230LE201 / NCT02847598

STATISTICAL ANALYSIS PLAN

STATISTICAL ANALYSIS PLAN

Product Studied: BIIB059

Protocol Number(s): 230LE201

A 2-Part Phase 2 Randomized, Double-Blind, Placebo-Controlled Study Evaluating the Efficacy and Safety of BIIB059 in Subjects with Systemic Lupus Erythematosus and Active Skin Manifestations and in Subjects with Active Cutaneous Lupus Erythematosus with or without Systemic Manifestations

Date of Protocol: 15 March 2019

SAP Version: 2.0

Date of Statistical Analysis Plan: 23 October 2019

Written By:

[redacted]
<Program Statistician: [redacted] PhD > 10/23/2019 Date

[redacted]
<SMT Statistician: [redacted] M.Sc. > 23-Oct-2019 Date

Approved By:

[redacted]
<Program Statistician: [redacted] PhD > 10/23/2019 Date

[redacted]
<SMT Physician, [redacted] MD> 23Oct2019 Date

STATISTICAL ANALYSIS PLAN

Product Studied: BIIB059

Protocol Number(s): 230LE201

A 2-Part Phase 2 Randomized, Double-Blind, Placebo-Controlled Study Evaluating the Efficacy and Safety of BIIB059 in Subjects with Systemic Lupus Erythematosus and Active Skin Manifestations and in Subjects with Active Cutaneous Lupus Erythematosus with or without Systemic Manifestations

Date of Protocol: 15 March 2019

SAP Version: 2.0

Date of Statistical Analysis Plan: 23 October 2019

Written By:

<Program Statistician: [redacted], PhD > Date

<SMT Statistician: [redacted], PhD > Date

Approved By:

<Program Statistician: [redacted], PhD > Date

<SMT Physician, [redacted], MD> Date

TABLE OF CONTENTS

2

List of Abbreviations

ACLE	Acute cutaneous lupus erythematosus
ACR	American College of Rheumatology
AE	Adverse event
AIC	Akaike information criterion
ALT	Alanine aminotransferase
ANA	Antinuclear antigen
ANCOVA	Analysis of covariance
AST	Aspartate aminotransferase
AUC	Area under the concentration-time curve
BDCA2	Blood dendritic cell antigen-2
BICLA	BILAG-2004 based Composite Lupus Assessment
BILAG	British Isles Lupus Activity Group
BLQ	Below the level of quantification
BOCF	Baseline observation carried forward
CCLE	Chronic cutaneous lupus erythematosus
CI	Confidence interval
CL	Clearance
CLASI	Cutaneous Lupus Erythematosus Disease Area and Severity Index
CLASI-A	CLASI–Activity
[redacted]	[redacted]
CLE	Cutaneous lupus erythematosus
Cmax	Maximum observed concentration
Cmin	Minimum observed concentration
CPK	Creatine phosphokinase
CRO	Contract research organization
CRP	C-reactive protein
CSR	Clinical study report
DHA	Directions for Handling and Administration
DLE	Discoid lupus erythematosus

dsDNA	double-stranded DNA
ECG	Electrocardiogram
eCRF	Electronic case report form
ESR	Erythrocyte sedimentation rate
ET	Early termination
FU	Follow-up visit
GCP	Good Clinical Practice
HBV	Hepatitis B virus
HCV	Hepatitis C virus
[redacted]	[redacted]
IA	Interim analysis
IC90	Concentration resulting in 90% inhibition of response
ICF	Informed consent form
ICH	International Council for Harmonization
IFN	Interferon
IFN-α	Interferon-alpha
IgG1	Immunoglobulin, class G, type 1
IHC	Immunohistochemistry
IRT	Interactive response technology
IM	Intramuscular
IP	Investigational product
ITT	Intent-to-treat
IV	Intravenous
LE	Lupus erythematosus
LOCF	Last observation carried forward
LS	Least squares
LTBI	Latent tuberculosis infection
[redacted]	[redacted]
MAD	Multiple ascending dose
MAR	Missing at random
MCMC	Markov chain Monte Carlo

MCP	Metacarpophalangeal
MCP-Mod	Multiple comparison procedure - modelling
[redacted]	[redacted]
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple imputation
MITT	Modified intent-to-treat population
MMF	Mycophenolate mofetil
MMRM	Mixed effect model repeat measurement
MNAR	Missing not at random
MPS	Mycophenolate sodium
mRNA	Messenger RNA
NRI	Non-responder imputation
[redacted]	[redacted]
NSAID	Nonsteroidal anti-inflammatory drug
PCS	Physical component summary
PD	Pharmacodynamic
pDC	Plasmacytoid dendritic cell
PGA	Physician's global assessment
[redacted]	[redacted]
PIP	Proximal interphalangeal
PK	Pharmacokinetics
PP	Per-protocol
PV1	Protocol version 1
PV2	Protocol version 2
OCS	Oral corticosteroid
Q2W	Every 2 weeks
Q4W	Every 4 weeks
RNA	Ribonucleic acid
PIP	Proximal interphalangeal joints
PT	Preferred term
PV	Protocol version

Q1	First quartile
Q3	Third quartile
RBC	Red blood cell
SAE	Serious adverse event
SAP	Statistical analysis plan
SC	Subcutaneous
SCLE	Subacute cutaneous lupus erythematosus
SD	Standard deviation
SELENA	Safety of Estrogen in Lupus: National Assessment
[redacted]	[redacted]
SLE	Systemic lupus erythematosus
SLEDAI-2K	Systemic Lupus Erythematosus Disease Activity Index 2000
SoC	Standard of care
SOC	System organ class
SRI	Systemic Lupus Erythematosus Responder Index
SRI-4	Systemic Lupus Erythematosus Responder Index Response of ≥4
SUSAR	Suspected unexpected serious adverse reaction
t1/2	Elimination half-life
TB	Tuberculosis
TD	Treatment day
TLG	Tables, listings and graphs
TLR	Toll-like receptor
TP	Treatment period
USA	United States of America
VAS	Visual analog scale
Vd	Volume of distribution
WHO	World Health Organization
WK	Week

1. STUDY OBJECTIVES AND ENDPOINTS

1.1. Primary Objective and Endpoint

Primary Objective

- To evaluate the efficacy of BIIB059 in reducing disease activity in SLE subjects with active skin manifestations and joint involvement (Part A), and in subjects with active CLE (subacute CLE [SCLE] or chronic CLE [CCLE], including discoid lupus erythematosus [DLE]) with or without systemic manifestations (Part B).

Primary Endpoints

- Part A: Change in active joint count (28-joint assessment) from baseline to Week 24. The active joint count is defined as the sum of the tender joint count and the swollen joint count.
- Part B: Percent change in Cutaneous Lupus Erythematosus Disease Area and Severity Index-Activity (CLASI-A) score from baseline to Week 16.

1.2. Secondary Objectives and Endpoints

Secondary Objective

- To evaluate additional efficacy parameters of BIIB059 in reducing SLE/CLE disease activity
- To evaluate the safety and tolerability of BIIB059
- To evaluate the pharmacokinetic (PK) of BIIB059

Secondary Endpoints

<u>Efficacy endpoints:</u>

- Proportion of subjects with a CLASI-50 response, defined as a 50% improvement from baseline in CLASI-A score, at Week 24 (Part A) and at Weeks 12 and 16 (Part B).
- Percent change in CLASI-A score from Baseline to Week 12, Week 16, and Week 24 (Part A) and Week 12 (Part B).

- Proportion of subjects with a ≥4-point reduction in CLASI-A score at Week 24 (Part A) and at Weeks 12 and 16 (Part B) compared with baseline.
- Proportion of subjects with a ≥7-point reduction in CLASI-A score at Week 24 (Part A) and at Weeks 12 and 16 (Part B) compared with baseline.
- Proportion of subjects achieving a SLE Responder Index (SRI) of ≥4 (SRI-4) at Week 24 (see Section 6.3.5 for a detailed description of this endpoint) (Part A only).
- Change from baseline to Week 24 in Systemic lupus erythematosus disease activity index 2000 (SLEDAI-2K) score (Part A only).
- Proportion of subjects with no new organ system affected, as defined by no new BILAG-2004 grade A and no more than one new BILAG-2004 grade B, from baseline at Week 24 (Part A only).
- Change from baseline to Week 24 in PGA of SLE (VAS) score (Part A only).

Pharmacokinetics Endpoints (Part A and Part B):

- PK parameters, including, but not limited to, clearance (Cl), volume of distribution (Vd) and absorption rate (ka).

Safety and Tolerability Endpoints (Part A and Part B):

- Incidence, nature, severity, and relationship to study treatment of adverse events (AEs) and serious adverse events (SAEs)
- Change/shift from baseline in standard laboratory parameters, vital signs and electrocardiogram (ECG) results
- Incidence of anti-BIIB059 antibodies in serum
- Absolute and percent change from baseline over time in immunoglobulin levels (IgA, IgG and IgM) and immunoglobin titers for tetanus, diphtheria, and Pneumococcus vaccines

[redacted]

- [redacted]
- [redacted]

- [redacted]
- [redacted]
- [redacted]

[redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted].
- [redacted]
- [redacted]
- [redacted]

[redacted]

- [redacted]
- [redacted]

[redacted]

- [redacted]
- [redacted]
  - [redacted]
  - [redacted].
  - [redacted]
  - [redacted]
  - [redacted]

- [redacted]

2. STUDY DESIGN

2.1. Study Overview

This is a 2-part randomized, double-blind, parallel-group, placebo-controlled, multicenter Phase 2 trial. The protocol version 1 (PV1) was designed to evaluate the efficacy and safety of BIIB059 in SLE subjects with active skin manifestations (Part A) or active CLE, defined as discoid or subacute subtypes, with or without SLE (Part B). The study was to be conducted at approximately 95 centers in the United States, Europe, Latin America and Asia. In PV1 Part A, 100 subjects were planned to be randomly assigned to 1 of 4 treatment groups in a 1:1:1:1 ratio and receive either fixed doses of BIIB059 (50, 150 or 450 mg) or placebo subcutaneously Q4W for 24 weeks with a loading dose at Week 2, for a total of 7 doses. Randomization was stratified by baseline CLASI-A score (≤ 10, >10) and prior oral corticosteroid usage (≤ 10 mg/d, >10 mg/d). In PV1 Part B, 30 subjects were planned to be randomly assigned to 1 of 2 treatment groups in a 2:1 ratio (20 subjects to BIIB059, 10 subjects to placebo) and receive either a fixed dose of BIIB059 (450 mg) or placebo SC Q4W for 12 weeks with a loading dose at Week 2, for a total of 4 doses. Randomization was stratified by CLASI-A score (≤ 10, >10) and SLE (presence, absence).

2.1.1. Protocol changes

The protocol was first amended on 16 May 2017. Protocol version 2 (PV2) was designed to evaluate the efficacy and safety of BIIB059 in subjects with SLE and active skin manifestations and joint involvement (Part A) or active CLE, defined as SCLE or CCLE, including DLE, with or without SLE (Part B). As part of the amendment, the number of sites was expanded to approximately 130 centers across United States, Europe, Latin America, and Asia. Two additional amendments were done after PV2 but were administrative in nature and didn't impact the study design and statistical analyses plan.

The key changes between PV1 and PV2 (and subsequent versions) are listed below:

<u>Part A</u>

- In PV2 (and subsequent versions), oral corticosteroid tapering in Part A is now mandatory. Under PV1, it was encouraged but was not mandatory.

- In PV2 (and subsequent versions), the subjects will be randomized to fixed dose of BIIB059 450 mg or placebo. In PV1, subjects were planned to be randomized to either fixed doses of BIIB059 (50, 150 or 450 mg) or placebo.

- In PV2 (and subsequent versions), change in inclusion criteria included: removal of the CLASI-A score requirement of 8 or higher at entry and addition of joint count, requiring at least 4 swollen and 4 tender joints based on assessment of 28 joints (see Section 4.3.3 of the protocol), with at least 4 occurring in the PIP, MCP, or wrist joints. Note: A joint that is both tender and swollen will be included in both categories (counts as 1 tender and 1 swollen).

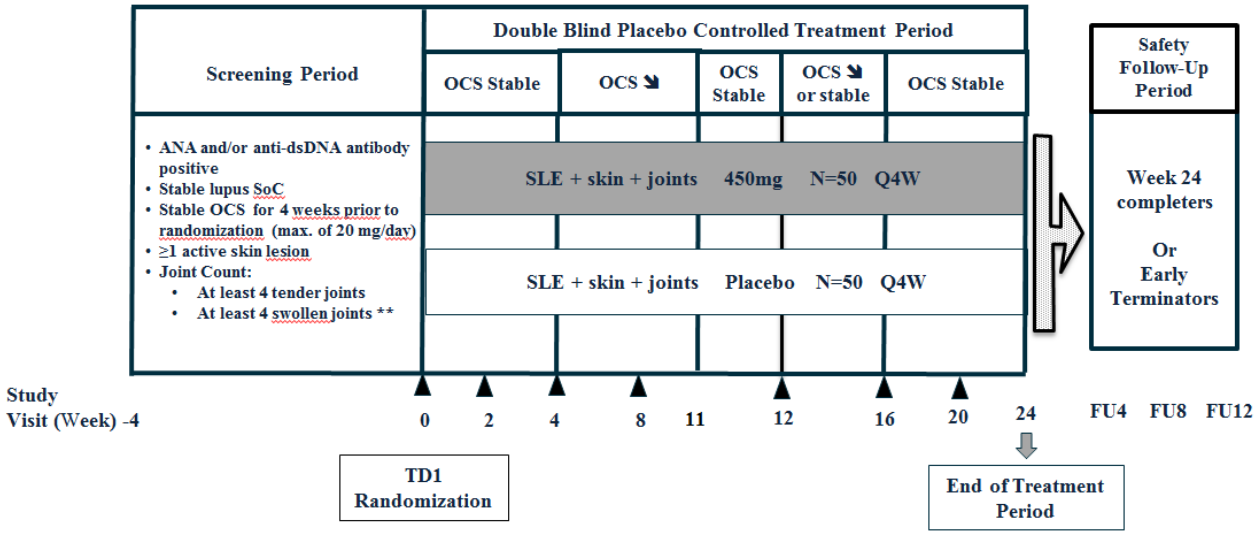
Part B

- In PV1, last dose was originally planned at week 8, with end of treatment period at week 12. Under PV2 (and subsequent versions), an additional week 12 dose was added (with end of treatment period now moved to week 16).

