The p-value for the difference between the BOS161721 and placebo groups at each nominal visit is based on the F-test.

Descriptive statistics for the tenderness and swelling derived sums will be summarized separately at baseline and each visit. Change from baseline will be summarized at each post-baseline visit. The LS Mean estimate, its standard error and 90% CI will be tabulated by treatment group and for the difference between treatment groups (BOS161721 versus placebo) for change at each visit.

The mean and standard deviation of the change from baseline in ACR-28 sum of tenderness and sum of swelling will be presented graphically over time by study phase and treatment group.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

Swollen and Tender Joints ACR-28 observed results will be presented in data listings for all randomized patients.

7.2.10. Results and Changes from Baseline in SLEDAI-2K

The SLEDAI-2K is a validated instrument that measures disease activity in SLE patients at the time of the visit and in the previous 30 days. It is a global index and includes 24 clinical and laboratory variables that are weighted by the type of manifestation, but not by severity. The total score falls between 0 and 105, with higher scores representing increased disease activity. The SLEDAI-2K has been shown to be a valid and reliable disease activity measure in multiple patient groups. A SLEDAI-2K of 6 or more generally represents moderately to severely active disease.

The analysis of SLEDAI-2K will be performed using an analysis of covariance (ANCOVA) model with change from baseline to each nominal visit as the dependent variable, treatment group as the factor, and the baseline SLEDAI-2K score as a covariate. The SLEDAI-2K total score is calculated and provided in the clinical database. A missing total SLEDAI-2K score will not be imputed. Observed results only will be presented for the change from baseline analysis.

The p-value for the difference between the BOS161721 and placebo groups at each nominal visit is based on the F-test.

Descriptive statistics for the SLEDAI-2K total score will be summarized separately at baseline and each visit. Change from baseline will be summarized at each post-baseline visit. The LS Mean estimate, its standard error and 90% CI will be tabulated by treatment group and for the difference between treatment groups (BOS161721 versus placebo) for change at each visit.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

SLEDAI-2K observed results will be presented in a data listing for all randomized patients.

7.2.11. Results and Changes from Baseline in SLICC/ACR Damage Index

The SLICC/ACR damage index is a validated instrument to assess damage, defined as irreversible impairment, continuously persistent for 6 months (ascertained by clinical assessment), occurring since the onset of lupus, and it is based on a weighted scoring system. This index records damage occurring in patients with SLE regardless of cause, with demonstrated content, face, criterion, and discriminant validity. It will be performed on Days 0 and 180.

The analysis of SLICC/ACR Damage Index total score will be performed using an analysis of covariance (ANCOVA) model with change from baseline to each nominal visit as the dependent variable, treatment group as the factor, and the baseline SLICC/ACR Damage Index total score as a covariate. The SLICC/ACR Damage Index total score is calculated and provided in the clinical database. A missing total SLICC/ACR Damage Index score will not be imputed. Observed results only will be presented for the change from baseline analysis. The p-value for the difference between the BOS161721 and placebo groups at each nominal visit is based on the F-test.

Descriptive statistics for the SLICC/ACR Damage Index total score will be summarized separately at baseline and each visit. Change from baseline will be summarized at each post-baseline visit. The LS Mean estimate, its standard error and 90% CI will be tabulated by treatment group and for the difference between treatment groups (BOS161721 versus placebo) for change at each visit.

These analyses will also be performed for the MAD Phase 1b with the exception of the statistical testing.

Observed SLICC/ACR Damage Index data will be presented in a by-patient listing for all randomized patients. SLICC Criteria for SLE at screening will also be presented in a by-patient listing for all randomized patients.

7.2.12. Medication Failure

7.2.12.1. Proportion of patients with medication failures

Patients who received prohibited medications or unallowable CS usage as described in Section 4.6 of the study protocol will be considered "medication failures" and will be treated as non-responders for the primary efficacy analysis and selected sensitivity and secondary efficacy analyses. Patients who received an allowable CS burst but having missing data at Day 210 will considered a medication failure. The determination of medication failures will be reviewed using blinded data and finalized prior to unblinding.

The proportion of patients with medication failures will be summarized by study phase and treatment group in the FAS overall and by visit. A patient will be counted as a medication failure at each visit occurring on or after the date in which a patient is considered a medication failure. Statistical methods described in Section 7.1.1 will be implemented for observed data only. LOCF and missing data imputation will not be performed.

Statistical testing data will be omitted from MAD Phase 1b treatment group displays.

7.2.12.2. Time to Medication Failure (TTMF)

Medication failure is defined in Section 4.6 of the study protocol and discussed in Section 7.1.1. Time to medication failure will be computed as the event date (as described below) – randomization date + 1.

The date of medication failure will be reviewed by Boston Pharmaceuticals and agreed upon prior to database lock and unblinding.

