Statistical Analysis Plan

Study M19-130

Systemic Lupus Erythematosus: A Phase 2 Study to Investigate the Safety and Efficacy of Elsubrutinib and Upadacitinib Given Alone or in Combination (ABBV-599 Combination) in Subjects with Moderately to Severely Active Systemic Lupus Erythematosus

Date: 26 July 2022

Version 4.0

Table of Contents

List of Tables

List of Figures

List of Appendices

1.0 Introduction

This statistical analysis plan (SAP) describes the proposed statistical analyses to be completed by the Data and Statistical Sciences Department for ABBV-599 Study M19-130 Primary Analysis performed after all subjects have completed the Week 24 visit or have discontinued the study and the Final Analysis performed after all subjects have completed or discontinued the study.

Pharmacokinetic and Biomarker analyses will be performed separately and reported in a separate document, as applicable, to further aid data interpretation.

The SAP will be updated if there are any protocol changes identified that impact the efficacy or safety analysis.

Unless noted otherwise, all analyses will be performed using SAS version 9.4 or later (SAS Institute Inc., Cary, NC 27513) under the UNIX operating system.

This SAP Version 3 includes changes to the analyses described in the amended RCT Study M19-130 (Version 7) protocol. Details are outlined in Section 13.0.

2.0 Study Design and Objectives

2.1 Objectives, Hypotheses and Estimands

The main objective is to evaluate the safety and efficacy of elsubrutinib, upadacitinib and the ABBV-599 (elsubrutinib/upadacitinib) combination vs. placebo for the treatment of signs and symptoms of Systemic Lupus Erythematosus (SLE) at 24 and 48 weeks in subjects with moderately to severely active SLE and to define dose(s) for further development.

Primary Efficacy Objective

The primary efficacy objective is to demonstrate a higher rate of achievement of SRI-4 and steroid dose ≤ 10 mg prednisone equivalent once a day (QD) at Week 24 of active

treatment groups (elsubrutinib, upadacitinib and ABBV-599 combination) when compared to placebo based on Full Analysis Set (FAS) as defined in Section 4.0.

Hypothesis corresponding to the primary efficacy objective and endpoint is:

- The proportion of subjects achieving SRI-4 and steroid dose ≤ 10 mg prednisone equivalent QD treated with active treatment groups (elsubrutinib, upadacitinib and ABBV-599 combination) is greater than those treated with placebo at Week 24.

The estimand corresponding to the primary efficacy objective is defined as follows:

- Difference in the proportion of subjects achieving SRI-4 and steroid dose ≤ 10 mg prednisone equivalent QD regardless of premature discontinuation of study drug and without initiation of rescue medications (see Section 8.2) between active treatment groups (elsubrutinib, upadacitinib and ABBV-599 combination) and placebo in FAS.

Secondary Efficacy Objective

The secondary efficacy objective is to demonstrate higher efficacy of active treatment groups (elsubrutinib, upadacitinib and ABBV-599 combination) when compared to placebo with respect to the secondary endpoints as specified in Section 3.2. The secondary efficacy objectives will be assessed based on FAS.

Hypotheses corresponding to the secondary efficacy objectives and endpoints are:

- For each binary secondary endpoint, a greater proportion of subjects with improvement for the endpoint is achieved in active treatment groups (elsubrutinib, upadacitinib and ABBV-599 combination) when compared to that of placebo.
- For each continuous secondary endpoint, greater mean change from baseline for the endpoint is achieved with active treatment groups (elsubrutinib, upadacitinib and ABBV-599 combination) when compared to that of placebo.

The estimands corresponding to the secondary efficacy objectives are defined as follows:

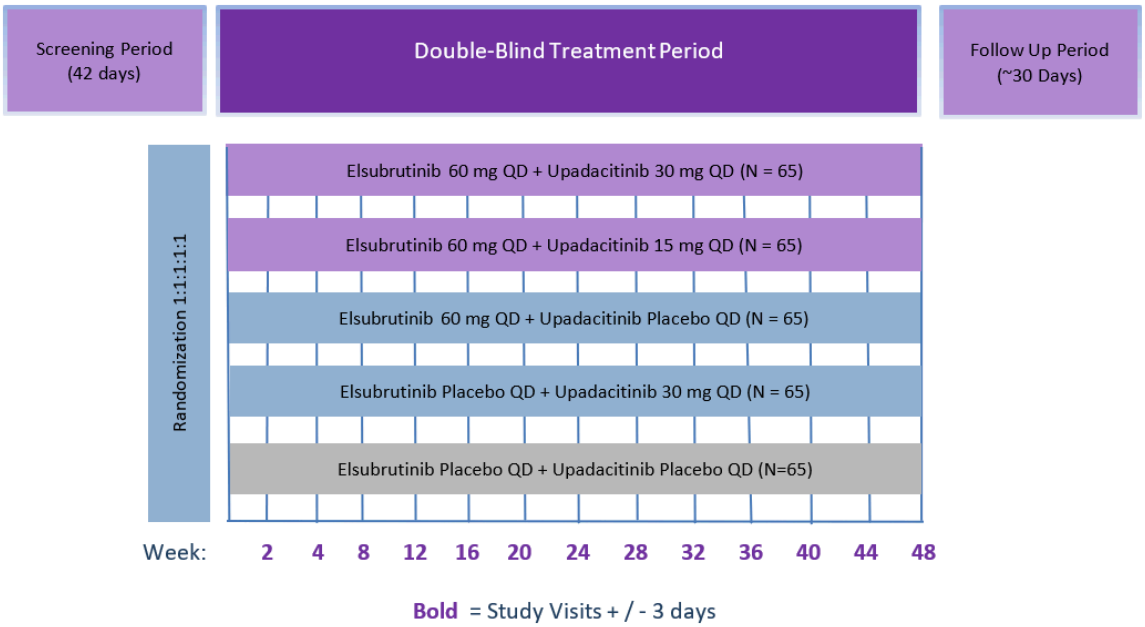
- For each binary secondary endpoint, the estimand is the difference in the proportion of subjects achieving the endpoint regardless of premature discontinuation of study drug and without initiation of rescue medications (see Section 8.2) between active treatment groups (elsubrutinib, upadacitinib and ABBV-599 combination) and placebo in FAS.
- For each continuous secondary endpoint, the estimand is the difference in the mean change from baseline of the endpoint regardless of premature discontinuation of study drug and if subjects would not initiate rescue medications (See Section 8.2) between active treatment groups (elsubrutinib, upadacitinib and ABBV-599 combination) and placebo in FAS.

2.2 Study Design Overview

This is a multi-center, randomized, double-blind, and placebo-controlled Phase 2 study, with a Week 24 primary endpoint, to investigate the safety and efficacy of elsubrutinib and upadacitinib given alone or in combination (ABBV-599 [elsubrutinib/upadacitinib] combination) in subjects with moderately to severely active systemic lupus erythematosus.