- In PV2 (and subsequent versions), the subjects will be randomized to either fixed doses of BIIB059 (50, 150 or 450 mg) or placebo. In PV1, subjects were to be randomized to fixed dose of BIIB059 450 mg or placebo.

The study design for PV2 (and subsequent versions) Part A and Part B is shown in Figure 1 and Figure 2 below:

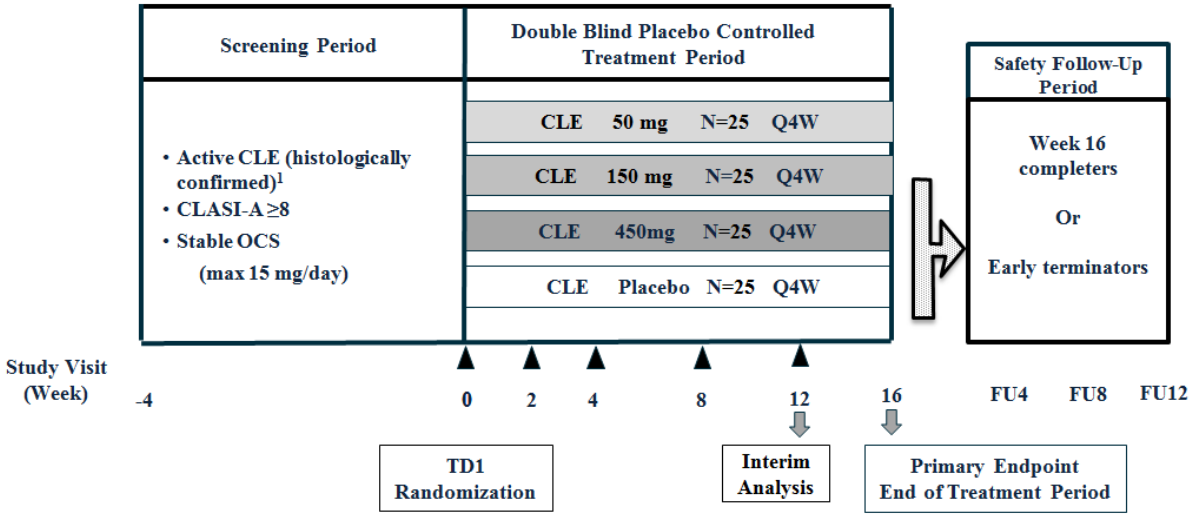
Figure 1: Study Design PV2 (and subsequent versions) Part A



Abbreviations: ANA=antinuclear antigen; dsDNA= double-stranded DNA; FU=follow-up visit; MCP=metacarpophalangeal joints; OCS=oral corticosteroid; PIP=proximal interphalangeal joints; Q4W=every 4 weeks; SoC=standard of care; TD1=Treatment Day 1.

** At least 4 of the swollen joints must be PIP, MCP, and/or wrist joints. A joint that is both tender and swollen will be included in both categories (counts as 1 tender and 1 swollen). Black triangles represent dosing visits.

Figure 2 Study Design PV2 (and subsequent versions) Part B



Abbreviations: CLASI-A=Cutaneous Lupus Erythematosus Disease Area and Severity Index- Activity Scale; CLE=cutaneous lupus erythematosus; FU=follow-up visit; OCS=oral corticosteroid; Q4W=every 4 weeks; SoC=standard of care; TD1=Treatment Day 1

1 Histologically confirmed in the past or at Screening.

Black triangles represent dosing visits

2.1.2. Subject recruitment and treatment

In Part A (PV2), approximately 100 subjects will be randomized to receive either subcutaneous injections of 450 mg BIIB059 or placebo in a 1:1 ratio every 4 weeks (Q4W) for 24 weeks, with an additional dose at Week 2 (for a total of 7 doses, Figure 1). Up to 70 additional subjects may be added at the interim analysis (see Section 10). Randomization will be stratified by oral corticosteroid usage (≤10 mg vs. >10 mg) and by geographic region (United States vs. Asia vs. Latin America and Europe). In Part A (PV1), prior to the amendment, subjects were randomly assigned to 1 of 4 treatment groups in a 1:1:1:1 ratio and received either fixed doses of BIIB059 (50, 150 or 450 mg) or placebo. Randomization, in Part A (PV1), was stratified by oral corticosteroid usage (≤10 mg vs. >10 mg) and by CLASI-A score (≤10 vs. >10). Subjects who have been enrolled in PV1 before the amendment were to remain on their original treatment assignment and were not allowed to reconsent under PV2 and subsequent version.

In Part B (PV2), approximately 100 subjects will be randomized to 1 of 4 treatment groups in a 1:1:1:1 ratio and will receive either fixed doses of BIIB059 (50, 150 or 450 mg) or placebo subcutaneously every 4 weeks (Q4W) for 16 weeks, with an additional dose at week 2, for a total of 5 doses (Figure 2). Randomization will be stratified by CLASI-A score (≤10 vs. >10) and DLE (presence, absence). In Part B (PV1), prior to the amendment, subjects were randomized to either 450 mg BIIB059 or placebo (in a 2:1 ratio). Randomization, in Part B (PV1), was stratified by SLE status (presence, absence) and by CLASI-A score (≤10 vs. >10). Subjects enrolled under PV1 were to remain on their original treatment assignment but could be reconsented under PV2 and subsequent versions.

Additional details of the study design can be found in full protocol of 230LE201.

3. STATISTICAL ANALYSIS

3.1. General Considerations

Three analyses are planned for this study and apply to PV2 and subsequent versions:

1. The IA for futility in Part B will be performed after approximately 45% of the subjects in Part B (PV2 and subsequent version) have completed their week 12 treatment visit. Further details are included in Section 10.
2. The final analysis for Part A will be performed when all subjects in Part A have completed the last study visit including follow-up period.
3. The final analysis of Part B will be performed when all subjects in part B have completed the last study visit including follow-up period.

A separate analysis will be performed for Part A and Part B, unless otherwise specified.

Within each part, data from PV1 will be pooled with PV2 (and subsequent versions), by treatment group, for each planned analysis as per the following criteria.

Part A:

- For the planned analyses of the primary and other endpoints associated with the joint count assessment, pool (by treatment group) only the data from PV1 and PV2 (and subsequent versions) subjects who satisfy the following inclusion criteria at baseline:
    - Must have at least 4 tender joints based on assessment of 28 joints.
    - Must have at least 4 swollen joints based on assessment of 28 joints, as described above, with at least 4 of the swollen joints occurring in the PIP, MCP, or wrist joints.
- For all the planned analyses of endpoints associated with CLASI-A assessment, pool (by treatment group) PV1 data with only the data from PV2 (and subsequent versions), subjects with baseline CLASI-A ≥ 8.

- For the planned analyses of all other endpoints pool (by treatment group) all the data from PV1 and PV2 (and subsequent version) subjects.

Part B: Pool (by treatment group) all the data from PV1 and PV2 (and subsequent versions), subjects.

A list of all the primary and secondary analyses and target population is provided in Appendix A1 (Part A) and Appendix A2 (Part B).

This statistical analysis plan covers the planned analyses for both the unblinded IA and the final analyses (Part A and Part B). All efficacy, safety and PK endpoints described in this plan will be included in the final analyses, while a subset of the endpoints will be included in the unblinded IA (see Appendix B for more details). For each endpoint, the same analysis method will be used for both the interim and final analyses, unless otherwise specified.

Within each part (Part A and Part B), all data will be summarized by treatment group. In addition for each part (Part A and Part B), the data described in Section 5 which includes patient disposition, baseline characteristics and demography data, and concomitant medications will be summarized overall, and the safety data described in Section 7 will also be summarized for all BIIB059 groups combined.

A pooled analysis (pooled data from Part A and Part B) may also be performed on Part B endpoints associated with CLASI-A assessment. Only the data from Part A subjects that meet the following criteria will be pooled (by treatment group) for the combined (Part A and Part B) analyses:

- Baseline CLASI-A ≥ 8

The pooled analyses will be performed using the same analysis method as for the CLASI-A endpoints within Part B, where applicable.

The statistical software, SAS® version 9.3 or above, will be used for all summaries and analyses.

3.2. Changes to Protocol Planned Analyses

- [redacted]
- [redacted]
- A secondary endpoint was added:
  - Proportion of subjects with a ≥7-point reduction in CLASI-A score at Week 24 (Part A) and at Weeks 12 and 16 (Part B) compared with baseline.
- [redacted]
  - [redacted]
  - [redacted]
  - [redacted]
  - [redacted]
  - [redacted]
  - [redacted]
  - [redacted]
  - [redacted]
  - [redacted]

- [redacted]
  - [redacted]
  - [redacted]
  - [redacted]
  - [redacted]
  - [redacted]
  - [redacted]
  - [redacted]
  - [redacted]
  - [redacted]
  - [redacted]
- [redacted]
  - [redacted]

[redacted]

- The baseline*visit interaction term was added to MMRM model as covariates used for analyses of continuous endpoints.
- In Part A, for binary endpoints related to CLASI-A score, CLASI-A score (≤10 vs. >10) was added to the logistic model as a covariate. In both Part A and B, for binary endpoints dichotomized from a continuous endpoint other than CLASI-A, the baseline of the corresponding continuous endpoint was added to the logistic model as a covariate.

3.3. Changes made in the Statistical Analysis Plan

- Interim analysis 2 is added to the statistical analysis plan according to the changes made in protocol version 5.0.
- Patients who have usage of disallowed concomitant therapy or SLE standard of care therapy that deviate from the protocol will be considered as major deviations that

could impact the efficacy assessments and will be excluded from the per-protocol population (Part A and B).

- Failing Inclusion criteria 8.1.3 is now considered to impact the efficacy analysis.
- Endpoint 'proportion of subjects with no worsening from baseline in subjects with lupus disease activity defined by <10% increase in PGA (VAS)' is now changed to 'by <1 point increase in PGA (VAS)', so as the corresponding components in SRI-4, -5, -6 and BICLA.
- [redacted]
- [redacted]
- Subgroup analysis will be modeled by MMRM model. To avoid inconsistency, if the MMRM model does not converge for some subgroups, ANCOVA model will no longer be performed for these same subgroups.

4. ANALYSIS POPULATIONS, STRATIFICATION FACTORS and VISIT DEFINITIONS

4.1. Definitions of Analysis Populations for Part A and Part B

Modified Intent-to-Treat (MITT) Population

The MITT population is defined as all randomized subjects who received at least 1 dose of study treatment (whether the subjects adhered to the protocol). Subjects will be analyzed according to the study treatment to which they were randomized. Analysis of other non-safety endpoints (e.g., demographics) will also be based on the MITT population.

Safety Population

The safety population will include the same subjects as defined for the MITT population. However, subjects will be analyzed by actual treatment received. If all subjects take the treatment as they were randomly assigned to, the safety population will be the same as the MITT population.

Per-Protocol (PP) Population

The per-protocol population will include the MITT subjects who do not have any major protocol deviations that could impact the efficacy assessments. Subjects will be analyzed by randomized treatment. Major deviations that could impact the efficacy assessments are:

Part A

- Actual treatment received is different from the planned treatment (e.g., patients received BIIB059 treatment while randomized to the placebo arm).
- Patients who have usage of disallowed concomitant therapy or SLE standard of care therapy that deviate from the protocol (specific rules for each concomitant therapy will be described in an addendum).
- Failing Inclusion/exclusion criteria that might potentially affect efficacy. The criteria listed in Table 1 are considered not going to affect efficacy, thus they are not considered for per-protocol population.
- < 80% treatment compliance with study medication

- > 120% treatment compliance with study medication

Part B

- Actual treatment received is different from the planned treatment (e.g., patients received BIIB059 treatment while randomized to the placebo arm).
- Patients who have usage of disallowed concomitant therapy or SLE standard of care therapy that deviate from the protocol (specific rules for each concomitant therapy will be described in an addendum).
- Failing Inclusion/exclusion criteria that might potentially affect efficacy. The criteria listed in Table 1 are considered not going to affect efficacy, thus they are not considered for per-protocol population.
- < 80% treatment compliance with study medication
- > 120% treatment compliance with study medication

Table 1 Inclusion/Exclusion Criteria not Considered for Per-Protocol Population

Section in Protocol	Item	Affect Efficacy
8.1.1 Inclusion Criteria (Parts A and B)	1. Ability of the subject to understand the purpose and risks of the study and provide signed and dated informed consent and authorization to use confidential health information in accordance with national and local subject privacy regulations.	Not related to efficacy
	2. Age 18 to 75 years old, inclusive, at the time of informed consent.	Relaxed to age 12 to 80
	3. All women of childbearing potential and all men must practice effective contraception during the study and for 16 weeks after their last dose of study treatment.	Not related to efficacy
8.2.1 Exclusion	5. Other unspecified reasons that, in the opinion of the Investigator or Biogen, make the subject unsuitable for enrollment.	Not related to efficacy