Kaplan-Meier methods will be used to estimate TTMF for each treatment group. Estimates of median TTMF will be provided along with 90% confidence intervals. Q1 and Q3 will also be tabulated. Comparisons between treatment groups will use the one-sided log-rank test. In addition, hazard ratio (HR) estimates with 90% CIs will be calculated from a Cox proportional hazards regression analysis that includes treatment as a factor. Censoring rules are defined as follows.

Situation	Date of Event or Censoring	Outcome
Patient is designated a medication failure (received prohibited meds or unallowable CS burst)	Date of first dose of medication leading to medication failure designation. These dates will be obtained from the date of deviation on the protocol deviation CRF form.	Event
Patient withdraws from study (no medication failure)	Censored at date of withdrawal	Censored
Patient completes the study (no medication failure)	Censored at last study visit	Censored
Patient dies during study (no medication failure)	Censored at the date of death	Censored

Kaplan-Meier curves of TTMF will be plotted over time.

These analyses will also be performed for the MAD Phase 1b with the exception of the statistical testing.

Time to medication failure data will be provided in a by-patient listing in the full analysis set.

7.2.13. Duration of Longest SRI-4 Response

Duration of longest SRI-4 response will be computed for patients who have been identified as a responder at least once based on criteria defined in Section 7.1.1. The duration of longest response will be defined as the longest period a patient meets the SRI-4 responder criteria at consecutive visits. This will be computed as the date of last consecutive SRI-4 response – date of first consecutive response + 1.

The duration of longest SRI-4 response will be analyzed using an analysis of variance (ANOVA) model with treatment group as the effect. The p-value for the difference between the BOS161721 and placebo group is based on the F-test.

The mean, standard deviation, median, minimum, maximum, Q3, Q4, LS Mean estimate, its standard error and 90% CI will be tabulated by treatment group and for the difference between treatment groups (BOS161721 versus placebo).

If a patient has missed a visit between two scheduled visits with a response at both, the patient will be counted as having a response at the missed visit.

Duration of longest SRI-4 response will be displayed in the SRI response listing.

These analyses will also be performed for the MAD Phase 1b with the exception of the statistical testing.

7.2.14. Time to First SRI-4 Response

SRI-4 response is defined in Section 7.1.1. Time to first SRI-4 response will be computed as the event date (as described below) – randomization date + 1.

Time to first SRI-4 response will be analyzed with Kaplan-Meier methods described for TTMF and censoring as described below. Kaplan-Meier curves will be plotted over time.

Situation	**Date of Event or Censoring**	**Outcome**
Patient has SRI-4 Response (prior to medication failure, withdrawal, or death)	First date where a patient has been documented to have achieved a SRI-4 responder designation	Event
Patient is designated a medication failure (received prohibited meds or unallowable CS burst) prior to SRI-4 response/non-responder	Date of first dose of medication leading to medication failure designation. These dates will be obtained from the date of deviation on the protocol deviation CRF form.	Censored
Patient withdraws from study (no SRI-4 response, no medication failure)	Censored at date of last visit with non-missing SRI-4 assessments (BILAG, SLEDAI-2K, and PGA) prior to withdrawal	Censored
Patient completes the study (no SRI-4 Response, no medication failure)	Censored at date of last visit with non-missing SRI-4 assessments (BILAG, SLEDAI-2K, and PGA)	Censored
Patient dies during study (no SRI-4 response, no medication failure)	Censored at the date of death	Censored

These analyses will also be performed for the MAD Phase 1b with the exception of the statistical testing.

Time to first SRI-4 response data will be provided in a by-patient listing in the full analysis set.

7.2.15. Time to First BILAG A Flare or >1 BILAG B Flare Compared to Baseline through Day 210

Time to first BILAG A Flare or >1 BILAG B Flare will be will be computed as the event date (as described below) – randomization date + 1. If a patient is designated as a medication failure prior to or on their day 210 visit, they will be considered as having a BILAG flare (event) as of the date of first dose of the medication leading to the medication failure designation due to prohibited medication or unallowable CS burst.

Time to first BILAG A Flare or >1 BILAG B Flare will be analyzed with Kaplan-Meier methods described for TTMF and censoring as described below. Kaplan-Meier curves will be plotted over time.

Situation	**Date of Event or Censoring**	**Outcome**
Patient has a BILAG A Flare or >1 BILAG B Flare compared to baseline (Noted as BILAG 1A/2B Flare) prior to or on their day 210 visit	Date of first BILAG 1A/2B Flare	Event
Patient is designated a medication failure (received prohibited meds or unallowable CS burst) prior to or on their day 210 visit	Date of first dose of medication leading to medication failure designation. These dates will be obtained from the date of deviation on the protocol deviation CRF form.	Censored
Patient withdraws from study (no BILAG 1A/2B Flare, no medication failure) prior to or on their day 210 visit	Censored at date of last non-missing BILAG assessment (prior to withdrawal) through day 210	Censored
Patient completes the study (no BILAG 1A/2B Flare, no medication failure)	Censored at date of last non-missing BILAG assessment through day 210	Censored
Patient dies during study (no BILAG 1A/2B Flare, no medication failure) prior to or on their day 210 visit	Censored at the day of death	Censored

These analyses will also be performed for the MAD Phase 1b with the exception of the statistical testing.