The schematic of the study is shown in Figure 1.

Figure 1. Study Schematic



QD = once daily

2.3 Treatment Assignment and Blinding

Approximately 325 male and female subjects, 18 to 65 years of age, inclusive, with a diagnosis of moderately to severely active SLE (clinical diagnosis of SLE at least 24 weeks prior to Screening, meeting at least 4 of the 11 revised Criteria for Classification of SLE according to the 1997 Update of the 1982 ACR OR meeting at least 4 of the 2012 Systemic Lupus Erythematosus International Collaborating Clinics [SLICC] classification criteria, including at least 1 clinical criterion and 1 immunologic criterion) despite standard of care therapy will be randomized in a 1:1:1:1:1 ratio of 5 treatment groups, as follows:

- Group 1: Elsubrutinib 60 mg QD and upadacitinib 30 mg QD (n = 65)
- Group 2: Elsubrutinib 60 mg QD and upadacitinib 15 mg QD (n = 65)
- Group 3: Elsubrutinib 60 mg QD and upadacitinib placebo QD (n = 65)

- Group 4: Elsubrutinib placebo QD and upadacitinib 30 mg QD (n = 65)
- Group 5: Elsubrutinib placebo QD and upadacitinib placebo QD (n = 65)

Randomization will be stratified based on the following factors:

- Baseline corticosteroid dose above 10 mg prednisone-equivalent (≤ 10 mg or > 10 mg)
- Screening Systemic Lupus Erythematosus Disease Activity Measure (SLEDAI)-2K (< 10 or ≥ 10)
- High versus low interferon score versus not applicable (approximately 80% high)
  - This stratification factor does not apply to subjects in China.
- Baseline immunosuppressant (azathioprine, tacrolimus, cyclosporine, methotrexate [MTX], mycophenolate) (yes/no)

Once approximately 20% of total subjects have been randomized who are interferon signature low, further screening of such subjects may be suspended. Interferon stratification will be utilized in all participating countries/regions except China due to local constraints.

The study duration will include a 42-day maximum screening period and a 48-week randomized, placebo-controlled, double-blind, parallel-group treatment period with a 30-day follow-up phone call. An interim analysis was performed after the first 50% of the subjects completed Week 24 visit or withdrew from the study by an independent team at AbbVie as planned (Section 11.2). The Week 24 Primary Analysis will be performed when all subjects have completed the Week 24 visit or have discontinued the study. The study team will remain blinded until the Week 24 Primary Analysis. The sites and subjects will remain blinded throughout the remainder of the study. A 30-day follow-up phone call from the last dose of study drug should occur to determine the status of any ongoing AEs/SAEs or the occurrence of any new AEs/SAEs.

2.4 Sample Size Determination

The planned sample size for this study is 325 subjects in total, with 65 subjects for each of the 5 arms, determined based on the primary endpoint, the composite of SLE Responder Index (SRI)-4 (See Section 3.1) and steroid dose ≤ 10 mg prednisone equivalent QD at Week 24. Assuming at Week 24 the placebo response rate is 25%, a two-sided Chi-squared Test with alpha=10% will provide 78% to 91% power to detect a difference of 20% to 25% between the active treatment groups and the placebo group.

3.0 Endpoints

3.1 Primary Endpoint(s)

The following is the primary efficacy endpoint:

- Achievement of SRI-4 and steroid dose ≤ 10 mg prednisone equivalent once a day (QD) at Week 24

SLE Responder Index (SRI)-4 is defined as follows with all criteria compared to baseline:

- ≥ 4-point reduction in Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K) score.
- No worsening of the overall condition (< 0.3 point increase in Physician's Global Assessment [PhGA])
- No new BILAG A or more than 1 new BILAG B disease activity scores (i.e., no organ system changes from baseline B/C/D/E to A and no more than 1 organ system changes from baseline C/D/E to B).

3.2 Secondary Endpoint(s)

The following secondary efficacy endpoints will be evaluated:

1. Achievement of SRI-4 at Week 24.

2. Achievement of BILAG-Based Combined Lupus Assessment (BICLA) response at Week 24.
3. Achievement of Lupus Low Disease Activity State (LLDAS) at Week 24
4. Steroid burden, assessed as change from Baseline at Week 24
5. Number of mild, moderate, or severe flares per patient-year (respectively and overall) by Safety of Estrogens in Lupus Erythematosus National Assessment (SELENA) SLEDAI Flare Index (SFI), assessed by number and types of flare (mild/moderate, severe, and any) per subject through Week 24

3.3 Other Efficacy Endpoint(s)

The following additional efficacy endpoints will be evaluated at all visits unless otherwise noted:

1. Achievement of SRI-4.
2. Achievement of SRI-5, -6, -7, -8 and steroid dose ≤ 10 mg prednisone equivalent QD at Weeks 24 and 48.
3. Achievement of SRI-7/-8 among subjects who have baseline SLEDAI-2K score >=7/>=8 and steroid dose ≤ 10 mg prednisone equivalent QD at Weeks 24 and 48.
4. Achievement of SRI-5, -6, -7, -8.
5. Achievement of SRI-7/-8 among subjects who have baseline SLEDAI-2K score >=7/>=8.
6. Achievement of BICLA response.
7. Achievement of LLDAS.
8. Achievement of ≥ 4-point decrease in SLEDAI-2K from Baseline.
9. Steroid burden, assessed as change from Baseline.

10. Number of mild, moderate, or severe flares per patient-year (respectively and overall) by SELENA SFI, assessed by number and types of flare per subject through Week 24 and Week 48.
11. Time to first flare (any severity) by SELENA SFI after first study drug administration up to Week 24 and Week 48.
12. Achievement of 50% reduction of tender or swollen lupus joints defined as ≥ 50% decrease in either tender or swollen joints (among those starting with total ≥ 6 affected joints).
13. Achievement of 50% reduction in Cutaneous Lupus Erythematosus Disease Area and Severity Index (CLASI) activity score (of those starting with CLASI ≥ 10).
14. Change in SLEDAI-2K from Baseline.
15. Change in BILAG total score from Baseline.
16. Change in PhGA from Baseline.
17. Change in Patient Global Assessment (PtGA) from Baseline.
18. Change from Baseline in Functional Assessment of Chronic Illness Therapy (FACIT)-Fatigue at Weeks 2, 12, 24, and 48.
19. Change from Baseline in 36-Item Short Form Health Survey (SF-36) at Weeks 2, 12, 24, and 48.
20. Change from Baseline Lupus Quality of Life questionnaire (LupusQoL) at Weeks 2, 12, 24, and 48.
21. Change from Baseline Pain Numerical Rating Scale (NRS) at Weeks 2, 12, 24, and 48.
22. Achievement of SRI-4 and steroid dose ≤ 10 mg prednisone equivalent QD.
23. The annualized BILAG-based flare rate through Week 24 and through Week 48.