Criteria (Parts A and B)	10. Subjects with a history of suicide attempt or suicidal ideation within 1 year prior to Screening.	Not related to efficacy
	13. History of severe herpes infection, such as herpes encephalitis, ophthalmic herpes, or disseminated herpes.	Not related to efficacy
	14. History of chronic, recurrent (3 or more of the same type of infection in a 12-month period), or recent serious infection (e.g., pneumonia, septicemia, herpes zoster) as determined by the Investigator and requiring anti-infective treatment within 12 weeks prior to Screening.	Not related to efficacy
	15. Signs of herpes or varicella zoster viral infection (specifically chicken pox, shingles, or herpes zoster) within 12 weeks prior to Screening.	Not related to efficacy
	16. History of or current diagnosis of active tuberculosis (TB), or untreated latent TB infection (LTBI), determined by a TB skin test with purified protein derivative as evidenced by induration ≥5 mm or a positive Quantiferon, positive or borderline T-SPOT (Elispot) test performed locally, either at Screening or documented with results within 12 weeks of the Screening Visit. Subjects who have previously completed appropriate and documented LTBI treatment will not be required to be tested. Subjects must have received complete LTBI treatment prior to Screening without evidence of re-exposure prior to entering the study. Subject with current household contacts with active TB will also be excluded unless the subject is being treated and there is evidence that household contacts are being treated. Indeterminate Quantiferon or T-SPOT tests may be repeated once and will be considered positive if retest results are positive or indeterminate. Subjects with documented BCG vaccination must perform a TB test at Screening and will be excluded if skin induration ≥5 mm or a positive Quantiferon or positive or borderline T-SPOT (Elispot) test.	Not related to efficacy
	18. Known hypersensitivity to BIIB059 or any of the components of the formulated BIIB059 or matching placebo.	Not related to efficacy
	19. History of, or ongoing, malignant disease, including solid tumors and hematologic	Not related to efficacy

	malignancies with the exception of basal cell carcinomas and squamous cell carcinomas and carcinoma in situ of the cervix that have been completely excised and considered cured >2 years prior to Screening.	
	21. History of substance abuse (except for cannabinoid) within 24 weeks prior to Screening, based on the Investigator's opinion.	Not related to efficacy
	22. Female subjects who are pregnant, currently lactating, have stopped lactating in the past 12 weeks, or are planning to become pregnant during the study and for 4 months after the last dose of study treatment.	Not related to efficacy

PK Population

The PK population will include the safety subjects who have at least one PK concentration measurement. Subjects will be analyzed by the actual treatment received.

4.2. Stratification Factors

The randomization stratification used for both Part A and B under each protocol versions (PV1 and PV2) is shown below:

Part A

PV1

- Oral Corticosteroid Usage (≤10 mg/d vs. >10 mg/d)
- CLASI-A score (≤10 vs > 10)

PV2 (or later)

- Oral Corticosteroid Usage (≤10 mg/d vs. >10 mg/d)
- Geographic Region (United States vs. Asia vs. Latin America and Europe)

Part B

PV1

- CLASI-A score (≤10 vs. >10)

- SLE (Yes/No)

PV2 (or later)

- CLASI-A score (≤10 vs. >10)
- DLE (Yes/No)

4.3. Visit/Analysis Week Definitions

The definition of visits below applies to both parts A and B.

Baseline

Baseline will refer to the last non-missing value prior to the first randomized dose.

Treatment Day [TD] (Days relative to the first randomized dose)

To facilitate the analysis of the data in the treatment period (TP), TD will be calculated for each day in the study as the number of days between the date of the first randomized dose and the specific day of interest. For a day that is on or after the day of the first randomized dose, the TD is a positive value and will be calculated as (date of study day – date of the first randomized dose +1). For a day that is prior to the date of the first randomized dose, the TD is a negative value and will be calculated as (date of study day – date of the first randomized dose). The study starts at TD1.

Analysis Week

Observations that are collected after the first randomized dose as scheduled visit will be used to the corresponding analysis week. Baseline corresponds to "Week 0" in the protocol or the latest non-missing value prior to first study drug administration if Week 0 assessment is missing.

Data from Early Termination or Unscheduled Visits

In general, when data are summarized by visit/analysis week, the Early Termination Visit or Unscheduled Visit will be assigned to an appropriate scheduled visit/analysis week (PV2 or later) by using a window scheme as follows.

The lower bound and the upper bound for the visit windows are defined as the midpoints of the scheduled visits (PV2 or later). If the date and time from early termination visit falls in between the lower bound and the upper bound for a scheduled visit, then it will be assigned to that visit. If there are 2 or more assessments available in the same analysis window for a subject, the assessment that is closest to the target visit day will be used for analysis. If there are 2 or more assessments in the same analysis window with the same distance from the target visit day, the earlier assessment will be used.

The windowing scheme for the primary endpoint in Part A and Part B is shown in Table 2 and Table 3 below.

Table 2 Visit Windows for Primary Endpoint (Joint Count) in Part A

Analysis visit	Target visit day	Analysis visit window
Baseline (Week 0)	1	Most recent non-missing pre-dose value
Week 2	15	[2, 22]
Week 4	29	[23, 43]
Week 8	57	[44, 71]
Week 12	85	[72, 99]
Week 16	113	[100, 127]
Week 20	141	[128, 155]
Week 24	169	[156, 197]
Follow-up 12	253	≥198

Table 3 Visit Windows for Primary Endpoint (CLASI) in Part B

Analysis visit	Target visit day	Analysis visit window
Baseline (Week 0)	1	Most recent non-missing pre-dose value
Week 2	15	[2, 22]
Week 4	29	[23, 43]
Week 8	57	[44, 71]
Week 12	85	[72, 99]
Week 16	113	[100, 141]
Follow-up 12	197	≥142
*For Part B subjects in PV1 who didn't reconsent to PV2 and hence don't have week 16 visit, the follow-up period days and lower and upper bound of visit window will shift back by 28 days		

5. STUDY SUBJECTS

5.1. General Consideration

Unless otherwise specified, all variables in Section 5 will be summarized for the MITT population for both Parts A and B.

5.2. Subject Disposition

For randomized subjects, the numbers and percentages of subjects who were randomized and dosed, discontinued study treatment in treatment period (including reasons for discontinuation), completed the study treatment, completed the study, and withdrew from the study in the treatment period (including reasons for withdrawal) will be summarized by treatment group.
Additionally, a listing of subjects who discontinued treatment and/or withdrew from study during the treatment period and the associated reasons for discontinuation/withdrawal will be provided.
The number and percentage of subjects in each treatment group will be summarized for the MITT population.

5.3. Demography and Baseline Disease Characteristics – Part A and B

The following demographic and baseline characteristics will be summarized for the MITT population only – note that if the safety population differs from the MITT population (i.e. if not all subjects receive their randomized treatment) then the summaries will be repeated for the safety population only.

Part A

- Age categories (<40 years, ≥40 - <65 years, ≥65)
- Age (years)
- Sex (male/female)
- Ethnicity (Hispanic or Latino/not Hispanic or Latino/ Not reported due to confidentiality regulations)
- Race (American Indian or Alaska Native/Asian/Black or African American/ Native

Hawaiian or Other Pacific Islander/White/Not reported due to confidentiality regulations/Other)

- Height (cm)
- Weight (kg)
- Body Mass Index (kg/m^2)
- Geographic region (United States, Asia, Latin America, Europe)
- CLE (Yes/No)
- CLE subtypes
    - ACLE
    - SCLE
    - DLE
    - CCLE other
        - Panniculitis
        - Chiblain
        - Tumidus

Additionally, the following baseline assessments will be summarized.

- Total active joint count
- Tender joint count
- Swollen joint count
- CLASI-A score
- [redacted]
- CLASI-A (≤10, >10)
- SLEDAI-2K score
- SLEDAI-2K (<10, ≥10)
- BILAG-2004 score
- BILAG-2004 grade (A, B, C, D, E)
- PGA of SLE (VAS) score

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- Concomitant medications
  - Antimalarials (Yes/No)
  - Corticosteroids (Yes/No)
  - Corticosteroids and antimalarials (Yes/No)
  - Azathioprine (Yes/No)
  - Methotrexate (Yes/No)
  - Mycophenolate (Yes/No)
  - Other allowed Medications (Yes/No)
- [redacted]
- [redacted]
- [redacted]
- [redacted]

Part B

- Age categories (<40 years, ≥40 - <65 years, ≥65 years)
- Age (years)
- Sex (male/female)
- Ethnicity (Hispanic or Latino/not Hispanic or Latino/ Not reported due to confidentiality regulations)

- Race (American Indian or Alaska Native/Asian/Black or African American/ Native Hawaiian or Other Pacific Islander/White/Not reported due to confidentiality regulations/Other)
- Height (cm)
- Weight (kg)
- Body Mass Index (kg/m^2)
- Geographic region (United States, Asia, Latin America, Europe)
- SLE (Yes/No)
- CLE subtypes
    - ACLE
    - SCLE
    - DLE
    - CCLE other
        - Panniculitis
        - Chiblain
        - Tumidus

Additionally, the following baseline assessments will be summarized.

- CLASI-A score
- CLASI-A (≤10, >10)
- [redacted]
- SLEDAI-2K score
- SLEDAI-2K (<10, ≥10)
- PGA of SLE (VAS) score
- [redacted]
- [redacted]
- [redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- Concomitant medications
    - Antimalarials (Yes/No)
    - Corticosteroids (Yes/No)
    - Corticosteroids and antimalarials (Yes/No)
    - Azathioprine (Yes/No)
    - Methotrexate (Yes/No)
    - Mycophenolate (Yes/No)
    - Other Allowed Medications (Yes/No)
- [redacted]
- [redacted]

All baseline characteristics will also be presented in subject listings.

5.4. Prior Medications

Prior medications used for SLE and/or CLE

Previous medications used for the treatment of SLE and/or CLE(NSAIDS/Immunosuppressants/corticosteroids/antimalarials/others) will be summarized. A listing of previous medications used for the treatment of SLE and CLE will be presented, including the previous medication used, duration of treatment, used at screening (yes/no), and reasons for discontinuation if not using at screening (lack of efficacy/side effects/or other).

Prior medications used for non SLE/CLE

Previous medications used for the non SLE/CLE related treatment will be summarized.

A listing of previous medications used for the non SLE/CLE related treatment will be presented, including the previous medication used, duration of treatment, used at screening (yes/no), and reasons for discontinuation if not using at screening (lack of efficacy/side effects/or other).

5.5. Concomitant Therapies

A concomitant therapy (including medication or non-drug therapy) is defined as any therapy (other than the treatment) that is taken on or after the date/time of the randomized dose. This includes therapies that are started prior to the date/time of randomized dose if their use continues on or after Treatment Day 1.

In order to define concomitant for therapies with missing start or stop date, the following additional criteria are defined:

- if both the start and stop dates of a therapy are missing, that therapy is considered concomitant;
- if the start date of a therapy is missing and the stop date of that therapy falls on or after the date of randomized dose, that therapy is considered concomitant;
- if the start date of a therapy is prior to the date/time of randomized dose and the stop date of that therapy is missing, that therapy is considered concomitant;
- if the start and the stop dates of a therapy is prior to the date/time of randomized dose, that therapy is considered prior.

All concomitant medications will be coded using the World Health Organization (WHO) medication dictionary (March 2017 Version). All concomitant non-drug treatments will be coded by Medical Dictionary for Regulatory Activities (MedDRA) preferred term (Version 21.1 or the most updated version at data base lock).

Concomitant medications will be summarized separately for medications used to treat SLE, medications used to treat CLE, and medications used for other indications. The number and

percentage of subjects taking each medication as collected in the eCRF will be summarized. All concomitant medications will be presented in the data listings.

5.6. Treatment Exposure and Compliance

The duration of treatment exposure during the treatment period will be based on the number of days between the date of first dose and the date of the last dose + 28 days. Each dose should be administered in the clinic during the randomization visit and recorded in the electronic case report form (eCRF).

The number and percentage of subjects in each of the following cumulative durations of exposure categories will be summarized for each part as follows:

Part A

At least 2, 4, 8, 12, 16, 20, weeks. In addition, days of exposure to the treatment will be summarized as a continuous variable with summary statistics.

Part B

at least 2, 4, 8, 12, weeks. In addition, days of exposure to the treatment will be summarized as a continuous variable with summary statistics.

BIIB059 treatment compliance during the treatment period will be assessed based on the drug administered as recorded in the eCRF. Overall compliance for each subject will be calculated and summarized for each treatment group. The compliance will be calculated using the following formula:

Percent Compliance = (actual number of doses administered)/(number of doses expected to be administered) *100

Note: For part B subjects, the "number of doses expected" in the formula above will depend on the protocol version to which the subject consents. For discontinued subjects, "number of doses expected" will be relative to the time of discontinuation.

5.7. Protocol Deviations

Protocol deviations identified during site monitoring will be captured in a Protocol Deviation Log and will be categorized as major or minor deviations based on protocol deviation classification prior to the final database lock. The major protocol deviations will be summarized and listed. The minor protocol deviations will also be listed. All summaries for protocol deviations will be presented only by treatment group.

Subjects that had at least one major protocol deviation that led to exclusion from the per-protocol population (see Section 4.1) will also be summarized.

5.8. Treatment Failure

For each part below, subjects who meet any of the following criteria will be considered as a treatment failure and considered as a non-responder in the efficacy analysis:

Part A

- Subjects who initiate or increase SLE standard of care therapy or other disallowed concomitant therapy (Table 8 of the latest protocol version) during the treatment period of the study.
- Subjects with corticosteroid dose increase above baseline during the treatment period except for what is allowed under rescue 1 during week 1 to 4.
- Subjects with more than 2 allowed rescues between week 1 to week 12.
- Subjects with any increase in corticosteroid during treatment period after week 20.