Time to first BILAG A Flare or >1 BILAG B Flare data will be provided in a by-patient listing in the full analysis set.

8. PHARMACOKINETICS/PHARMACODYNAMICS

Pharmacokinetics and pharmacodynamics are secondary endpoints for the Phase 1b MAD study and exploratory endpoints for the Phase 2 POC study.

PK concentrations and PD results reported as less than the lower limit of quantification (LLOQ) will be counted as 0 to calculate summary linear statistics, 0.5*LLOQ for geometric statistics, and will be reported as <LLOQ in listings.

PK blood sampling as collected on the CRF will be provided in by-patient data listings. All observed PK concentrations and PD data will also be presented in by-patient listings, including CCI , nAb, and ADA. PK parameters, as provided by Human Predictions, LLC, will be displayed in by-patient data listings.

Additional details surrounding the PK and PD analysis, modeling, and evidence of PK/PD relationships will be provided in a separate CPAP.

8.1. Pharmacokinetics

BOS161721 PK concentrations will be summarized and listed in the PK set by study phase, treatment group, visit, and timepoint. Summary statistics will include the geometric mean, geometric standard deviation, geometric % coefficient of variation, n, mean, standard deviation, median, minimum, and maximum. BOS161721 concentrations will be plotted over time by treatment group and visit.

PK parameters for the PK analysis set will be calculated by Human Predictions, LLC and will be provided to Array Biostatistics to prepare outputs for the MAD Phase 1b. Details surrounding the PK parameter calculations will be provided in a separate CPAP. Summary statistics will be presented for C_{max}, T_{max}, AUC, $t_{½}$, CL, and V_d and will include the same statistics as PK concentrations except Tmax which will include n, median, minimum, and maximum, only.

8.2. Pharmacodynamics

Descriptive statistics for the observed results of the following PD endpoints/results will be summarized separately at baseline and each visit. Change from baseline will be summarized at each post-baseline visit. In cases where both numeric and qualitative results are provided, the summaries will be based on the numeric results. Graphical presentations of mean and standard deviation of the change from baseline over time will also be presented for the below parameters.

PD Endpoint	**Applicable Study Portion(s)**	**Parameters to be Included in Summary Tables and Figures ('NA (all)' is displayed if all collected parameters are to be included)**
Phosphorylated signal transducer and activator of transcription 3 (pSTAT3)	MAD only	% pSTAT3+ Lymphocytes –Stimulated; Ratio MEF (Stimulated / Unstimulated)
CCI [redacted]	MAD and POC	NA (all)
Leukocyte immunophenotype (B and NK Cells)	MAD and POC	CD19+% of CD45+ Lymphocytes; IgD+C27-% of CD19+; IgD+CD27-CD38++CD24++% of CD19+; IgD-CD27-% of CD19+; IgD+CD27+% of CD19+; IgD-CD27+% of CD19+; IgD-CD27+CD38++CD138-% of CD19+; IgD-CD27+CD38++CD138+% of CD19+; CD56+% OF CD45+ Lymphocytes
Leukocyte imunophenotype (T-Cells)	MAD and POC	CD4+CD8-% of CD3+; CXCR5+% of CD4+CD8-; PD-1+% of CD4+CD8-; ICOS+% of CD4+CD8-; CD25+CD127-% of CD4+CD8-; CXCR5+% of CD25+CD127-; PD-1+% of CD25+CD127-; ICOS+% of CD25+CD127-; CD8+CD4-% of CD3+
CCI [redacted]	MAD and POC	CCI [redacted]
CCI [redacted]	MAD and POC	NA (all)

PD Endpoint	Applicable Study Portion(s)	Parameters to be Included in Summary Tables and Figures ('NA (all)' is displayed if all collected parameters are to be included)
CCI	MAD and POC	NA (all)
CCI	MAD and POC	NA (all)
CCI	MAD and POC	NA (all)
CCI	MAD and POC	All parameters for 'average target CP' values

CCI

Additional immunogenicity and PD results may be plotted over time by treatment group and visit.

9. QUALITY OF LIFE

9.1. CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

9.2. CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

CCI

10. SAFETY

All safety analysis reporting will be based on the Safety Analysis Set.

10.1. Adverse Events

Adverse events (AEs) will be collected and recorded for each patient from the date of the first dose of study drug until the end of their participation in the study, including the safety follow up period. Severity of adverse events will be graded according to National Cancer Institute

Common Terminology Criteria for Adverse Events (NCI CTCAE) version 4.03. AEs will also be assessed for relationship to study drug, and seriousness. Adverse events will be marked as adverse events of special interest (AESIs), dose-limiting toxicities (DLTs), or injection site reactions (ISRs), where applicable. AEs will be considered treatment-emergent adverse events (TEAEs) if they started any time after first dose or if the start date was missing. Any missing severity assessments will be assumed to be Grade 3, missing relationship assessments will be assumed to be related, and missing seriousness assessments will be assumed as serious.