24. Achievement of ≥ 50% improvement in joint counts (defined as ≥ 50% decrease in the swollen and tender joint count from baseline in patients with ≥ 8 swollen and ≥ 8 tender joints at baseline).
25. Achievement of Clinical SLEDAI (defined as ≥ 4-point reduction in clinical components (no laboratory components) of the SLEDAI) Among subjects with baseline SLEDAI-2K clinical score ≥ 4.
26. Achievement of prednisone equivalent steroid daily dose ≤ 7.5 mg among subjects with daily dose ≥10 mg at Baseline through Week 48.
27. Achievement of prednisone equivalent steroid daily dose ≤ 7.5 mg sustained from Week 36 to Week 48 among subjects with daily dose ≥ 10 mg at Baseline.
28. Achievement of no increase in SLEDAI-2K total score from baseline.
29. Achievement of < 0.3 point increase in PhGA from baseline.
30. Achievement of no new BILAG A or more than 1 new BILAG B comparing to baseline (as defined in SRI-4).
31. Achievement of improvement in all baseline BILAG A and BILAG B (as defined in BICLA)

3.4 Safety Endpoint(s)

Safety evaluations include adverse event (AE) monitoring, physical examinations, vital sign measurements, electrocardiograms (ECGs), and clinical laboratory testing (hematology, chemistry, and urinalysis) as measures of safety and tolerability for the entire study duration.

An independent external Cardiovascular Adjudication Committee (CAC) will adjudicate all blinded cardiovascular and cerebrovascular events, embolic/thrombotic events and deaths, as defined in the CAC charter.

4.0 Analysis Populations

The Full Analysis Set (FAS) includes all randomized subjects who have received at least 1 dose of study drug. Subjects will be grouped according to treatment as randomized. The FAS will be used for all efficacy analyses and summary of baseline characteristics.

The Safety Analysis Set consists of all subjects who have received at least 1 dose of study drug. For the Safety Analysis Set, subjects are assigned to a treatment group based on the treatment actually received, regardless of the treatment randomized.

5.0 Subject Disposition

The total number of subjects who were screened, enrolled (randomized), and treated will be summarized.

A summary of subject accountability will be provided where the number of subjects in each of the following categories will be summarized for each treatment group:

- Subjects enrolled (randomized) in the study.
- Subjects who took at least one dose of study drug.
- Subjects who completed study drug up to Week 24.
- Subjects who completed study drug up to Week 48.
- Subjects who prematurely discontinued study drug (all reasons and primary reason).
- Subjects who completed Week 24 visits.
- Subjects who completed the study.
- Number of subjects who elect to participate in the long-term extension Study M20-186.

For end of study participation, the number and percentage of subjects who completed the protocol-defined follow-up period (or did not with associated reasons) will be summarized overall and by treatment group.

6.0 Study Drug Duration and Compliance

6.1 Study Drug Duration

The duration of exposure to study drug will be summarized by each treatment group using the Safety Analysis Set population.

- Duration of exposure = the last study drug dose date – the first study drug dose date + 1.

For the Safety Analysis Set, duration of treatment will be summarized for each treatment group and for all investigational study drug dose groups combined. The duration of treatment is defined for each subject as last dose date minus first dose date + 1. Duration of treatment will be summarized using the number of subjects, mean, standard deviation, minimum, median and maximum for duration of Study M19-130. In addition, the number and percentage of subjects exposed to study drug will be summarized for the following exclusive duration intervals:

- ≥ 1 day
- ≥ 29 days
- ≥ 57 days
- ≥ 85 days
- ≥ 113 days
- ≥ 141 days
- ≥ 169 days
- ≥ 197 days
- ≥ 225 days
- ≥ 253 days
- ≥ 281 days
- ≥ 309 days
- ≥ 337 days

6.2 Study Drug Compliance

Study drug compliance will be summarized for each treatment group for SA population. Descriptive statistics will be provided including sample size, mean, standard deviation, median, minimum, and maximum.

Treatment compliance (TC) will be calculated using the following formula for tablets and capsules respectively:

$$TC = \frac{Total\ number\ of\ tablets/\text{capsules}\ taken}{(last\ dose\ date - first\ dose\ date + 1) \times < number\ of\ tablets/\text{capsules}\ per\ day >} \times 100\%$$

- The total number of tablets/capsules taken will be calculated as the total number of tablets/capsules dispensed minus the total number of tablets/capsules returned.

  Note: If the bottle is not returned, total number of tablets/capsules returned is assumed to be 0.
- In M19-130: number of tablets per day is 1 and number of capsules per day is 3.

7.0 Demographics, Baseline Characteristics, Medical History, and Prior/Concomitant Medications

Demographics, baseline or disease characteristics, medical history, and prior and concomitant medications will be summarized for the FAS overall and by treatment group. Categorical variables will be summarized with the number and percentage of subjects; percentages will be calculated based on the number of non-missing observations. Continuous variables will be summarized with descriptive statistics (number of non-missing observations, mean and standard deviation, median, minimum and maximum).

7.1 Demographics and Baseline Characteristics

The following demographic and baseline characteristics, as measured at baseline of the study, will be summarized.

Demographic Characteristics

- Sex (male, female)
- Age (years)
- Age category [18 – < 40 years old, 40 – < 65 years old, ≥ 65 years old]
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- Race (White, Black or African American, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Asian, Other)
- Geographic region (North America, South/Central America, Western Europe, Eastern Europe, Asia Japan, Asia Other, Other)
- Weight (kg)
- Weight Categories (< 60 kg, ≥ 60 kg)
- Height (cm)
- Body Mass Index (BMI) (kg/m^2)
- BMI Category (kg/m^2) (BMI < 25, BMI ≥ 25)

Other Baseline Characteristics

The following continuous variables will be summarized:

- SLEDAI-2K score
- Patient's Global Assessment of Disease Activity score
- Physician's Global Assessment of Disease Activity score
- LupusQoL score
- SF-36 score
- Time since onset of lupus (years), which will be calculated as follows:

Time since onset of lupus (years) = year of first dose – year of initial diagnosis of SLE

- Number of tender joints (from 28-tender joint count)
- Number of swollen joints (from 28-swollen joint count)
- Daily prednisone equivalent dose of corticosteroid (mg/day)
- CLASI Total Activity Score
- CLASI Total Damage Score
- Complement C3 and C4 levels
- Anti-dsDNA level
- Urine protein to urine creatinine ratio
- Serum immunoglobulin (Ig)G and IgM concentrations
- Serum creatinine

The following categorical variables will be summarized:

- Anti-dsDNA status (positive defined as >= 30 IU/mL)
- SLEDAI-2K status (< 10 or ≥ 10)
- British Isles Lupus Assessment Group (BILAG) organ system involvement at baseline (yes or no for each organ system domain). Involvement requires a baseline BILAG disease activity score of A or B.
- BILAG A organ system involvement at baseline (yes or no for each organ system domain).
- SLEDAI-2K organ system involvement at baseline (yes or no for each organ system domain).
- C3 status (less than the lower limit of normal [LLN]) (< 90.0 mg/dL)
- C4 status (less than LLN [< 10.0 mg/dL])
- Anti-Smith antibodies (≥ 30 AU/mL) (yes or no)
- Corticosteroid use (yes or no), and within those taking corticosteroids, ≤ 10 mg/day or > 10 mg/day.
- Corticosteroid use (yes or no), and within those taking corticosteroids, ≤ 7.5 mg/day or > 7.5 mg/day.