Part B

- Subjects who initiate disallowed concomitant therapy during treatment period of Part B (see Table 8 of the latest protocol version).
- Subjects who initiate or increase SLE/CLE standard of care therapy during the treatment period of the study.
- Subjects who initiate or increase corticosteroid dose above baseline during the treatment period of the study (oral or topical except those in Table 6 of the latest protocol version).

Note: Treatment failure subjects would be allowed to remain in the trial and receive investigational product, as applicable)

More detailed definition regarding specific concomitant therapy is described in and addendum to this SAP. For both parts, subjects' data considered as treatment failure will be handled as described in Section 6.6.

6. EFFICACY ANALYSES

6.1. General Considerations

For both Part A and Part B, the efficacy analyses will be performed using the MITT population and will be considered as the primary analyses. For each endpoint, additional conditions may apply for pooling of data across protocol versions (see Section 3.1). The final model for each analysis in the MITT population will also be used for the PP analysis. A complete list of all the planned analyses and target population for the primary and secondary endpoints is provided in Appendix A1 (Part A) and Appendix A2 (Part B).

Unless otherwise specified, statistical testing across all efficacy endpoint will be performed to compare each BIIB059 dose (Part A: 450 mg; Part B: 50, 150 or 450 mg) with placebo at the 2-sided, 0.05 significance level. P-values and corresponding 95% confidence intervals (CI) will be provided. P-values will be rounded to three decimal points before assessing statistical significance.

Summary statistics will be presented for all endpoints at all visits including follow-up visits. For continuous variables, the summary statistics will include number of subjects with data, mean, standard deviation (SD), median, first quartile (Q1), third quartile (Q3), minimum and maximum. For categorical endpoints, the summary statistics will generally include: number of subjects in the corresponding analysis population, number and percentage of subjects in each category.

Unless otherwise specified, week 2, 4, 8, 12, 16, 20 and, 24 will be included in all Part A "by visit" MMRM analyses.
Unless otherwise specified, week 2, 4, 8, 12 and,16 will included in all Part B "by visit" MMRM analyses.
Additional study part specific efficacy analyses considerations are described in Section 6.1.1 (Part A) and Section 6.1.2 (Part B)

6.1.1. Part A Efficacy Analyses Considerations

Continuous endpoint

All Part A continuous efficacy endpoints will be analyzed using the following described method, unless otherwise specified.
A Mixed Effect Model Repeat Measurement (MMRM) model will be used as the primary method to analyze endpoint of interest using treatment group (BIIB059 450 mg vs. placebo), study visit (week 2, 4, 8, 12, 16, 20, 24), baseline corticosteroid usage level (≤ 10 mg, >10 mg), region (USA, Asia, Latin America & Europe), study visit-by-treatment interaction, baseline value of the endpoint, and baseline-by-visit interaction as fixed effect covariates. An unstructured covariance matrix will be used to model the within-patient variance-covariance errors. If the unstructured covariance structure matrix results in a lack of convergence, the heterogeneous Toeplitz covariance structure followed by the heterogeneous first-order autoregressive covariance structure will be used. The Kenward-Roger approximation will be used to estimate the denominator degrees of freedom. The primary treatment comparison will be the difference between the BIIB059 450 mg dose and placebo at the end of the Week 24

visit. Handling of missing data and data from treatment failures subjects is provided in Section 6.6.

Binary endpoint

The binary endpoints will be analyzed using logistic regression with covariates of treatment (BIIB059 450 mg vs. placebo), baseline corticosteroid usage level (≤ 10 mg, >10 mg) and region (USA, Asia, Latin America & Europe). For binary endpoints related to CLASI-A score, CLASI-A score (≤10 vs. >10) will be added to the above covariates list. For endpoints dichotomized from a continuous endpoint (such as PGA VAS < 1 point increase) the baseline of the continuous endpoint will be added to the above covariates list. The odds ratio, its 95% CI and the p-value for BIIB059 450 mg dose compared to placebo will be presented. Early discontinuation(s) and/or treatment failures will be handled using non-responder imputation approach as discussed in Section 6.6.

[redacted]

6.1.2. Part B Efficacy Analyses Considerations

Continuous endpoint

All Part B continuous efficacy endpoints will be analyzed using the following described method, unless otherwise specified.

A Mixed Effect Model Repeat Measurement (MMRM) model will be used as the primary method to analyze endpoint of interest using treatment group, study visit (week 2, 4, 8, 12, 16),

study visit-by-treatment interaction, DLE (Yes/No), CLASI-A score (≤10 vs. >10), baseline value of the endpoint and baseline-by-visit interaction as fixed effect covariates. An unstructured covariance matrix will be used to model the within-patient variance-covariance errors. If the unstructured covariance structure matrix results in a lack of convergence, the heterogeneous Toeplitz covariance structure followed by the heterogeneous first-order autoregressive covariance structure will be used. The Kenward-Roger approximation will be used to estimate the denominator degrees of freedom. The primary treatment comparison between each BIIB059 treated group and placebo will be at the end of the Week 16. Handling of missing data and data from treatment failures subjects is provided in Section 6.6.

Binary endpoint

The binary endpoints will be analyzed using logistic regression with covariates of treatment and DLE (Yes/No) and CLASI-A score (≤10 vs. >10). For binary endpoints dichotomized from a continuous endpoint (such as PGA VAS < 1 point increase, the baseline of the continuous endpoint will be added to the above covariates list. The odds ratio, its 95% CI and the p-value for each BIIB059 dose compared to placebo will be presented. Early discontinuation(s) and/or treatment failures will be handled using non-responder imputation approach as discussed in Section 6.6.

[redacted]

Censoring rules: Subjects who do not experience an event by Week 16 or early termination will be censored at the date of last assessment during the treatment period. The start date for calculation of day to censor or event will be the date of first dose.

6.2. Analysis of Primary Endpoint

6.2.1. Primary Analysis

Part A

The primary analysis will be performed on MITT subjects with a baseline and at least one post-baseline joint count assessment. For joint count assessments, the analyses will be conducted using pooled subjects' data from PV2 and PV1 who meet the criteria as described in Section 3.1.

The continuous primary endpoint, change from baseline to week 24 in the active joint count, will be analyzed as described for continuous endpoint analysis in Section 6.1.1

Part B

The primary analysis will be performed on MITT subjects with a baseline and at least one post-baseline CLASI-A assessment. The analyses will be conducted using pooled subjects' data from PV2 and PV1 who meet the criteria as described in Section 3.1.

The primary analysis of the primary endpoint, percent change from baseline to week 16 in CLASI-A score, is a test of dose-response, using the Multiple Comparison Procedure – Modelling (MCP-Mod) methodology [Bretz 2005].

Five dose-response trends will be tested using the appropriate contrasts, as determined by the MCP-Mod methodology, on the treatment effects obtained from the MMRM model. The MMRM model will be applied to percent change from baseline in CLASI-A score as described for endpoint analysis in Section 6.1.2

For MCP-Mod Candidate Models, two families of dose-response models (or 'candidate models') will be considered:

- SigEmax models (4 candidates) - parameters to estimate: e0, emax, h and ED50

*SigEmax — μd = e0 + emax * d^h/(ED50^h + d^h),*

where: *emax* = asymptotic maximum effect; *ED50* = Dose giving half of the asymptotic maximum effect; *e0* = placebo effect and *h* = shape parameter.

- Beta model (non-monotonic, 1 candidate)

  Beta model — $\mu d = e0 + emax\ B(\delta 1, \delta 2)\ x\ ^{\delta 1}\ (1-x)^{\delta 2}$

where: *emax* = asymptotic maximum effect; *e0* = placebo effect;

$B(\delta 1, \delta 2) = (\delta 1 + \delta 2)^{(\delta 1 + \delta 2)} / (\delta 1^{\delta 1}\ \delta 2^{\delta 2})$; x = d/scale

The initial parameters for candidate models are as follows:

- SigEmax model 1: (ED50 = 25, h = 2.5)
- SigEmax model 2: (ED50 = 50, h = 2.1)
- SigEmax model 3: (ED50 = 90, h = 2.5)
- SigEmax model 4: (ED50 = 450, h = 1)
- Beta Model: (scale = 720) ($\delta 1 = 0.6425, \delta 2 = 1.235$)

The maximum treatment effect (over placebo within the dose range) is assumed to be 27.5 at week 16. The four candidate SigEmax models were considered in order to capture different ED50 values (ranging from 25 to 450 mg). The beta model was considered to capture the scenario where maximum benefit is not observed at the highest but is observed at a lower dose.

If at least one model is significant, there is an indication for evidence of a dose-response effect. A candidate model will be selected using model selection criteria AIC (Akaike Information Criteria). The selected model will be used to produce estimates of target doses using modeling and simulations.

To further understand the dose-response relationship, based on the model type (sigmoid Emax model or Beta model) selected from the MCP-Mod method, a sigmoid Emax model or Beta model will be fitted using data from week 16. For subjects who are missing with week 16 visit data, the predicted value from the MMRM model will be used to fit the does-response model.

The dose-response model parameters, such as Emax and ED50, will be estimated as well as the treatment effect of each dose group.

[redacted]

[redacted]

6.2.2. Sensitivity Analysis

A sensitivity analysis will be performed on both part A and part B primary endpoints using the same methods as described in Section 6.2.1 but using a different missing data handling approach, reference based multiple imputation, as described in Section 6.6.

6.3. Analysis of Secondary Efficacy Endpoints

6.3.1. CLASI-50 Response (Yes/No)

Part A

The percent reduction from baseline to Week 24 in the CLASI-A score will be calculated for each subject. The proportion of subjects who had CLASI-50 response, defined as a ≥50% reduction in the CLASI-A score from baseline, will be analyzed as described for binary endpoint analysis in Section 6.1.1

Part B

The percent reduction from baseline to Week 12 and 16 in the CLASI-A score will be calculated for each subject. The proportion of subjects who had CLASI-50 Response at week 12 and 16, defined as a ≥50% reduction in the CLASI-A score from baseline, will be analyzed separately for each time point as described for binary endpoint analysis in Section 6.1.2.

6.3.2. A 4-Point Reduction in CLASI-A Response (yes/no) at Week 24 (Part A) and Weeks 12 and 16 (Part B)

Part A

The reduction from baseline to Week 24 in the CLASI-A score will be calculated for each subject. The proportion of subjects who had ≥4-point CLASI-A response at week 24, defined as a ≥4-point reduction in the CLASI-A score from baseline at week 24, will be analyzed as described for binary endpoint analysis in Section 6.1.1.

Part B

The reduction from baseline to Week 12 and 16 in the CLASI-A score will be calculated for each subject. The proportion of subjects who had ≥4-point CLASI-A response at week 12 or week 16, defined as a ≥4-point reduction in the CLASI-A score from baseline at week 12 or 16, will be analyzed separately for each time point as described for binary endpoint in Section 6.1.2.

6.3.3. A 7 Point Reduction in CLASI-A Response (yes/no) at Week 24 (Part A) and Weeks 12 and 16 (Part B)

Part A

The reduction from baseline to Week 24 in the CLASI-A score will be calculated for each subject. The proportion of subjects who had ≥7-point CLASI-A response at week 24, defined as a ≥7-point reduction in the CLASI-A score from baseline at week 24, will be analyzed as described for binary endpoint analysis in Section 6.1.1

Part B

The reduction from baseline to Week 12 and 16 in the CLASI-A score will be calculated for each subject. The proportion of subjects who had ≥7-point CLASI-A response at week 12 or week 16, defined as a ≥7-point reduction in the CLASI-A score from baseline at week 12 or 16, will be analyzed separately for each time point as described for binary endpoint in Section 6.1.2.

6.3.4. Percent Change in CLASI-A Score from Baseline at Week 12, 16 and 24 (Part A) and Week 12 (Part B)

The change from baseline to each post-baseline week in the CLASI-A score will be analyzed will be analyzed as described for continuous endpoint analysis in Section 6.1.1 (Part A) and Section 6.1.2 (Part B). Summary statistics for actual values and change from baseline will also be presented by week based on observed data.

6.3.5. Proportion of Subjects with an SLE Responder Index (SRI) Response of ≥4 at Week 24 (Part A Only)

SLE Responder Index (SRI) of ≥4 (SRI-4) at Week 24 is a categorical response variable (yes/no) incorporating the following criteria for achievement of responder status (i.e., all criteria must be met to achieve responder status):

- A reduction from baseline of ≥4 points in Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K), and
- No new organ system affected, as defined by no new British Isles Lupus Activity Group (BILAG)-2004 grade A and no more than one new BILAG-2004 grade B, and
- No worsening from baseline in subject's lupus disease activity defined by <1 point increase in the Physician's Global Assessment (PGA) [visual analog scale (VAS)] (in the scale of 0-10) and
- No changes to protocol specified medications rules as described below (all criteria must be met):
    - No initiation or increase of SLE standard of care therapy or other disallowed concomitant therapy (see table 8 from latest protocol for protocol-prohibited list).
    - Concomitant corticosteroid dosage at Week 24 to be ≤10 mg/day.
    - Concomitant corticosteroid dosage at Week 24 to be ≤Day 1.
    - No increase in corticosteroid dose between weeks 17 and 24.

The proportion of subjects who had a composite response, as defined above, will be analyzed as described for binary endpoint analysis in Section 6.1.1. Proportion of subjects who responded to each of the 4 criteria will also be reported.

6.3.6. Change from Baseline to Week 24 in SLEDAI-2K Score (Part A Only)

The change from baseline to each post-baseline week in the SLEDAI -2K score will be analyzed as described for continuous endpoint analysis in Section 6.1.1.