An overview of TEAEs will be produced, including counts and percentages of patients with any incidences of TEAE, TEAEs related to study treatment, Serious adverse events (SAEs), SAEs related to study treatment, CTCAE grade 2 or higher TEAEs, CTCAE grade 2 or higher TEAEs related to study treatment, CTCAE grade 3 TEAEs, CTCAE grade 3 or higher TEAEs related to study treatment, CTCAE grade 4 TEAEs, CTCAE grade 4 or higher TEAEs related to study treatment, TEAEs leading to study treatment discontinuation, Related TEAEs leading to study treatment discontinuation, TEAEs of special interest, Dose-limiting toxicities, TEAEs resulting in death, and injection site reactions.

Adverse events will be coded based on the Medical Dictionary for Regulatory Affairs (MedDRA) for reporting by system organ class (SOC) and preferred term (PT) in descending order of overall incidence. The MedDRA version used is defined in the study data management plan.

Summaries of adverse events by SOC and PT will include the following types:

- TEAEs;
- TEAEs related to study treatment;
- SAEs;
- CTCAE Grade 3 or higher TEAEs related to study treatment;
- TEAEs leading to treatment discontinuation; and
- Dose-Limiting Toxicities
- TEAEs of Special Interest

A summary of TEAEs by SOC, PT, and maximum CTCAE grade will also be prepared. For these summaries, TEAEs will be sorted for each patient by SOC/PT and CTCAE grade; patients will be counted once within a SOC/PT based on their maximum CTCAE grade. Missing CTCAE grade will be counted in a missing row.

A summary of TEAEs by PT and maximum CTCAE grade will also be prepared with sections for any TEAE and related TEAEs. For these summaries, patients will be counted once within a PT based on their maximum CTCAE grade. Missing CTCAE grade will be counted as grade 3. Preferred terms will be included if at least one patient has a CTCAE grade 3 or higher for the PT. PTs will also be included if 10% or more of patients experience a CTCAE grade 1 or 2 TEAE. The summary will be sorted by the total incidence of the PT in the any TEAE section.

A comprehensive listing of all AEs will be provided in a by-patient data listing for the safety analysis set. In addition, the following listings will be provided for the safety analysis set:

- TEAEs related to study treatment;
- SAEs;
- TEAEs leading to treatment discontinuation; and
- Adverse Events of Special Interest

10.2. Clinical Laboratory Evaluations

Laboratory tests will be performed at times defined in the protocol Schedule of Assessments. Clinical chemistry and hematology parameters will be reported based on the International System of Units (SI). The following laboratory evaluations will be reported in data summaries, with asterisks indicating those that will be graded using CTCAE. Grading will be performed by the Central Lab based on the details provided in Appendix 13.2 for relevant laboratory parameters. If a grade is not present in the central lab database for a graded lab parameter, the grade will be analyzed as grade 0.

Hematology: Hemoglobin*, Hematocrit, RBC count, RDW, MCV, MCH, MCHC, Platelet count*, WBC count*, CD4+ count*, Total neutrophils (Abs)*, Eosinophils (Abs), Monocytes (Abs), Basophils (Abs), Lymphocytes (Abs)*

Clinical chemistry: Alanine Aminotransferase*, Albumin*, Alkaline Phosphatase*, Aspartate Aminotransferase*, Bicarbonate, Bilirubin*, Blood urea nitrogen, C Reactive Protein, Calcium, Chloride, Creatine Kinase*, Creatinine*, Glucose (fasting)*, Glucose (random)*, Potassium*, Protein, Sodium*, Uric acid*, eGFR (Cockcroft-Gault)*, eGFR (MDRD)*, Total cholesterol (fasting)*, LDL-C (fasting), HDL-C (fasting), Triglycerides (fasting)*, Gamma Glutamyl Transferase*.

Coagulation: Prothrombin Intl. Normalized Ratio*, Activated Partial Thromboplastin Time*

Urinalysis: pH, Glucose (qual), Protein (qual)*, Blood (qual)*, Ketones, Nitrites, Leukocyte esterase, Urobilinogen, Urine bilirubin, Microscopy

Observed values and changes from baseline for hematology, clinical chemistry, and coagulation laboratory evaluations will be summarized at each visit and most extreme change. Endpoint analyses will be based on changes from baseline assessments at Days 120 and 210.

The number and percent of patients with a NCI-CTCAE toxicity grade of 3 or higher will be tabulated by laboratory evaluation with defined NCI-CTCAE grading at each visit for hematology, clinical chemistry, and coagulation. Patients with missing values post-baseline will be excluded from the summary. The denominator for percentages will be the number of patients with at least one post-baseline assessment for the laboratory parameter in question.

Hematology, clinical chemistry, and coagulation shift tables displaying the shift from baseline to the worst value of NCI-CTCAE grade will be presented based on the most extreme change as it relates to the relevant NCI-CTCAE definition. NCI-CTCAE relating to “high/hyper” conditions will depend on the maximum post-baseline value while NCI-CTCAE “low/hypo” will be

reported based on the minimum post-baseline value. Separate shift tables will be prepared for parameters with bi-directional toxicity grading. Lab parameters which are optional, such as gamma glutamyl transferase, will not be included in the shift tables.