- Immunosuppressant use (yes or no):
  - Mycophenolate use (yes or no)
  - Azathioprine use (yes or no)
  - Methotrexate use (yes or no)
  - Tacrolimus use (yes or no)
  - Cyclosporine use (yes or no)
  - Leflunomide use (yes or no)
- Antimalarial use (yes or no):
  - Hydroxychloroquine (yes or no)
  - Chloroquine (yes or no)
  - Quinacrine (yes or no)
- Clinical Tests at Screening:
  - Chest x-ray
  - Tuberculin PPD skin test, QuantiFERON TB Gold test
  - Serum pregnancy test
- Immunization History
  - Herpes Zoster immunization
  - Hepatitis B immunization
- Tobacco/Nicotine and Alcohol Use
  - Tobacco/Nicotine Use (current, former, never, or unknown)
  - Alcohol Use (current, former, never, or unknown)

7.2 Medical History

Medical history data will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). The actual version of the MedDRA coding dictionary will be noted in the statistical tables and clinical study report. Medical history data will be summarized and presented for the FAS population using body systems and conditions/diagnoses as captured on the CRF. The body systems will be presented in alphabetical order and the conditions/diagnoses will be presented in alphabetical order within each body system.

The number and percentage of subjects with a condition/diagnosis will be summarized for each treatment group as well as overall. Subjects reporting more than one condition/diagnosis within a body system will be counted only once for that body system. No statistical comparison will be performed for medical history reporting.

7.3 Prior and Concomitant Medications

Prior and concomitant medications will be summarized by each treatment group as well as overall for the FAS. Prior medications are those medications taken prior to the first dose of study drug. These include medications with a start date before the first study drug administration date, regardless of the end date of these medications. Medications taken on the day of the first dose of study drug are not counted as prior medications. Concomitant medications are those medications, other than study drug, taken after the first dose of study drug and within 1 day of the last dose of study drug. This includes medications with a start date between first study drug administration and last study drug administration + 1 day, as well as medications with a start date prior to first dose of study drug and which are ongoing after first dose of study drug. Medications taken on the day of the first dose of study drug are counted as concomitant medications.

The number and percentage of subjects who received a prior medication and the number and percentage of subjects who received a concomitant medication will be tabulated separately by the generic name assigned by the most current version of the World Health Organization (WHO) Drug Dictionary.

8.0 Efficacy Analyses

8.1 General Considerations

All efficacy analysis will be performed in the FAS.

AbbVie may terminate this study prematurely at any time, either in its entirety or partially (discontinue 1 or more treatment groups based on interim analysis), at any site. In the

event the study is partially terminated, efficacy analysis will be performed separately for the terminated treatment groups and for the continued treatment groups:

Terminated Treatment Groups

For terminated group(s), the descriptive statistics will be provided by treatment group for the following efficacy endpoints using observed case (OC) (defined in Section 8.3) through Week 48:

- Achievement of SRI-4 and steroid dose ≤ 10 mg prednisone equivalent QD.
- Achievement of SRI-4.
- Achievement of BICLA.
- Number of mild, moderate, or severe flares per patient-year (respectively and overall) by SELENA SFI, assessed by number and types of flare per subject through Week 24 and Week 48.
- Steroid burden, assessed as change from Baseline.

The statistics include number of observations, mean, standard deviation, minimum, median, and maximum for continuous variables, and number and percentage of subjects for categorical variables.

Continued Treatment Groups

For continued treatment groups, the primary and secondary efficacy endpoints will be analyzed at the alpha level of 0.1 (two-sided). Multiplicity control is not planned for this study.

For all efficacy endpoints, the descriptive statistics will be provided by treatment group. The statistics include number of observations, mean, standard deviation, minimum, median, and maximum for continuous variables, and number and percentage of subjects for categorical variables.

Unless otherwise specified, any subject who is randomized based on a wrong stratum will be analyzed according to the actual stratum that the subject belongs to.

"Baseline" refers to the last non-missing observation before the first administration of study drug, or randomization if no study drug is given.

8.2 Handling of Potential Intercurrent Events

Potential intercurrent events considered in this study include 1) premature discontinuation of the study drug; 2) initiation of rescue medications; 3) initiation or dose escalation of SLE-related corticosteroid. Intercurrent events will be handled using the following methods for the efficacy analysis:

8.2.1 Premature Discontinuation of Study Drug

Data collected will be used regardless of premature discontinuation of the study drug.

8.2.2 Initiation of Rescue Medications or Initiation or Dose Escalation of SLE-Related Corticosteroid

Data collected on or after the occurrence of the rescue medications or initiation or dose escalation of SLE-related corticosteroid will be handled using the methods below for the efficacy analysis:

1. [redacted]
2. [redacted]
3. [redacted]

4. [redacted]
5. [redacted]
6. [redacted]

Intercurrent events above apply to all categorical endpoints. For continuous endpoints, data after intercurrent events above will be excluded from modeling (i.e., predicted from the model for analysis).

8.3 Handling of Missing Data

Missing data could occur due to various reasons, including missing visits/assessments, early withdrawal from the study, or missing due to Coronavirus disease 2019 (COVID-19) infection or logistic restriction.

The COVID-19 pandemic is interfering with the conduct of many ongoing trials, with potential impacts on treatment duration and the collection, analysis and the interpretation of clinical trial data. Some protocol-specified visits in the clinical trials may be impacted due to COVID-19 infection or logistical restrictions during the pandemic. For example, some scheduled visits may be missed due to self-quarantine or local government restrictions on travel; some visits may also be delayed or canceled due to healthcare resource constraints during the pandemic. Impacted visits due to COVID-19 will be recorded in the database. The probability of having missed visits and missing data due to COVID-19 infection or logistical restrictions related to the COVID-19 pandemic can be reasonably assumed to be unrelated to the unobserved values. Therefore, for the purpose of statistical analysis, it is reasonable to assume that these missing data are missing at random (MAR) and the statistical models that require MAR assumption are appropriate. The intent is to provide reliable estimates of the treatment effects targeted in the protocol

under the scenario without the impact of COVID-19 pandemic. Number of subjects with missing values due to COVID-19 will be presented.

Missing data and intercurrent events will be handled using the following methods for the efficacy analysis.