6.3.7. Proportion of Subjects with No New BILAG-2004 A and no More than One New BILAG-2004 B at Week 24 (Part A Only)

The proportion of subjects who had no new BILAG-2004 A and no more than one new BILAG-2004 B from baseline to Week 24 will be analyzed as described for binary endpoint analysis in Section 6.1.1.

6.3.8. Change from Baseline to Week 24 in PGA of SLE (VAS) Score (Part A Only)

The change from baseline to week 24 in the PGA of SLE (VAS) score will be analyzed as described for continuous endpoint analysis in Section 6.1.1.

[redacted]

- [redacted]
- [redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
  - [redacted]
    - [redacted]

- [redacted]
- [redacted]
- [redacted]
  - [redacted]
  - [redacted]
  - [redacted]
  - [redacted]

- [redacted]
- [redacted]
- [redacted]
  - [redacted]
    - [redacted]

[redacted]

- [redacted]
  - [redacted]
  - [redacted]
  - [redacted]
    - [redacted]
    - [redacted]
    - [redacted]
    - [redacted]

- [redacted]
- [redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

- [redacted]
- [redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]

- [redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

- [redacted]

6.5. Multiplicity Adjustment

Part A

There will be no adjustment for multiplicity at interim or final analysis for the primary or secondary endpoints in Part A.

Part B

Multiple Comparison Procedure-Modeling (MCP-Mod) methodology (Bretz et al, 2005) will be used to test the dose-response relationship across 3 doses of BIIB059 and PBO for the percent change from baseline in CLASI-A score at interim and final analysis. Five dose-response trends will be tested using the optimal contrasts as determined by the MCP-Mod methodology such that the overall Type I error rate is controlled at 0.05 (one-sided testing).

There will be no adjustment for multiplicity with secondary efficacy endpoints at interim or final analysis.

6.6. Handling of Missing Data

The handling of missing data and data from treatment failure subjects is summarized here and applies to all the endpoints unless otherwise specified. For composite endpoints (e.g.: SRI-4), the missing values will be imputed at each component endpoint level, as applicable, using the rules described below.

6.6.1. Continuous Endpoints

The primary approach described below will apply to both part A and part B continuous endpoints, unless otherwise specified. The sensitivity approach will be applied to the primary endpoints.

Primary approach

For subjects who are considered as treatment failures, as defined in Section 5.8, the worse observation of the baseline and the last observation before treatment failure carried forward will be used as the primary method to impute values for all the visits post treatment failure visit. Under this approach, subject's post treatment failure data will be censored and then imputed with the subject's worst case non-missing data prior to treatment failure.

Data after treatment discontinuation will be censored and treated as missing data. For all other monotone and non-monotone missing data, including missingness post treatment discontinuation, mixed-effects model repeated measures (MMRM) model based approach will be used as the primary approach. MMRM assumes data is missing at random (MAR).

A few subjects from PV1 part B might have completed treatment up to week 12 but could not reconsent to PV2 and therefore will have missing week 16 data. These subjects will have week 16 data imputed using predicted value from the MMRM model.

Sensitivity analyses approach

For subjects who are considered as treatment failures, as defined in Section 5.8, the worse observation of the baseline and the last observation before treatment failure carried forward will be used to impute values for all the visits post treatment failure visit. Under this approach, subject's post treatment failure data will be censored and then imputed with the subject's worse case between the baseline and the last non-missing data prior to treatment failure.

For all the other missing data, including missingness post treatment discontinuation, reference-based Multiple Imputation (MI) approach will be used to consider a missing not at random (MNAR) mechanism for monotonic missing data. Imputed values of subjects in the BIIB059 group will be based on the Placebo group. This approach assumes no sustained benefit of BIIB059 after discontinuation of study treatment and limits post-discontinuation effects to that of placebo.

Intermittent (non-monotonic) missing data will be imputed based on the MAR assumption and with an imputation model using Markov chain Monte Carlo (MCMC) method within each treatment group. The imputation models will include baseline value and the values at each post-baseline week for endpoint of interest. The MCMC method in the MI procedure in SAS will be used with a single chain, with a burn-in of 1000, and a thinning of 100 and non-informative priors for all parameters.

The remaining, monotonic missing data for all subjects who discontinue study treatment early will be imputed using sequential regression MI model estimated based on data from the placebo group only. Each sequential regression model (i.e., for imputation of values at a given week) will include terms for baseline score, the stratification factors, and all the continuous endpoint scores. Missing values at a given week in placebo and BIIB059 groups will be imputed from the same imputation model. No rounding or range restrictions will be applied to imputed continuous values.

200 imputations will be performed. Each of the 200 imputed datasets will be analyzed using the following method.

Results from analysis of each imputed dataset, i.e., least square (LS) mean differences from placebo and their SEs, will be combined using Rubin's method [Rubin 1987] as implemented in the SAS MIANALYZE procedure to produce a pooled LS mean estimate of treatment difference, its 95% CI, and a pooled p-value. A p-value will only be presented for the primary

comparison at primary timepoint. The LS means of change from baseline to each post-baseline week will be plotted by treatment group.

A figure of the change from baseline to the primary timepoint will be presented; for data imputed using the MI methodology described above the least square mean of the 200 imputed values will be used.

A few subjects from PV1 part B might have completed treatment up to week 12 but could not reconsent to PV2 and therefore will have missing week 16 data. These subjects' week 16 data will be considered MAR and imputed using MMRM.

6.6.2. Binary Endpoints

The non-responder imputation (NRI) approach will apply to both part A and part B binary endpoints, unless otherwise specified.

Subjects who are considered as treatment failures, as defined in Section 5.8, will be considered as non-responders for all the visits post treatment failure visit.

Subjects who discontinued treatment will be classified as non-responders for all the visits following treatment discontinuation. Intermittent (non-monotonic) missing data will be imputed based on last observation carried forward (LOCF) approach. Subjects who completed treatment but had a missing score at primary timepoint (Week 24 of Part A or Week 12, 16 of Part B) will also be classified as non-responder for that timepoint, except for subjects from PV1 Part B who completed treatment up to week 12. These subjects will have the binary response derived based on the imputation of week 16 data under MAR assumption as discussed in Section 6.6.1.

6.6.3. Missing Data in Subcomponents

For missingness in subcomponents of efficacy instruments, such as subcomponents in joint counts, CLASI, SLEDAI-2K, BILAG and etc., imputation will be performed for the missing

subcomponents first. If there are ≤50% subcomponents missing at one visit, missing data of subcomponents will be imputed based on LOCF approach. If there are >50% subcomponents missing at one visit, this visit will be considered missing and the missingness will be imputed at the component level based on approaches described in Section 6.6.1 and Section 6.6.2. The same rule will apply to subcomponents of each component of the composite endpoints (e.g.: SRI-4). For endpoints with more than 1 level of subcomponents, where the overall score consists of multiple domains and each domain consists of multiple subcomponents, this rule will be applied to subcomponents under each domain.

[redacted]

[redacted]

[redacted]

6.7. Per-Protocol Analyses

Part A

The primary efficacy endpoint, active joint count, and the following secondary efficacy endpoints will be analyzed for the per-protocol population using the same analysis methods as for the MITT population.

- Percent change in CLASI-A score from Baseline
- SRI-4 response

Part B

The primary efficacy endpoint, percent change from baseline in CLASI-A score, and the following secondary efficacy endpoints will be analyzed for the per-protocol population using the same analysis methods as for the MITT population.

- Proportion of subjects with a CLASI-50 response, defined as a 50% improvement from baseline in CLASI-A score by visit
- Proportion of subjects with a ≥4-point reduction in CLASI-A score by visit
- Proportion of subjects with a ≥7-point reduction in CLASI-A score by visit

[redacted]

[redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

[redacted]

[redacted]

6.10. Additional Analysis

The following additional analyses will be performed on primary continuous endpoints of both part A and part B.

"Retrieved dropout" analysis

Part A

The primary analysis and sensitivity analyses of the primary endpoint only utilize the active joint count data observed before treatment failure visit or subjects discontinued the study treatment. In this study, subjects who were considered as treatment failures or discontinued the study treatment were encouraged to stay in the study and complete the assessments. [redacted] The change in joint count score will be computed based on all active joint count scores collected, and the same MMRM model as for the primary analysis.

The "retrieved dropout" analysis addresses a different estimand (effectiveness estimand) than the primary and sensitivity analyses (efficacy estimand) of the primary endpoint, by using joint count scores collected after cessation of study treatment, including while being treated with other therapies. Therefore, results of this analysis will not be considered as a sensitivity analysis of the primary endpoint but as an additional analysis addressing a complementary question.

To examine the pattern of missing data, the number and percentage of subjects with an active joint count score at each treatment week will be presented by treatment group for both the observed on-treatment data and the retrieved dropout data.

Part B

The primary analysis and sensitivity analyses of the primary endpoint only utilize the CLASI-A score observed before treatment failure visit or subjects discontinued the study treatment. In this study, subjects who were considered as treatment failures or discontinued the study treatment were encouraged to stay in the study and complete the assessments. [redacted] The change in CLASI-A score will be computed based on all CLASI-A score scores collected, and the same MMRM model as for the primary analysis.

The "retrieved dropout" analysis addresses a different estimand (effectiveness estimand) than the primary and sensitivity analyses (efficacy estimand) of the primary endpoint, by using CLASI-A scores collected after cessation of study treatment, including while being treated with other therapies. Therefore, results of this analysis will not be considered as a sensitivity analysis of the primary endpoint but as an additional analysis addressing a complementary question.

To examine the pattern of missing data, the number and percentage of subjects with an active joint count score at each treatment week will be presented by treatment group for both the observed on-treatment data and the retrieved dropout data.

7. SAFETY ANALYSES

7.1. General Considerations

All safety analyses will be performed for the safety population. No statistical testing will be performed on the safety data.

7.2. Analysis of Adverse Events (AEs)

All AEs will be coded by system organ class and preferred term using the MedDRA dictionary (Version 21.1 or the most updated version at data base lock). In this study, only treatment emergent AEs (i.e. those that occurred or worsened after the first randomized dose through the end of study) will be analyzed. Such AEs will have an onset date on or after the date of first randomized dose or will have presented prior to the first randomized dose and subsequently worsened. Therefore, whenever an analysis of summary of AEs is mentioned, it is intended that this is in reference to treatment-emergent AEs.

The number and percentage of subjects experiencing one or more treatment-emergent AEs in the following categories will be summarized.

- Any AE
- Any mild AE
- Any moderate AE
- Any severe AE
- Any related AE
- Any serious AE
- Any related serious AE
- Any AE that led to withdrawal of study drug
- Any AE that led to study withdrawal
- Any AE that led to drug interrupted, dose not changed
- Deaths

The following summaries will be presented.

- AEs by system organ class (SOC) and preferred term (PT)
- AEs by PT
- AEs by PT with an incidence of 5% or more in any of the treatment groups
- AEs by PT with at least 2% higher in incidence for either BIIB059 group compared to placebo
- Serious AEs by SOC and PT
- Treatment-related serious AEs by SOC and PT
- Severe AEs by SOC and PT
- AEs by relationship to study treatment by SOC and PT
- Treatment-related AEs by SOC and PT
- AEs that led to discontinuation of study drug by SOC and PT
- AEs that led to withdrawal from study by SOC and PT
- AEs by maximum severity by SOC and PT
- Suspected unexpected serious adverse reactions (SUSAR) by SOC and PT

To count the number of subjects who experience each AE, a subject experiencing the same AE multiple times will only be counted once for that preferred term. Similarly, if a subject experiences multiple AEs within the same system organ class, that subject will be counted only once for that system. For subjects experiencing the same AE more than once, the occurrence of the AE with the greatest severity will be used in the calculation of incidence by severity.

The following listings will be presented.

- Listing of serious AEs
- Listing of AEs that led to discontinuation of study drug
- Listing of AEs that led to withdrawal from study
- Listing of all AEs

7.3. Analysis of Laboratory Tests

7.3.1. Shift Analysis for Lab Tests

Laboratory abnormalities will be summarized with shift from baseline tables. Each hematology and chemistry value will be flagged as "low", "normal", or "high" relative to the normal ranges of the laboratory that performed the assay, or as "unknown" if no results are available. Each urinalysis value will be flagged as "positive", "negative", or "unknown" if no values are available.

Shift from baseline to high/low status for hematology and chemistry parameters and shifts from baseline to high/positive for urinalysis will be presented. In each summary, the denominator for the percentage is the number of patients at risk for the shift. The number at risk for a shift to low is the number of subjects whose baseline value was not low and who had at least one post-baseline value. The number at risk for a shift to high is the number of subjects whose baseline value was not high and who had at least one post-baseline value. Subjects will be counted only once for each parameter and each type of shift regardless of how many post-dosing assessments had that type of shift. A shift to high includes normal to high, low to high, and unknown to high; a shift to low includes normal to low, high to low, and unknown to low. A shift to positive includes 'negative' to 'positive' and 'unknown' to 'positive'.

A listing of normal ranges will be provided for all laboratory parameters.

7.3.2. Analysis of Change from Baseline and Actual Values

The absolute and percent changes from baseline by visit will be summarized for quantitative laboratory parameters using descriptive statistics. Actual values will also be summarized in a similar way.

7.3.3. Analysis of Potentially Clinically Significant Abnormalities

The number and percentage of subjects with potentially clinically significant abnormal values for laboratory parameters will be summarized. A listing of potentially clinically significant abnormal

values will also be presented. Potential clinical significance will be determined for selected parameters based on the criteria listed below.