Hematology and clinical chemistry laboratory data which is not graded will also be summarized in shift tables of baseline to each visit and most extreme change based on range categories of low (below lower limit of normal [LLN], normal, and high [above upper limit of normal [ULN]]).

Selected hematology and clinical chemistry laboratory results will be plotted over time for treatment group for MAD phase 1b portion of the study in the safety analysis set. The mean and standard deviation of the change from baseline for selected hematology and clinical chemistry laboratory results will be plotted over time for the POC Phase 2 portion of the study in the safety analysis set.

All laboratory parameters as well as a summary of central laboratory tests and chest x-ray will be provided in patient data listings for all randomized patients. By patient listings of clinical chemistry and hematology data will be presented in the safety analysis set for the DMC and interim analysis only.

10.3. Other Safety Evaluations

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

The C-SSRS is a low-burden measure of the spectrum of suicidal ideation and behavior that was developed by Columbia University researchers for the National Institute of Mental Health Treatment of Adolescent Suicide Attempters Study to assess severity and track suicidal events through any treatment. It is a clinical interview providing a summary of both ideation and behavior that can be administered during any evaluation or risk assessment to identify the level and type of suicidality present. The C-SSRS can also be used during treatment to monitor for clinical worsening.[i] The C-SSRS evaluation will be performed as specified in the protocol Schedule of Assessments.

The C-SSRS has 11 binary (yes/no) outcomes corresponding to five categories of suicidal ideation, five categories of suicidal behavior, and the category of self-injurious behavior without suicidal intent:

Suicidal ideation is present if there is a “yes” response to any of the 5 suicidal ideation category questions. Suicidal behavior is present if there is a “yes” response to any of the 5 suicidal behavior category questions. Suicidal ideation or behavior is present if there is a “yes” response to any of the 10 suicidal or ideation or behavior category questions.

The number and percentage of patients with any post-baseline occurrence of the following ideations and behaviors will be presented by study phase and treatment group:

- Suicidal ideation (overall and by individual question),
- Suicidal behavior (overall and by individual question),

- Suicidal ideation or behavior,

A similar tabulation of the occurrence of any of the above ideations and behaviors at baseline or as part of their lifetime history/past 12-month history (collected at the screening visit) will be presented by study phase and treatment group.

A shift table displaying the shift from worst pre-treatment category to the worst post-baseline category will be presented by study phase and treatment group. Best to worst C-SSRS category is defined in the following order: no suicidal ideation or behavior, suicidal ideation, and suicidal behavior. The worst post-baseline shift is calculated based on all post-baseline visits, including unscheduled visits. Worst pre-treatment will include all C-SSRS assessments prior to first dose date/time (screening or baseline). The denominator will be the number of patients with at least one non-missing baseline and post-baseline C-SSRS assessment.

Missing data will not be imputed.

All C-SSRS individual items will be presented in data listings for screening and post-screening (since last visit), separately. A separate listing will be presented for patients with suicidal ideation, suicidal behavior, or non-suicidal self-injurious behavior for all randomized patients.

10.3.2. Vital Signs

Vital signs include: heart rate (beats/min); temperature (^{o}C); systolic and diastolic blood pressure (mmHg); and weight (kg). Observed values and changes from baseline for vital signs will be summarized at each visit and time point, as well as for most extreme change by study phase and treatment group.

All vital signs data will be presented in patient data listings for all randomized patients.

10.3.3. 12-Lead Electrocardiogram (ECG)

ECG will be assessed as specified in the protocol Schedule of Assessments. The following ECG parameters will be collected: PR interval (msec), QRS interval (msec), RR interval (msec), QT interval (msec), and QT interval corrected for heart rate using Fridericia's formula (QTcF) (msec). ECG parameters and ECG evaluation will come directly from the database and will not be calculated during analysis.

Observed values and changes from baseline for ECG parameters will be summarized at each visit, as well as for most extreme change by study phase and treatment group.

The number and percent of patients with the following parameters at any post-baseline visit and at each visit will be summarized:

- Abnormal, not clinically significant ECGs
- Abnormal, clinically significant ECGs

The number and percent of patients with each ECG evaluation result will be summarized by visit. The worst ECG evaluation at any post-baseline visit will also be summarized in the order of: Normal, Abnormal, not clinically significant, Abnormal, clinically significant.

For the summary of incidence at any post-baseline visit, the denominators for percentages will include patients with one or more non-missing value at any post-baseline visit. For the by-visit summaries, the denominators for percentages will include patients with a non-missing value at that visit.

Investigator reported ECG result shifts from baseline to each visit and worst case post-baseline will be summarized. Worst case post-baseline will be based on the most abnormal observed value on or after the randomization date.