Non-Responder Imputation (NRI) incorporating Multiple Imputation (MI) to handle missing data due to COVID-19 (NRI-C)

- The NRI-C will categorize any subject who does not have an evaluation during a pre-specified visit window (either due to missing assessment or due to early withdrawal from the study) as a non-responder for the visit. There are two exceptions which will be handled in the order below:
  1) when a subject is a responder at both before and after the visit window, the subject will be categorized as a responder for the visit.
  2) missing data due to COVID-19 infection or logistical restriction and not satisfying 1) will be handled by MI.

Of note, subjects will be counted as non-responders, thereafter, and will not be imputed using 1) or 2), on or after the intercurrent event as specified in Section 8.2.2.

- Details on MI method are described below:

  MI will be performed using observed data excluding data on or after intercurrent events specified in Section 8.2.2. Markov Chain Monte Carlo (MCMC) will be first applied to augment data into monotonic missing pattern, where applicable, and PROC MI will be used to generate 30 datasets using the regression method. If the binary endpoints are derived from the continuous variables, PROC MI will be applied to the continuous variables. The variables to be included in the imputation model are: treatment group, randomization stratification factors (baseline corticosteroid dose above 10 mg prednisone-equivalent (≤ 10 mg or > 10 mg), screening SLEDAI-2K (< 10 or ≥ 10), baseline interferon score (high or low or not applicable), baseline immunosuppressant (yes or no)), gender, pooled race (White, Black, Asian, or Other), ethnicity, pooled geographic region (North America/South America,

Europe, or Asian), age, baseline BMI, and if applicable, baseline measurement and post-baseline measurements at each visit up to the end of the analysis period. The random seed for MCMC and the random seed for PROC MI are specified in Appendix E. The imputed post-baseline measurements will be rounded to the same precision as the observed data before the determination of responder status. Prior to applying CMH model, subjects will be characterized as responders or non-responders based on MI imputed values if missing due to Covid-19, otherwise subjects will be considered as non-responder for missing due to other reasons in the NRI-C approach. Using the CMH model adjusted by randomization stratification factors, the endpoints will be analyzed using each of the 30 imputed datasets. SAS PROC MIANALYZE will be used to generate the final inferences of the risk difference using Rubin's rule. The rules of handling intercurrent events, as specified in Section 8.2.2, will be applied before and after applying MI for efficacy assessments.

Mixed-Effect Model Repeated Measurements (MMRM)

MMRM analysis will be conducted using a mixed-effect model including observed measurements at all visits for continuous variables. Data collected on or after initiation of rescue treatment or after initiation or dose escalation of SLE-related corticosteroid as defined in Section 8.2.2 will not be included. The mixed model includes the categorical fixed effects of treatment, visit and treatment-by-visit interaction, randomization stratification factors, and the continuous fixed covariate of baseline measurement. An unstructured variance covariance matrix will be used. The parameter estimations are based on the assumption of data being missing at random and using the method of restricted maximum likelihood (REML).

As Observed (AO)

AO analysis will not impute values for missing evaluations, and thus a subject who does not have an evaluation on a scheduled visit will be excluded from the AO analysis for that visit. Regardless of premature discontinuation of study drug or initiation of rescue treatment or initiation or dose escalation of SLE-related corticosteroid as defined in

Section 8.2.2, all observed data will be used in the analysis. As observed analysis may be used for both categorical variables and continuous variables for sensitivity analysis.

Multiple Imputation (MI)

As a sensitivity analysis, MI will be utilized for the primary endpoint. Data collected on or after the occurrence of intercurrent events defined in Section 8.2.2 will be considered as non-responder. When a subject is a responder at both before and after the visit window, the subject will be categorized as a responder for the visit. Otherwise, subjects who have no available measurements will be handled by multiple imputations. Details on MI can be found in NRI-C approach.

Observed Case (OC)

The OC analysis will not impute values for missing evaluations and, thus, a subject who does not have an evaluation on a scheduled visit will be excluded from OC analysis for that visit. In addition, the OC analysis will not use values on or after the occurrence of intercurrent events defined in Section 8.2.2.

Censoring

For the time-to-event endpoint, i.e., time to flare, missing data will be considered as censored at the last available time for the information.

8.4 Primary Efficacy Endpoint and Analyses

8.4.1 Primary Efficacy Endpoint

The primary efficacy endpoint is the composite of SRI-4 and steroid dose ≤ 10 mg prednisone equivalent QD at Week 24.

8.4.2 Main Analysis of Primary Efficacy Endpoint(s)

The attributes of the estimand corresponding to the primary efficacy endpoint are summarized in Table 1.

Table 1. Summary of the Estimand Attributes of the Primary Efficacy Endpoint(s)

	Attributes of the Estimand				
Estimand Label	**Treatment**	**Endpoint**	**Population**	**Handling of Intercurrent Events**	**Statistical Summary**
Primary	• Elsubrutinib 60 mg QD and upadacitinib 30 mg QD • Elsubrutinib 60 mg QD and upadacitinib 15 mg QD • Elsubrutinib 60 mg QD and upadacitinib placebo QD • Elsubrutinib placebo QD and upadacitinib 30 mg QD • Elsubrutinib placebo QD and upadacitinib placebo QD	Achievement of SRI-4 and steroid dose ≤ 10 mg prednisone equivalent once a day (QD) at Week 24.	FAS	[redacted]	Difference in the percentage of subjects achieving SRI-4 and steroid dose ≤ 10 mg prednisone equivalent once a day (QD) between treatment groups.

The primary analysis of the primary endpoint will be conducted on FAS based on randomized treatment groups (each active treatment group vs placebo) using NRI-C for missing data handling. Point estimate and 95% CI using normal approximation will be provided for the response rate for each randomized treatment group. The difference between the treatment groups in the primary efficacy endpoint will be assessed using the Cochran-Mantel-Haenszel (CMH) test and will be stratified by baseline corticosteroid dose above 10 mg prednisone-equivalent (≤ 10 mg or > 10 mg), screening SLEDAI-2K (< 10 or ≥ 10), baseline interferon score (high or low or not applicable), baseline immunosuppressant (azathioprine, tacrolimus, cyclosporine, methotrexate [MTX], mycophenolate) (yes or no). Baseline interferon score for patients in China will be

considered as not applicable. Point estimate, 95% CI and nominal p-value for the treatment comparison will be presented.

ABBV-599 treatment arms will be compared to the corresponding monotherapy treatment arms, in addition to comparison to placebo. All comparisons or analyses are summarized below,

- Each ABBV-599 vs placebo
- Upadacitinib vs placebo
- Elsubrutinib vs placebo
- Each ABBV-599 vs Elsubrutinib
- Each ABBV-599 vs upadacitinib
- Elsubrutinib 60 mg QD and upadacitinib 30 mg QD vs Elsubrutinib 60 mg QD and upadacitinib 15 mg QD

There will be no multiplicity adjustment for the multiple comparisons.