Parameter Name	PCS Low	PCS High
Hematology		
White Blood Cells	<3.0 x 10^3/mm^3	>16 x 10^3/mm^3
Neutrophils	<1.5 k/mm^3	>13.5 k/mm^3
Lymphocytes	<0.8 k/mm^3	>12 k/mm^3
Monocytes	N/A	>2.5k/mm^3
Eosinophils	N/A	>1.6 k/mm^3
Basophils	N/A	>1.6 k/mm^3
Hemoglobin for Females For Males	 ≤ 9.5 g/dL ≤ 11.5 g/dL	 ≥ 17.5 g/dL ≥ 19.0 g/dL
Hematocrit for Females for Males	 ≤ 32% <37%	 ≥ 54% ≥ 60%
Red Blood Cells (RBC)	≦ 3.5 x 10^6/mm^3	≧ 6.4 x 10^6/mm^3
Platelet count	≤ 75 x 10^3/mm^3	≥ 700 x 10^3/mm^3
Chemistry		
Sodium	≤ 126 mEq/L	≥ 156 mEq/L
Potassium	≤ 3.0 mEq/L	≥ 6 mEq/L
Chloride	≤ 90 mEq/L	≥ 118 mEq/L
Bicarbonate	≤ 16 mEq/L	≥ 35 mEq/L
Calcium	<8.0 mg/dL	≧ 12 mg/dL
Magnesium	<1.2 mEq/L	>2.3 mEq/L
Inorganic Phosphorous	≦ 1.7 mg/dL	≧ 5.3 mg/dL
Aspartate aminotransferase (AST)	N/A	≥ 3 xULN
Alanine Aminotransferase (ALT)	N/A	≥ 3 x ULN
Alkaline phosphatase	N/A	≧ 3 x ULN

Creatinine	N/A	>1.5xULN
Total Bilirubin	N/A	>1.5xULN
Total Protein	≤ 4.5 g/dL	≥ 10 g/dL
Albumin	≤ 2.5 g/dL	N/A
Uric Acid for Females for Males	N/ A	≥ 8.5 mg/dL ≥ 10.5 mg/dL
Glucose (Fasting) (Non-fasting)	≦ 40 mg/dL ≤ 40 mg/dL	≧ 160 mg/dL ≧ 250
Creatinine Phosphokinase (CPK)	N/A	>3x ULN
Total Cholesterol	N/A	>300 mg/dL

7.4. Vital Signs

The number and percentage of subjects with clinically relevant abnormalities in vital signs post-baseline will be summarized. Subjects evaluated will be safety subjects with a baseline and at least one post-baseline assessment for that vital sign. Abnormal vital signs will be determined by the following criteria.

Vital Sign	Criteria to determine clinically relevant abnormal vital signs
Temperature	>38°C and an increase from pre-dose of at least 1°C
Pulse	>100 beat per minute (bpm) and an increase from baseline of more than 20 bpm, or <50 bpm and a decrease from baseline of more than 20 bpm.
Systolic Blood Pressure	>160 mmHg and an increase from baseline of more than 40 mmHg, or <90 mmHg and a decrease from baseline of more than 30 mmHg
Diastolic Blood Pressure	>100 mmHg and an increase from pre-dose of more than 30 mmHg, or <45 mmHg and a decrease from pre-dose of more than 20 mmHg

A listing of all vital sign values will be presented for subjects with any vital signs abnormalities. Clinically relevant abnormal vital signs will be flagged in the listing.

7.5. ECG

The number and percentage of subjects with shifts from normal at baseline to abnormal in ECG results will be summarized. If there is no ECG result at baseline, the subject will be counted in the unknown category. A shift to abnormal will include a shift from “unknown” or “normal” at baseline to “abnormal” post- baseline.

7.6. Physical Examination

Abnormal findings which are noted after subjects received study treatment and are deemed by the investigator as clinically significant will be reported as AEs and included in AE analyses.

8. PK ANALYSES

For the PK population, serum concentration of BIIB059 will be summarized descriptively using nominal time points. The mean values (+/- SD) will be plotted over time by treatment group on the linear and log linear scale. Measurements that are below the level of quantification (BLQ) will be set to missing before statistics are calculated, if the BLQ occurs before the first measurable concentration (such as predose), it should be set to zero. If the value of the concentration is "Non-Determinable", then the concentration value will be set to missing. Additionally, a population PK analysis will be conducted to characterize the PK of BIIB059 in subjects with SLE and CLE, estimate population PK parameters of BIIB059, estimate random inter- and intrasubject variability, and estimate random residual variability associated with the estimation of PK model parameters. The PK parameters calculations are detailed in a separate PK analysis plan. PK parameters will be summarized with the following statistics: n, mean, SD, percent coefficient of variation, median, minimum, maximum, and geometric mean.

The population PK analysis may also be used to identify sources of variability (continuous and categorical covariates) that may influence the PK of BIIB059 in patients. In addition, exposure-response (efficacy) analysis may be performed to evaluate the relationship between exposure of BIIB059 and clinical endpoints, which will be used to justify doses for Phase 3 studies. Further details can be found in the PK analysis plan.

9. OTHER ANALYSES

The quantitative or qualitative analysis of the relationship between the BIIB059 exposure and the primary and secondary endpoints will be performed. As described in Section 8, details of the analysis will be included in a separate PK analysis plan for the PKPD modeling.

[redacted]

9.1. Immunogenicity Data

The analysis population for immunogenicity is defined as all subjects who are randomized and who have been dosed with study treatment and have at least 1 post-dose immunogenicity sample collected.

Summary of positive anti-BIIB059 serum antibodies will be presented by visits and by treatment group.

To define immunogenicity status the following rules will be used

- The baseline value is defined as the latest immunogenicity data collected at any time prior to the first dose. If no immunogenicity data are collected prior to the first dose, the baseline value is missing and will be imputed as anti-drug antibody negative for immunogenicity analyses.
- Subjects with at least one confirmed post-treatment positive result will be considered positive for anti-BIIB059 antibodies if their baseline result is negative.
- Subjects where none of the post-treatment samples were positive for anti-BIIB059 antibodies will be considered negative regardless of their baseline result.
- For subjects that are confirmed positive at baseline and have at least one post-treatment sample with a ≥ 4-fold increase in titer will be considered positive for anti-BIIB059 antibodies. Subjects that are positive at baseline, with subsequent post-treatment samples titers that are within 2-fold will be considered negative for anti-BIIB059 antibodies.
- In addition, for subjects that are considered anti-BIIB059 antibody positive with final immunogenicity data, the following may be evaluated:
    - Persistence of the anti-BIIB059 antibody response: More than one positive evaluation that are ≥112 days apart or a positive evaluation at the last time point with no further samples available.
    - Transient anti-BIIB059 antibody response: A single positive evaluation or more than one positive evaluation <112 days apart.

The impact of anti-BIIB059 antibodies on PK parameters may be evaluated by considering antibody status (positive or negative, persistent/transient) as a co-variate in the analysis. The impact on efficacy may be evaluated by summarizing key efficacy endpoints by antibody status (positive or negative, persistent/transient). In addition, to determine the relationship (if any) of

anti-BIIB059 antibodies on safety, incidence of adverse events, serious adverse events and events of particular interest will be presented by antibody status (positive or negative, persistent/transient) as well as by timing of the events.

10. INTERIM ANALYSES

10.1. Interim Analysis 1 (IA1): Futility

An unblinded interim analysis for futility Part B (and possible expansion of Part A) will be performed. The IA will be performed after approximately 45% of the subjects in Part B (PV2 and subsequent version) have completed their Week 12 Treatment Visit.

At the time of the IA, it is estimated that approximately 50% of subjects in Part A will have completed their Week 12 treatment visit. Depending on the outcome of the IA, BIIB059 doses may be adjusted (downwards or upwards), with predicted exposures within what was shown to be safe and well tolerated. In addition, up to 70 subjects may further be enrolled into Part A, which objectives are as follow:

1) potentially include an expanded SLE patient population based on more inclusive criteria,
2) potentially adjust the dose of BIIB059 and,
3) To provide greater precision for the Week 24 endpoints, including the SRI-4.

Unblinded subject baseline characteristics (e.g., Joint Count, SLEDAI-2K) will be assessed. The randomization stratification factors will be re-evaluated in the expanded enrollment to achieve balance between treatment groups.

The IA will be performed by independent personnel (internal and/or external) not involved in the conduct of the study. Unblinded tables, listings, or graphs (TLGs) will be generated by independent personnel (internal or external) not involved in the conduct of the study.

The team that performs the unblinded interim analysis must not be involved in the study conduct between the interim DBL and the final DBL. They can however be part of the final analysis team after the final DBL has occurred. At the discretion of Biogen, unblinded TLGs and/or data may be shared with Sponsor personnel who are not directly involved in the conduct of the study. Additional information will be provided in an unblinding plan.

The endpoints and analysis to be included in the IA1 are listed in Appendix B.

10.2. Interim Analysis 2 (IA2): End of Treatment Periods

After the last subjects complete the Double-Blind Treatment Periods for Part A (at Week 24) and Part B (at Week 16), an interim database lock will occur and an IA of efficacy and safety data will be conducted. All Investigators, study site personnel, and subjects will remain blinded to treatment assignments until after the end of the study. Members of the Biogen study team not in direct contact with the study sites will have access to the study data after all subjects complete all Week 24 assessments in Part A and all Week 16 assessments in Part B. Members of the Biogen study team who become unblinded to the data for this IA will have no further contact with study sites until after the final database lock.

Unblinding details will be provided in the unblinding plan.

The endpoints and analysis to be included in the IA1 and IA2 are listed in Appendix B.

11. SAMPLE SIZE JUSTIFICATION

Part A:

The sample size for Part A is 100 subjects, randomized in a 1:1 ratio, with 50 subjects allocated to each treatment group. This sample size will provide approximately 71% power to detect a statistically significant difference in the Week 24 absolute change from baseline active joint count, assuming a standard deviation of 6, a maximal difference of BIIB059 over placebo of 2.5, a 20% dropout rate, and a 2-sided testing at the 0.2 level significance. If the study is not stopped for futility at the IA, then up to an additional 70 subjects may be enrolled into Part A. These may increase the power to approximately 87% at Week 24.

Part B:

The planned sample size for Part B is 100 subjects, randomized in a 1:1:1:1 ratio, with 25 subjects per dosage arm of BIIB059 (50 mg, 150 mg, and 450 mg) and placebo. This sample size will provide approximately 90% power to detect a dose-response relationship in the Week 16 percent change from baseline score in CLASI-A score, assuming a standard deviation of 30, a maximal difference of BIIB059 over placebo of 27.5%, and a 20% dropout rate. Five different

dose-response relationships will be tested at the 2-sided 10% significance level with the MCP-Mod method being used to control for multiplicity. Further details about the dose-response analyses can be found in Section 6.2. [redacted] The power calculations are based on assuming a standard deviation of 30, a maximal difference of BIIB059 over placebo of 27.5% at 450 mg, and a 20% dropout rate.