All ECG data will be presented in a by patient data listing for all randomized patients. A separate by patient listing for patients with abnormal, clinically significant 12-lead ECGs will also be provided in the safety analysis set for DMC meetings and the interim analysis only.

10.3.4. Physical Examinations

Targeted physical examination data and full physical examination data will be presented in patient data listings. A separate by patient listing for patients with abnormal, clinically significant physical examination findings will also be provided in the safety analysis set for DMC meetings and the interim analysis only.

10.3.5. Anti-Drug Antibodies (ADAs)

The serum samples to measure the presence of ADA will be collected as specified in the protocol Schedule of Assessments. ADAs will be summarized and listed according to details provided in Section 8.

11. CHANGES TO THE PLANNED ANALYSIS

No changes have been made to the planned analysis.

12. REFERENCES

i Posner K, Melvin G a, Stanley B, Oquendo M a, Gould M. Factors in the assessment of suicidality in youth. *CNS Spectr*. 2007;12(2):156-162.

CCI [redacted]

International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH E3 Guideline: Structure and content of clinical study reports questions & answers (R1). 2012. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_QAs_R1_Step4.pdf. Accessed 13 Jan 2016.

International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonised Tripartite Guideline: Structure and content of clinical study reports E3. Step 4.
1995. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf. Accessed 13 Jan 2016.

International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH E9: Statistical principles for clinical trials.
1998. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf . Accessed 15 June 2016.

FDA Guidance for Industry Systemic Lupus Erythematosus – Developing Medical Products for Treatment.
2010. https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm072063.pdf. Accessed 23February2018.

Committee for Medicinal Products for Human use (CHMP). EMA Guideline on clinical investigation of medicinal products for the treatment of systemic lupus erythematosus and lupus nephritis.
2015. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/03/WC500184889.pdf. Accessed 04March2018.

CCI [redacted]

Nilsson M, Suryawanshi S, Gassmann-Mayer C, Dubrava S, McSorley P, Jiang K. Columbia-Suicide Severity Rating Scale Scoring and Data Analysis Guide. V2
2013. https://cssrs.columbia.edu/wp-content/uploads/ScoringandDataAnalysisGuide-for-Clinical-Trials-1.pdf. Accessed 18May2018.

CCI [redacted]

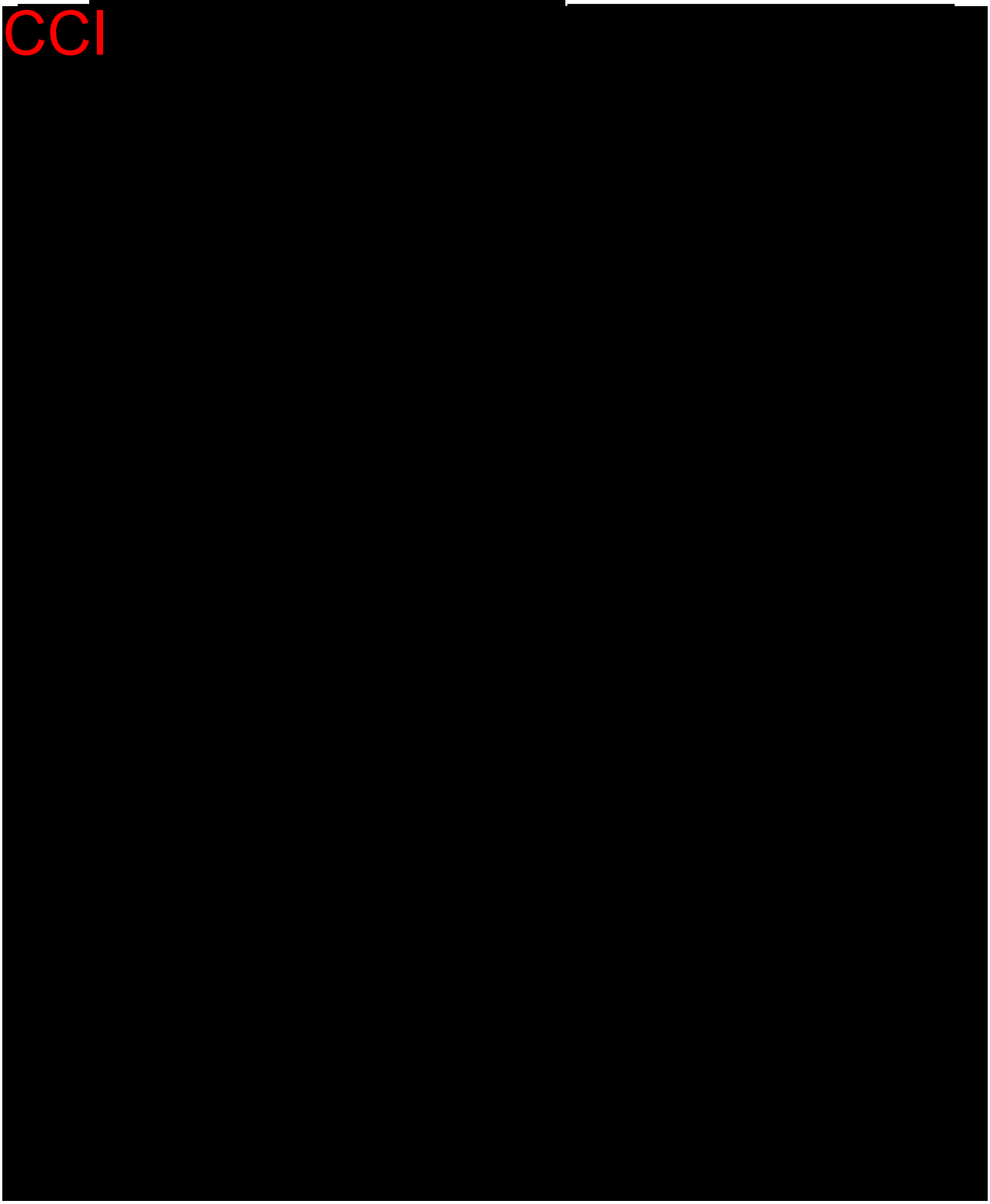
CCI [redacted]