In the event the study is partially terminated, terminated groups will be removed from the comparisons above.

8.4.3 Sensitivity and Supplementary Analyses of the Primary Efficacy Endpoint

The same CMH analysis will be repeated based on FAS:

- Using AO data handling without any imputation as additional analysis.
- Using MI approach as additional analysis.

8.5 Secondary Efficacy Endpoints and Analyses

8.5.1 Secondary Efficacy Endpoints

The secondary efficacy endpoints are defined in Section 3.2.

8.5.2 Main Analyses of Secondary Efficacy Endpoints

The attributes of the estimands corresponding to the secondary efficacy endpoints are summarized in Table 2.

Table 2. Summary of the Estimand Attributes of the Secondary Efficacy Endpoints

	Attributes of the Estimand				
Estimand Label	**Treatment**	**Endpoint**	**Population**	**Handling of Intercurrent Events**	**Statistical Summary***
Binary Secondary Endpoint	• Elsubrutinib 60 mg QD and upadacitinib 30 mg QD • Elsubrutinib 60 mg QD and upadacitinib 15 mg QD • Elsubrutinib 60 mg QD and upadacitinib placebo QD • Elsubrutinib placebo QD and upadacitinib 30 mg QD • Elsubrutinib placebo QD and upadacitinib placebo QD	Achievement of SRI-4 at Week 24. Achievement of BILAG-Based Combined Lupus Assessment (BICLA) response at Week 24. Achievement of Lupus Low Disease Activity State (LLDAS) at Week 24.	FAS	[redacted]	Difference in the percentage of subjects achieving each binary secondary endpoint between groups.
Continuous Secondary Endpoint	• Elsubrutinib 60 mg QD and upadacitinib 30 mg QD • Elsubrutinib 60 mg QD and upadacitinib 15 mg QD • Elsubrutinib 60 mg QD and upadacitinib placebo QD • Elsubrutinib placebo QD and upadacitinib 30 mg QD • Elsubrutinib placebo QD and upadacitinib placebo QD	Steroid burden, assessed as change from Baseline at Week 24.	FAS	[redacted]	Difference in the mean change from Baseline in Continuous endpoint between groups.

Table 2. Summary of the Estimand Attributes of the Secondary Efficacy Endpoints (Continued)

	Attributes of the Estimand				
Estimand Label	**Treatment**	**Endpoint**	**Population**	**Handling of Intercurrent Events**	**Statistical Summary***
Count Secondary Endpoint	• Elsubrutinib 60 mg QD and upadacitinib 30 mg QD • Elsubrutinib 60 mg QD and upadacitinib 15 mg QD • Elsubrutinib 60 mg QD and upadacitinib placebo QD • Elsubrutinib placebo QD and upadacitinib 30 mg QD • Elsubrutinib placebo QD and upadacitinib placebo QD	Number of mild, moderate, or severe flares per patient-year (respectively and overall) by Safety of Estrogens in Lupus Erythematosus National Assessment (SELENA) SLEDAI Flare Index (SFI), assessed by number and types of flare (mild/moderate, severe, and any) per subject through Week 24 and through Week 48.	FAS	[redacted]	Ratio in the mean incidence rate of mild, moderate, or severe flares per patient-year between groups.

* Note: In the event the study is partially terminated, terminated groups will be removed from the comparison.

The binary secondary endpoints will be analyzed using the same approach as that for primary endpoint as specified in Section 8.4.2. Specifically, CMH adjusting for randomization stratification factors will be used to construct the treatment difference, the associated 95% CI and p-value between treatment group comparison. The NRI-C will be the primary approach for missing data handling in the analyses of binary secondary efficacy endpoints.

All continuous endpoints will be evaluated through the change from Baseline. For continuous endpoints with multiple post-Baseline visits, the treatment effect will be assessed through Mixed-effect Model Repeated Measures (MMRM) with treatment, visit

and treatment-by-visit interaction, randomization stratification factors, and the continuous fixed covariate of baseline measurement. An unstructured variance covariance matrix will be used. For continuous endpoints with a single post-baseline visit, the treatment effect may be assessed through analysis of covariance (ANCOVA) with baseline value, treatment, and randomization stratification factors as covariates. In addition, the descriptive statistics for continuous variables will be provided including sample size, mean, standard deviation, median, minimum and maximum.

For the count endpoint (e.g., number of flares by SLENAI SFI), negative binomial regression model will be used to assess treatment effect with treatment, and randomization stratification factors as covariates.

8.5.3 Sensitivity and Supplementary Analyses for Secondary Efficacy Endpoints

For all secondary endpoints, analyses using AO data will be performed as additional analyses based on FAS.

8.6 Additional Efficacy Analyses

For the annualized flare rate, a negative binomial regression model will be used with the response variable being the number of flares over the 24-week or 48-week treatment period using both OC and AO data. The model will include covariates of treatment group, and randomization stratification factors. The logarithm (to base e) of the follow-up time will be used as an offset variable in the model to adjust for subjects having different exposure times. The estimated treatment effect and the corresponding 95% CI, as well as the 2-sided p-value will be presented.

A summary of the annualized flare rate by descriptive statistics as well as a summary of the number and percentage of subjects with no flares, at least one flare, 1 flare, 2 flares, and 3 or more flares, respectively, will be presented by treatment group.

Time to event (time to first flare) is defined as number of days from the date of the first dose to the date of the first occurrence of an event. It is calculated as (date of first

occurrence of Event - date of the first dosing + 1). Subjects who discontinue from study prior to experiencing the event are censored at the time of discontinuation. Subjects who are not experiencing the event are censored at the date of last non-missing measurement of SELENA SFI. Data collected on or after the occurrence of intercurrent events as specified in Section 8.2.2 will be excluded. Pairwise comparison between each treatment group and placebo will be performed by log-rank test and Cox proportional hazards regression analysis including treatment and baseline stratification factors as covariates. Same analysis will be performed using AO data as additional analysis.

All other efficacy endpoints specified in Section 3.3 will be analyzed based on FAS using the same approach as that for secondary endpoints as specified in Section 8.5.

8.7 Efficacy Subgroup Analyses

To evaluate the consistency of the efficacy over demographic and other baseline characteristics, subgroup analysis will be performed for the primary endpoint for the subgroups listed in Table 3 below. If any of the resulting subgroups except for age, sex and race has fewer than 5% of the Analysis Population, the subgroup analyses for that category will not be presented.

Point estimate and 95% CI for each treatment group as well as point estimate and 95% CI for treatment differences between treatment groups will be presented using the same analysis method as in the primary analysis. P-value for treatment and subgroup factor interaction using Breslow-Day test will also be provided. In addition:

- A logistic regression with treatment group, subgroup, randomization stratification factors, and treatment by subgroup interaction as factors will be performed to evaluate the significance of the interaction between subgroup and treatment group.
- A stepwise logistic regression with treatment group, all subgroup factors, randomization stratification factors will be performed to evaluate potential important subgroup factor(s).