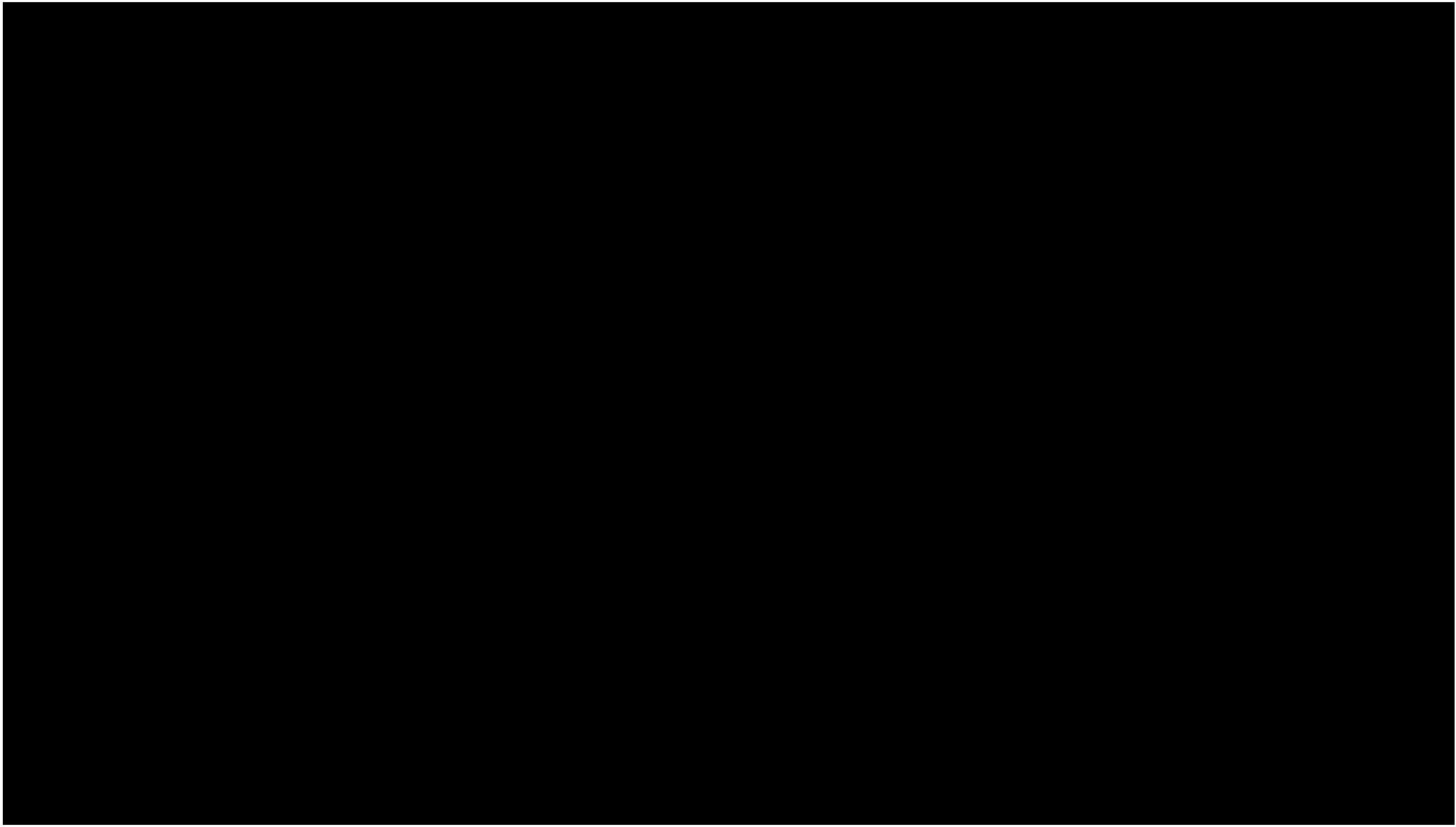
12. REFERENCES

- Ball EM, Bell AL. Lupus arthritis--do we have a clinically useful classification? Rheumatology (Oxford). 2012 May;51(5):771-9.
- Bretz, F; Pinheiro, JC; Branson, M. Combining multiple comparisons and modeling techniques in dose-response studies. Biometrics. 2005;61(3):738-48.
- Bonilla-Martinez, Z.; Albrecht, J.; Troxel, AB; Taylor, L.; Okawa, J; Dulay, S; Werth, VP. The Cutaneous Lupus Erythematosus Disease Area and Severity Index. Arch Dermatol. 2008; 144(2): 173-180.
- [redacted]
- Gladman, DD; Goldsmith, CH; Urowitz, MB; Bacon, P; Bombardier, C; Isenberg, D; Kalunian, K; Liang, MH; Maddison, P; Nived, O; et al. Sensitivity to change of 3 Systemic Lupus Erythematosus Disease Activity Indices: international validation. J Rheumatol. 1994 Aug;21(8):1468-71.
- Gladman, DD; Ibañez, D; Urowitz, MB. Systemic lupus erythematosus disease activity index 2000. J Rheumatol. 2002 Feb; 29(2):288-91.
- [redacted]
- Isenberg DA, Rahman A, Allen E, Farewell V, Akil M, Bruce IN, D'Cruz D, Griffiths B, Khamashta M, Maddison P, McHugh N, Snaith M, Teh LS, Yee CS, Zoma A, Gordon C. BILAG 2004. Development and initial validation of an updated version of the British Isles Lupus Assessment Group's disease activity index for patients with systemic lupus erythematosus. Rheumatology (Oxford). 2005 Jul;44(7):902-6.
- Klein R, Moghadam-Kia S, Taylor L, Coley C, Okawa J, LoMonico J, Chren MM, Werth VP. Quality of life in cutaneous lupus erythematosus. J Am Acad Dermatol. 2011 May;64(5):849-58.

- Little RJA, Rubin DB. Statistical Analysis with Missing Data. 1st ed. New York: John Wiley & Sons, Inc; 1987.
- LupusQoL instrument scoring, RWS Life Sciences. http://www.corptransinc.com/sites/lupusqol/instrument-information
- McElhone K, Abbott J, Shelmerdine J, Bruce IN, Ahmad Y, Gordon C, Peers K, Isenberg D, Ferenkeh-Koroma A, Griffiths B, Akil M, Maddison P, Teh LS. Development and validation of a disease-specific health-related quality of life measure, the LupusQol, for adults with systemic lupus erythematosus. Arthritis Rheum. 2007 Aug 15;57(6):972-9.
- Optum PRO CoRE smart measurement system manual, version 1.4. Optum, Inc., Johnston, RI; 2019 Mar.
- Rubin, D. B. Multiple Imputation for Nonresponse in Surveys. New York: John Wiley & Sons; 1987
- [redacted]
- Ware JE Jr. SF-36 health survey update. Spine (Phila Pa 1976). 2000 Dec 15;25(24):3130-9.
- Yee CS, Cresswell L, Farewell V, Rahman A, Teh LS, Griffiths B, Bruce IN, Ahmad Y, Prabu A, Akil M, McHugh N, D'Cruz D, Khamashta MA, Isenberg DA, Gordon C. Numerical scoring for the BILAG-2004 index. Rheumatology (Oxford). 2010 Sep;49(9):1665-9.

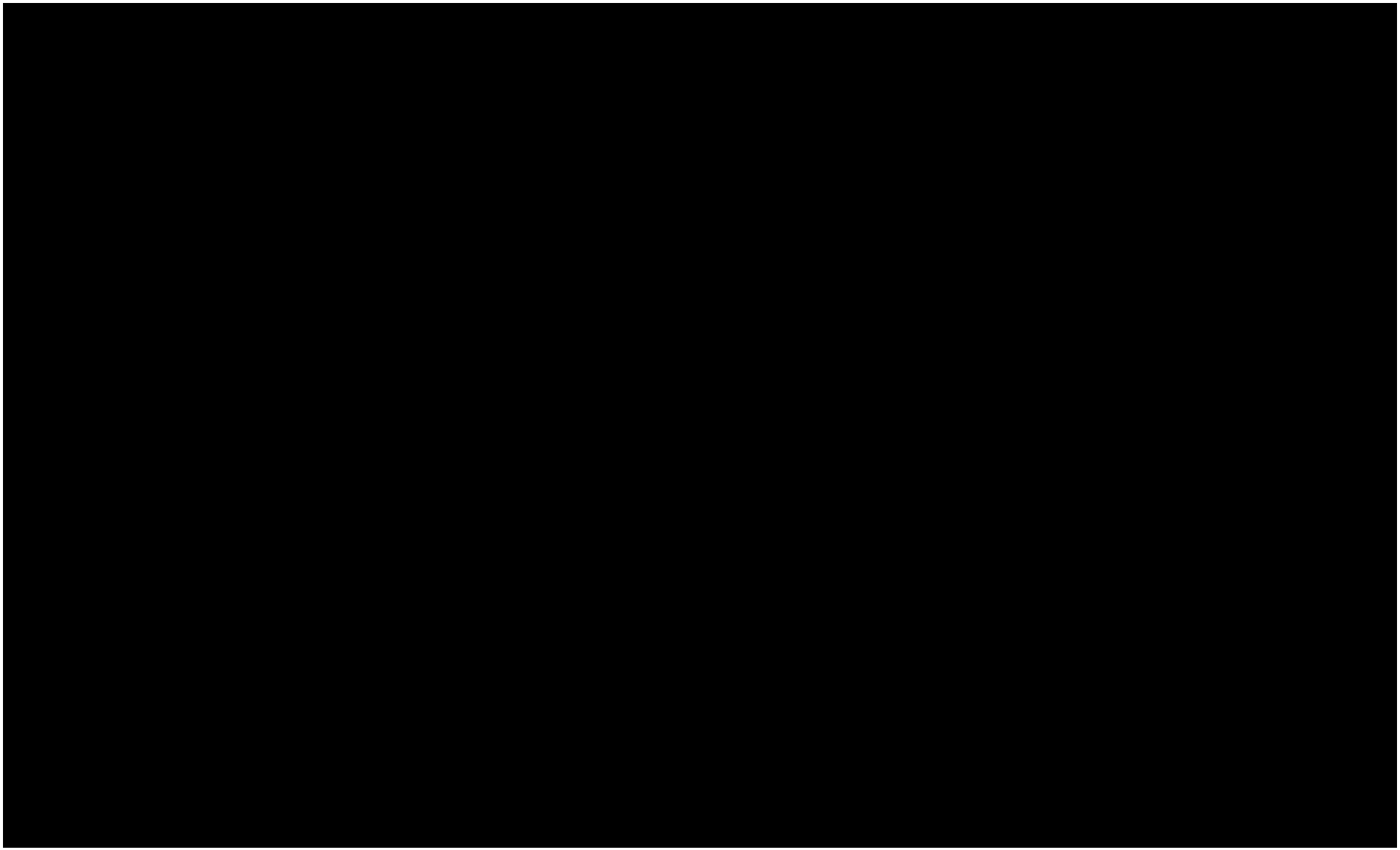
APPENDICES

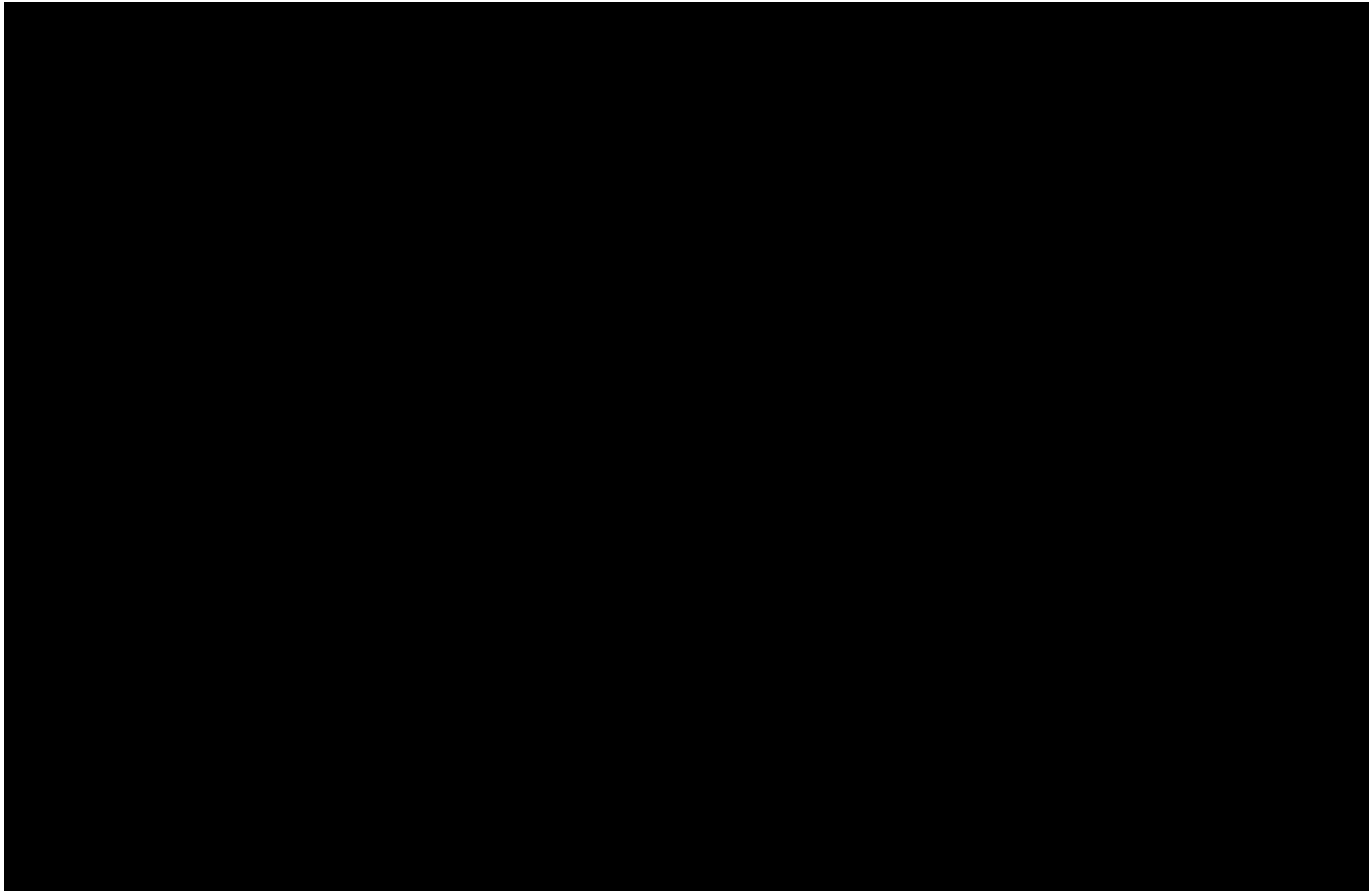
[redacted]