CCI [redacted]

13. APPENDICES

13.1. CCI

CCI



CCI

CCI

CCI

CCI

13.2. CTCAE Grading Details from the Central Laboratory

CTCAE CHECKS

NCI Common Terminology Criteria (CTC) for Adverse Events (AE)
Laboratory Tests

The NCI CTC for AE is a descriptive terminology which can be utilized for AE reporting. A grading (severity) scale is provided for each AE term. (CTCAE, Version 4.03)

Grade I: Mild AE
Grade II: Moderate AE
Grade III: Severe AE
Grade IV: Life-threatening or disabling AE

LLN = Lower Limit of Normal range
ULN = Upper Limit of Normal range

NOTES: The ranges in this table are adjusted to the precision and units used by BARC. In case the grading ranges are in absolute values, the ranges are listed in the different units in which the parameter can be reported. In case the grading ranges are expressed as a factor of the lower and/or upper limit of the normal ranges, no units are shown. It is possible that in some cases, the lower or upper limit of normal range is at the same time a Grade I AE.
Please note CTC checks only apply on parameters taken within the scope of the same visit.
More information about NCI CTC can be accessed on the CTEP website "http://ctep.cancer.gov/reporting/ctc.html"

Parameter	Grade I	Grade II	Grade III	Grade IV	Units
		HEMATOLOGY			
CD4 count	500 - < LLN[4]	200 - < 500	50 - < 200	< 50	/mm³
Hemoglobin	10.0 - < LLN	8.0 - < 10.0	< 8.0	-	g/dL
	100 - < LLN	80 - < 100	< 80	-	g/L
	ULN + > 0.0-2.0	ULN + > 2.0-4.0	ULN + > 4.0	-	g/dL
	ULN + > 0-20	ULN + > 20-40	ULN + > 40	-	g/L
Leukocytes	3.0 - < LLN	2.0 - < 3.0	1.0 - < 2.0	< 1.0	10^3/µL or 10^9/L
	-	-	> 100.0	-	10^3/µL or 10^9/L
Lymphocytes	0.80 - < LLN	0.50 - < 0.80	0.20 - < 0.50	< 0.20	10^3/µL or 10^9/L
	-	> 4.00 – 20.00	> 20.00	-	10^3/µL or 10^9/L
Neutrophils	1.50 - < LLN	1.00 - < 1.50	0.50 - < 1.00	< 0.50	10^3/µL or 10^9/L
Platelets	75 - < LLN	50 - < 75	25 - < 50	< 25	10^3/µL or 10^9/L
		COAGULATION			
Fibrinogen	0.75xLLN - < LLN	0.5xLLN - < 0.75xLLN	0.25xLLN - < 0.5xLLN	< 0.25xLLN	
PT (INR)	> ULN - 1.5xULN	> 1.5xULN – 2.5xULN	> 2.5xULN	-	
APTT	> ULN - 1.5xULN	> 1.5xULN – 2.5xULN	> 2.5xULN	-	
		BIOCHEMISTRY			
Albumin	3.0 - < LLN	2.0 - < 3.0	< 2.0	-	g/dL
	30 - < LLN	20 - < 30	< 20	-	g/L
Alkaline phosphatase	> ULN - 2.5xULN	> 2.5xULN - 5xULN	> 5xULN - 20xULN	> 20xULN	
ALT	> ULN – 3xULN	> 3xULN - 5xULN	> 5xULN - 20xULN	> 20xULN	
Amylase	> ULN - 1.5xULN	> 1.5xULN - 2xULN	> 2xULN - 5xULN	> 5xULN	
AST	> ULN – 3xULN	> 3xULN - 5xULN	> 5xULN - 20xULN	> 20xULN	
Bilirubin, total	> ULN - 1.5xULN	> 1.5xULN - 3xULN	> 3xULN - 10xULN	> 10xULN	