Table 3. Subgroups

Subgroup Factor	Categories
Age at screening	18 – < 40 years old 40 – < 65 years old ≥ 65 years old
Sex	Male Female
BMI	< 25 ≥ 25
Race	White Black Asian Other
Baseline corticosteroid dose	≤ 10 mg > 10 mg
Screening SLEDAI-2K	< 10 ≥ 10
Baseline immunosuppressant	Yes No
Baseline interferon score	High Low NA
Region	North America/South America, Europe, Asia

9.0 Safety Analyses

9.1 General Considerations

Safety analyses will include reporting of adverse events, laboratory, and vital signs measurements. Safety analyses will be carried out using the Safety Analysis Set for both continued and terminated groups except for mean change from baseline analysis of laboratory and vital signs, which will be provided for continued groups only.

All summaries of adverse events will be provided using data through Week 24, through Week 48 and through Week 48 per 100 patient-years.

The following summary statistics will be presented for subjects who have both baseline and post-baseline values for laboratory parameters and vital signs: the mean value at baseline and at each respective protocol specified visit, and the mean, standard deviation and median for changes from baseline. Categorical data will be summarized using frequencies and percentages. The number of non-missing values will be given.

For the safety analyses, subjects are assigned to a treatment group based on the treatment actually received, regardless of the treatment randomized. A subject's actual treatment will be determined by the most frequent dose regimen received.

For the safety analyses, baseline is defined as the last measurement prior to the first dose of study drug.

Missing safety data will not be imputed.

9.2 Adverse Events

Adverse events (AEs) will be summarized and presented using primary MedDRA System Organ Classes (SOCs) and preferred terms (PTs) according to the version of the MedDRA coding dictionary used for the study at the time of database lock. The actual version of the MedDRA coding dictionary used will be noted in the AE tables and in the clinical study report. Specific adverse events will be counted once for each subject for calculating percentages, unless stated otherwise. In addition, if the same adverse event occurs multiple times within a subject, the maximum toxicity grade and level of relationship to investigational product will be reported.

9.2.1 Treatment-Emergent Adverse Events

Treatment-emergent Adverse Events (TEAE) are defined as any AE that begins either on or after the first dose of study drug in M19-130 and within 30 days after the last dose administration of the study drug in M19-130.

Events where the onset date is the same as the study drug start date are assumed to be treatment-emergent. If an incomplete onset date was collected for an adverse event, the

event will be assumed to be treatment-emergent unless there is other evidence that confirms that the event was not treatment-emergent (e.g., the event end date was prior to the study drug start date).

All treatment-emergent AEs will be summarized overall, as well as by primary MedDRA SOC and Preferred Term. The SOCs will be presented in alphabetical order, and the PTs will be presented in alphabetical order within each SOC.

The number and percentage of subjects experiencing TEAEs will be summarized.

9.2.2 Adverse Event Overview

An overview of AEs will be presented consisting of the number and percentage of subjects experiencing at least one event for each of the following AE categories will be summarized for each treatment group:

- Any treatment-emergent AEs.
- Any COVID-19 related TEAEs.
- Any treatment-emergent AEs related to study drug according to the investigator.
- Any severe treatment-emergent AEs.
- Any serious treatment-emergent AEs.
- Any treatment-emergent AEs leading to discontinuation of study drug.
- Any treatment-emergent AEs leading to death.
- TEAEs of special interest (AESIs) (as defined in Appendix B).
- All deaths.
    - Any COVID-19 related deaths.
    - Occurring <= 30 days after the last dose.
    - Occurring > 30 days after the last dose.

Additional AEs may be considered for tabulation/summary based on recommendations from Clinical and Safety as deemed appropriate.

An overview of AEs will be presented based on data through Week 24 and Week 48. In addition, an overview of AEs through Week 48 per 100 patient-years of study exposure will be presented for the AE categories defined above. The number of TEAEs reported, the total number of years of study drug exposure, and the TEAE rate per 100 patient-years will be presented.

9.2.3 Treatment-Emergent Adverse Events by SOC and/or PT

Treatment-emergent adverse events, including COVID-19 related TEAEs, will be summarized by SOC and PT; by maximum relationship to study drug as assessed by the investigator (e.g., reasonable possibility or no reasonable possibility) and SOC and PT; by maximum toxicity grade and SOC and PT; and by subject number and SOC and PT. Specific adverse events will be counted once for each subject for calculating percentages, unless stated otherwise. In addition, if the same adverse event occurs multiple times within a subject, the maximum toxicity grade and level of relationship to investigational product will be reported.

In addition, TEAEs will be summarized by PT and sorted by decreasing frequency for the total active group.

TEAEs through Week 24, through Week 48 and through Week 48 per 100 patient-years of study exposure will be presented for the AE categories defined above.

9.2.4 Treatment-Emergent Adverse Events per Patient-Years of Exposure

Exposure-adjusted AEs through Week 48 per 100 patient-years will be provided, where AEs per 100 patient-years of exposure are defined as the number of AEs divided by the total exposure in 100 patient-years, i.e.,

$$100 * (\text{Number of TEAEs}) / (\text{Total Patient Years})$$

where total patient years is defined as the sum of the study drug exposure of all subjects, normalized by 365.25, and rounded to 1 decimal place. Note that one event per preferred

term per day per subject will be counted in the calculation of the number of AEs (i.e., a preferred term will not be counted twice on the same day for the same subject).

The TEAE through Week 48 per 100 patient-years will be provided for the following AE categories by SOC and PT:

- Any treatment-emergent AEs.
- Any COVID-19 related TEAEs.
- Any severe treatment-emergent AEs.
- Any serious treatment-emergent AEs.

9.2.5 SAEs (Including Deaths) and Adverse Events Leading to Study Drug Discontinuation

SAEs (including deaths) and AEs leading to study drug discontinuation will be summarized by SOC and PT and in listing format.

9.2.6 Adverse Events of Special Interest

Adverse events of special interest (AESI) through Week 24, through Week 48 and through Week 48 per 100 patient-years will be summarized by SOC and PT. AESI will also be provided in listing format. It will be identified based on standardized or company MedDRA queries (SMQs or CMQs specified in Appendix B).

9.2.7 Adverse Events by "Reasonably Possibly Related" Relationship

TEAEs and reasonably possibly related AEs through Week 24, through Week 48 and through Week 48 per 100 patient-years occurring for more than 2% of the subjects in any of the treatment groups will also be summarized by MedDRA SOC and PT. If a subject has an AE with an unknown relationship, then the subject will be counted in as 'related.'

9.2.8 Frequent (> 5%) Adverse Events and Reasonably Possibly Related Adverse Events by Preferred Term in Decreasing Frequency

TEAEs and reasonably possibly related AEs through Week 24, through Week 48 and through Week 48 per 100 patient-years occurring for more than 5% of the subjects in any of the treatment groups will be summarized by MedDRA PT in decreasing frequency separately.