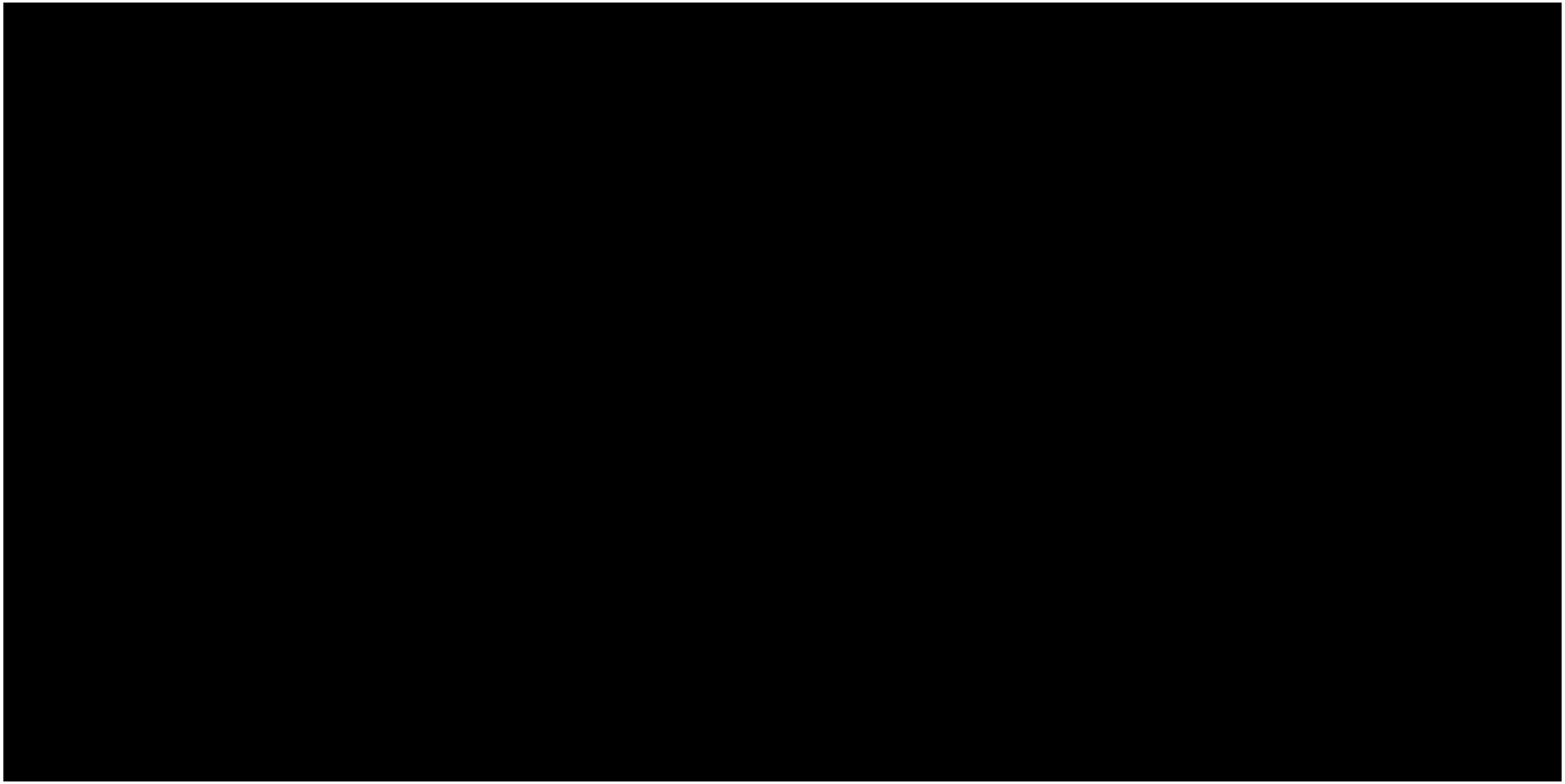


[redacted]

[redacted]







Appendix B: Summary of Additional Interim Analyses

	Section of the SAP	Note:	Study Part
Subject disposition	5.2	IA1and IA2, MITT population	Part A and B
Demography and baseline disease characteristics	5.3	IA1and IA2, MITT population	Part A and B
Previous medications used for the treatment of SLE and/or CLE	5.4	IA1and IA2	Part A and B
Treatment exposure	5.5	IA1and IA2	Part A and B
Concomitant therapies	5.6	IA1and IA2	Part A and B
Protocol Deviations	5.7	IA2: Summary of major protocol deviations	Part A and B
Primary endpoint: Primary analysis	6.2	IA1 and IA2: Part B: Model based and summary statistics over time Part A, IA1 Descriptive statistics by visit; IA2: Model based and summary statistics over time.	Part A and B
Secondary and additional efficacy endpoints	6.3, 6.4	Part A, IA1: Descriptive statistics by visit of the following endpoints- SRI-4, TJC, SJC, PGA of SLE (VAS), SLEDAI-2K, BILAG-2004, CLASI-A based endpoints, [redacted] Part A, IA2: Model based and summary statistics by visit of the following endpoints- SRI-4, TJC, SJC, PGA of	Part A

		SLE (VAS), SLEDAI-2K, proportion of subjects with no new organ system affected, CLASI-A based endpoints, BILAG based severe flare; Descriptive statistics by visit for the following endpoints- BILAG-2004 score, [redacted]	
Secondary and additional efficacy endpoints	6.3, 6.4	Part B, IA1: Model based and summary statistic over time- CLASI-A based endpoints; Descriptive statistics by visit of the following endpoints- SLEDAI-2K, [redacted], Physicians Global assessment of SLE (VAS), [redacted]. Part B, IA2: Model based and summary statistic over time - CLASI-A based endpoints, SLEDAI-2K; Descriptive statistics by visit of the following endpoints - [redacted], PGA of SLE (VAS), [redacted]	Part B
Immunoglobulins (IgG, IgA, IgM) change from baseline by visit	7.3.2	Part A and B, IA1: Descriptive statistics by visit.	Part A and B

[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
Safety analyses	7	IA2: Selection of adverse events analysis and laboratory analysis	Part A and B
PK concentrations and parameters	9	IA1 and IA2: PK concentration data needed for Pop PK analyses	Part A and B

Appendix C Efficacy Instruments

C.1 Cutaneous Lupus Erythematosus Disease Area and Severity Index (CLASI)

The CLASI instrument was developed to specifically evaluate lupus skin manifestations. It differentiates and separately scores disease activity (CLASI-A) [redacted]. The activity scale includes measurements of erythema, scale and hypertrophy, and mucous membrane disease, whereas the damage scale measures hyperpigmentation, atrophy, and scarring alopecia. Each part of the body is listed separately, from the scalp to the feet, in addition to sections focusing on mucous membrane involvement and alopecia. For the activity score, points are given for the presence of erythema, scale, mucous membrane lesions, recent hair loss, and inflammatory alopecia. For the damage score, points are given for the presence of dyspigmentation, scarring, and scarring alopecia. Total dyspigmentation scores are doubled when most of the dyspigmentation has been present for more than 1 year. Scores for each area are assigned based on the most severe lesion within the area of interest. Of note, affected body parts are weighted equally regardless of surface area and number of lesions present. Separate composite scores for activity and damage are calculated by simply summing the individual component scores. CLASI-A scores of 0-9, 10-20, and 21-70 represent disease severity of mild, moderate, and severe, respectively [Bonilla-Martinez 2008; Klein 2010].

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

The SLEDAI-2K is a reliable, valid, simple, 1-page activity index that measures disease activity and records feature of active lupus as present or not [Gladman 2002]. It is a modification of the SLEDAI to reflect persistent, active disease in those descriptors that had previously only considered new or recurrent occurrences. The SLEDAI-2K has been validated against the SLEDAI, which has been shown to be reliable at different levels of disease activity [Gladman 1994].

The SLEDAI-2K uses a weighted checklist to assign a numeric score based on the presence or absence of 24 symptoms at the time of assessment or during the previous 28 days. Each symptom present is assigned between 1 and up to 8 points based on its usual clinical importance, yielding a total score that ranges from 0 points (no symptoms) to 105 points (presence of all defined symptoms). However, if scored correctly, it is rare for even the sickest patients to score more than 20 points. The assessor is also requested to assess the subject's symptoms using the visual analog scale (VAS) for the Physician's Global Assessment (PGA).

C.3 British Isles Lupus Activity Group - 2004

The BILAG-2004 Disease Activity Index evaluates SLE activity in a number of organ systems, based on the principle of "physician's intention to treat" [Isenberg 2005]. The primary purpose of the BILAG in this study is to assess possible worsening in specific organ systems. Additional analyses of improvements in disease activity as assessed by the BILAG-2004 numeric scoring will also be performed [Yee 2010].

A separate alphabetic score is assigned to each organ system, corresponding in general to the following definitions:

- BILAG A: Severe disease activity requiring systemic high-dose oral corticosteroids, intravenous pulse corticosteroids, systemic immunosuppressants, therapeutic high-dose anticoagulation in the presence of high-dose corticosteroids, or prednisone ≥20 mg/day. Note that in the context of a clinical trial protocol with medication restrictions and blinded study treatment, the term "requiring" is not taken literally but indicates that if all else were equal this would be the degree of treatment indicated. It is also understood that some patients respond differently to levels of medication than do others, so that in assessing patients with the BILAG, "intent- to- treat" really means that most patients with this degree of symptom would require this level of treatment, not necessarily the patient in question.

- BILAG B: Moderate disease activity requiring treatment with systemic low-dose oral glucocorticoids, intramuscular or intra-articular or soft tissue corticosteroid injection, topical corticosteroid or immunosuppressants, or symptomatic therapy such as antimalarial medications or NSAIDs

- BILAG C: Mild disease

- BILAG D: System previously affected but now inactive

- BILAG E: System never involved

The BILAG-2004 is evaluated by scoring each item of a list of signs and symptoms as:

1. Improving

2. Same

3. Worse

4. New

0 Not present

(ND) Not done

For some items, appropriate responses may be:

- Y/N OR numeric values where indicated

- Y/N and confirm this is due to SLE activity

All signs and symptoms scored must be due to SLE. Use of a glossary provided with the BILAG-2004 instrument and training of assessors in use of the instrument are essential to obtaining reliable and consistent results. Use of the BILAG-2004 index for evaluating flares has been identified as a robust way of evaluating the efficacy of drugs; this judgment has been corroborated by external advisors and regulatory authorities.

BILAG assessments should be conducted by a trained evaluator, with documentation of training.

A copy of the BILAG-2004 and glossary is provided in the Study Reference Guide.

C.4 Tender and Swollen Joint Counts

Joint involvement is a frequent manifestation in patients affected by SLE and occurs in up to 90% of patients at the onset or during the course of the disease [Ball and Bell 2012]. An active joint is defined as a joint with pain and signs of inflammation (e.g., tenderness, swelling or effusion). A 28-joint assessment will be performed. The joints to be assessed bilaterally include the shoulders, elbows, wrists (radiocarpal, carpal, and carpometacarpal bones are considered as a single unit), metacarpophalangeal (MCP) joints (MCP 1, 2, 3, 4, and 5), proximal interphalangeal (PIP) joints (PIP 1, 2, 3, 4 and 5), and the knees. Artificial and ankylosed or fused joints will be excluded from tenderness and swelling assessments.

C.5 Physician Global Assessment of SLE (VAS)

The PGA is used to quantify disease activity and is measured using an anchored VAS. The PGA asks the Investigator to assess the subject's current disease activity from a score of 0 (none) to 3 (severe), with the assessment made relative not to the subject's most severe state but the most severe state of SLE that might exist at study visits.

A cop y of the PGA of SLE is provided in the Study Reference Guide.

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

- [redacted]
- [redacted]

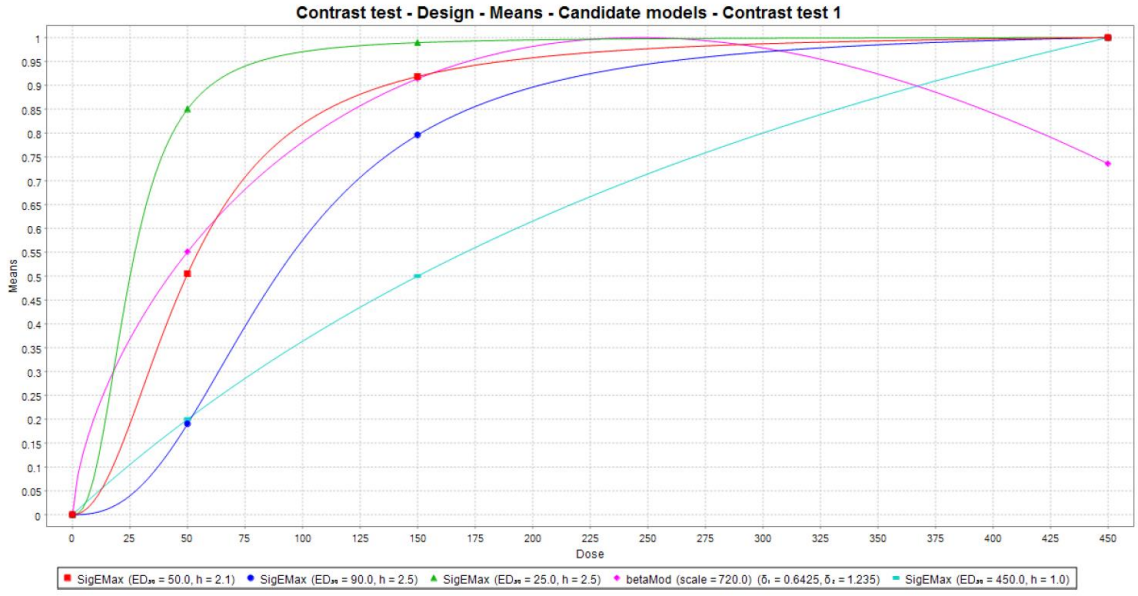
- [redacted]

- [redacted]

- [redacted]

[redacted]

Appendix D MCP-MOD candidate models



Appendix E Oral Corticosteroid Tapering Schedule

Timing by Weeks (Days)	Allowable OCS Dosing (mg)
Week 1-4 (Day 1 to 28)	Stable; 1 OCS rescue for SLE activity allowed up to maximum 10mg/day above Day 1 dosage, but must return to Day 1 level within 7 days
Week 5 (Day 29 – 35)	17.5
Week 6 (Day 36 – 43)	17.5
Week 7 (Day 44 – 50)	15
Week 8 (Days 51 to 57)	15
Week 9 (Days 58 – 64)	12.5
Week 10 (Days 65 to 71)	12.5
Week 11 (Days 72 to 78)	10
Week 12 (Days 79 to 85)	10
Week 13-16 (Days 86 to 113)	10
Week 17-24 (Days 114 to 169)	10
Note 1: Two rescues are allowed out of the three described. Note 2: Between Weeks 1 and 4 (Rescue 1). Once a dose is reduced it must not be re-increased, with the exception of the 1-time rescue. Note 3: Between Weeks 5 and 7 (Rescue 2) and between Weeks 9 and 11 (Rescue 3), the dose can be increased to the previous dose level for up to 1 week before restarting the tapering schedule.	