Parameter	Grade I	Grade II	Grade III	Grade IV	Units
Corrected calcium	8.0 - < LLN	7.0 - < 8.0	6.0 - < 7.0	< 6.0	mg/dL
	2.00 - < LLN	1.75 - < 2.00	1.50 - < 1.75	< 1.50	mmol/L
	> ULN - 11.5	> 11.5 - 12.5	> 12.5 - 13.5	> 13.5	mg/dL
	> ULN - 2.90	> 2.90 - 3.10	> 3.10 - 3.40	> 3.40	mmol/L
Cholesterol, total	> ULN - 300	> 300 - 400	> 400 - 500	> 500	mg/dL
	> ULN - 7.75	> 7.75 - 10.34	> 10.34 - 12.92	> 12.92	mmol/L
Creatine kinase	> ULN - 2.5xULN	> 2.5xULN - 5xULN	> 5xULN - 10xULN	> 10xULN	
Creatinine (serum)	> ULN - 1.5xULN	> 1.5xULN - 3xULN	> 3xULN - 6xULN	> 6xULN	
eGFR	60 - < LLN[4]	30 - 59	15 - 29	< 15	mL/min (/1.73m^2)
GGT	> ULN - 2.5xULN	> 2.5xULN - 5xULN	> 5xULN - 20xULN	> 20xULN	
Glucose (2h PP; random)	55 - < LLN[1]	40 - < 55	30 - < 40	< 30	mg/dL
	3.00 < LLN[1]	2.20- < 3.00	1.70- < 2.20	< 1.70	mmol/L
Glucose (fasting)	55 - < LLN	40 - < 55	30 - < 40	< 30	mg/dL
	3.00 < LLN	2.20- < 3.00	1.70- < 2.20	< 1.70	mmol/L
	> ULN - 160	> 160 - 250	> 250 - 500	> 500	mg/dL
	> ULN - 8.90	> 8.90 - 13.90	> 13.90 - 27.80	> 27.80	mmol/L
Haptoglobin	< LLN	-	-	-	
Lipase	> ULN - 1.5xULN	> 1.5xULN - 2xULN	> 2xULN - 5xULN	> 5xULN	
Magnesium	1.2 - < LLN	0.9 - < 1.2	0.7 - < 0.9	< 0.7	mg/dL
	0.50 - < LLN	0.40 - < 0.50	0.30 - < 0.40	< 0.30	mmol/L
	> ULN – 3.0	-	> 3.0 – 8.0	> 8.0	mg/dL
	> ULN - 1.23	-	> 1.23 - 3.30	> 3.30	mmol/L
Phosphorus	2.5 - < LLN	2.0 - < 2.5	1.0 - < 2.0	< 1.00	mg/dL
	0.80 - < LLN	0.60 - < 0.80	0.30 - < 0.60	< 0.30	mmol/L
Potassium	3.0 - < LLN[2]	3.0 - < LLN[2]	2.5 - < 3.0	< 2.5	mmol/L
	> ULN - 5.5	> 5.5 - 6.0	> 6.0 - 7.0	> 7.0	mmol/L
Sodium	130 - < LLN	-	120 - < 130	< 120	mmol/L
	> ULN - 150	> 150 - 155	> 155 - 160	> 160	mmol/L
Triglycerides	150 - 300	> 300 - 500	> 500 - 1000	> 1000	mg/dL
	1.71 - 3.42	> 3.42 - 5.70	> 5.70 - 11.40	> 11.40	mmol/L
Troponin I	> 0.01 - < 0.30 (♀)		≥ 0.30 (♀ & ♂)		ng/mL or µg/L
	> 0.03 - < 0.30 (♂)				ng/mL or µg/L
Troponin T	> 0.01 - <0.10	-	≥ 0.10	-	ng/mL or mg/L
Uric acid	> ULN - 10.0[3]	-	> ULN - 10.0[3]	> 10.0	mg/dL
	> ULN - 590[3]	-	> ULN - 590[3]	> 590	µmol/L
URINALYSIS					
Creatinine clearance	60 - < LLN	30 - 59	15 - 29	< 15	mL/min
Blood	+(++)	-	-	-	
Protein	+	++ or +++ (≥ 18y)	-	-	
	< 1000	1000 – 3499 (≥18y)	≥ 3500 (≥ 18y)	-	mg/24h
Protein/creatinine ratio (< 18y)		0.5 – 1.9	> 1.9		

[1] Glucose: BARC reference range has no LLN, therefore 74 mg/dL (4.11 mmol/L) is applied as LLN in Grade I flagging (reference Tietz Clinical guide to laboratory tests 4th ed.)
[2] Potassium: Grade II when intervention is indicated, else Grade I
[3] Uric acid: Grade III, when in combination with physiologic consequences, else Grade I
[4] CD4 count & eGFR: LLN is lower or equal to lower threshold for Grade I flagging, therefore only the exact lower threshold value is flagged as Grade I

13.3. Table of Contents for Tables, Listings, and Figures

The table of contents for tables, listings, and figures will be provided in a corresponding mock shell document.