9.2.9 Listing of Adverse Events

The following additional summaries of AEs will be prepared.

- Listing of Subjects with Treatment-Emergent AESIs.
- Listing of Subjects with Pretreatment SAEs.
- Listing of Subjects with Treatment-Emergent SAEs.
- Listing of Treatment-Emergent AEs that led to discontinuation of study drug.
- Listing of all deaths.

9.3 Analysis of Laboratory Data

Data collected from central and local laboratories, including additional laboratory testing due to an SAE, will be used in all analyses, except for baseline, where SAE-related laboratory assessments on or before the first dose of study drug will be excluded. The clinical laboratory tests defined in the protocol operations manual (e.g., hematology and clinical chemistry) will be summarized.

For the analysis of laboratory data, values observed up to 30 days after the last dose of study drug will be included.

Each laboratory variable will be summarized for all time points (starting with baseline) with the number of non-missing observations, mean and standard deviation, median, minimum and maximum.

For continued groups, mean change from baseline to each applicable post-baseline visit will be summarized for selected laboratory variables, with the number of observations, baseline mean, and visit mean. The change from baseline mean, standard error, and 95% confidence interval will be presented for the mean change from baseline within each treatment group and difference between treatment groups.

Shift tables will be provided for each hematology, clinical chemistry parameter. Shifts from Baseline to the following endpoints will be considered: post-Baseline value and final value. This analysis will be performed only for category of "low or normal" at Baseline which shift to 'high' and category of "high or normal" at Baseline which shift to category of 'low.'

For selected laboratory variables including hemoglobin, neutrophils, lymphocytes, platelets, leukocytes, creatinine, ALT, AST, bilirubin, CPK, a listing of all observations collected will be generated for subjects that had at least one post-baseline observation meeting pre-defined criteria for potentially clinically important values. The number and percentage of subjects in each treatment group who have at least one post-baseline observation meeting the pre-defined criteria for potentially clinically important values will be provided for each variable. The criteria for potentially clinically important laboratory values will be determined by NCI CTCAE version 4.03 of greater than Grade 3 or greater than Grade 4.

For the purpose of assessing for potential Hy's law cases, the frequencies and percentages of subjects with post baseline liver-specific function test values that meet the following criteria of potential clinical interest will be summarized by treatment group for both short-term and long-term analyses:

- ALT $\geq 3 \times$ ULN
- ALT $\geq 5 \times$ ULN
- ALT $\geq 10 \times$ ULN
- ALT $\geq 20 \times$ ULN
- AST $\geq 3 \times$ ULN

- AST ≥ 5 × ULN
- AST ≥ 10 × ULN
- AST ≥ 20 × ULN
- TBL ≥ 2 × ULN
- Alkaline phosphatase ≥ 1.5 × ULN
- ALT and/or AST > 3 × ULN and TBL > 1.5 × ULN
- ALT and/or AST > 3 × ULN and TBL > 2 × ULN

A listing of possible Hy's Law cases, defined as those who meet all of the following conditions at any post-baseline visit will be provided: ALT > 3 × ULN or AST > 3 × ULN and Total Bilirubin ≥ 2 × ULN.

9.4 Analysis of Vital Signs

Vital sign measurements of systolic and diastolic blood pressure, pulse rate, body temperature, and weight will be summarized.

For the analysis of vital sign data, values observed up to 30 days after the last dose of study drug will be included.

Each vital sign variable will be summarized for all time points (starting with Baseline) with the number of non-missing observations, mean and standard deviation, median, minimum and maximum.

For continued groups, mean change from baseline to each applicable post-baseline visit will be summarized for each vital sign variable, with the number of observations, baseline mean, and visit mean. The change from baseline mean, standard error, and 95% confidence interval will be presented for the mean change from baseline within each treatment group and difference between treatment groups.

Vital sign variables will be evaluated based on potentially clinically important (PCI) criteria (Appendix C). For each vital sign PCI criterion, the number and percentage of subjects who have a vital sign value meeting the criteria will be summarized. Listings

will be provided to summarize subject-level vital sign data for subjects meeting PCI criteria.

9.5 Safety Subgroup Analyses

No planned safety subgroup analyses.

9.6 Other Safety Analyses

ECG is collected at Screening and Weeks 2, 4, 8, 12, 16, 20, 24 and 48. The number and percentage of subjects with QTcF prolongation (QTcF > 500 msec) will be summarized by treatment groups using Safety Analysis Set population.

10.0 Other Analyses

No other analyses are planned.

11.0 Interim Analysis

11.1 Data Monitoring Committee

An unblinded internal data monitoring committee (DMC) has been established to ensure the overall integrity and conduct of the study. DMC reviews will be conducted quarterly. In addition to the unblinded internal DMC, a multi-disciplinary internal Safety Review team will also conduct regular reviews in a blinded fashion during the study and at study completion. AbbVie will also conduct regular systematic reviews of emerging safety data from all clinical studies with elsubrutinib, upadacitinib and ABBV-599 (elsubrutinib/upadacitinib) combination.

A separate DMC charter describes the roles and responsibilities of the DMC members, frequency of data reviews, relevant data to be assessed, and general operations.

Since there are no efficacy analyses for early stopping, no alpha adjustment is needed.

11.2 Interim Analysis

An unblinded interim analysis will be performed after the first 50% of the subjects complete Week 24 visit or withdraw from the study by an independent team at AbbVie that does not include members of the blinded M19-130 Study team. The objective of the interim analysis is to reassess the treatment regimens in the rollover Study M20-186. The Interim Analysis results will be reviewed by an Internal Executive Review Committee (IERC) following a separate IERC charter and one or more treatment groups may be terminated. The IERC members are not involved in the study but have the depth of technical knowledge and experience to make decisions based on Interim Analysis results.

An interim unblinding plan will be developed separately describing the analysis to be presented to the IERC and will include execution logistics, unblinded analysis team, and the data chain of custody to protect the integrity of the clinical study.

11.2.1 Interim Analysis Population

Efficacy Analysis Population

- Full Analysis Set for Interim Analysis (FAS_IA): all randomized subjects with at least one dose of study drug before the cutoff date (20JUN2021). Subjects will be analyzed as randomized.
    - FAS_IA50 Subset: the subset of FAS_IA which includes the first 50% of planned randomized subjects.

Safety Analysis Population

- Safety Analysis Set for Interim Analysis (SA_IA) includes same subjects as FAS_IA. Subjects will be analyzed as treated.
    - SA_IA50 Subset includes same subjects as FAS_IA50 Subset. Subjects will be analyzed as treated.

Analysis of efficacy endpoints and summaries of baseline demographics and disease characteristics will be based on the FAS_IA50 Subset for primary analysis and on the FAS_IA for additional supportive analysis.

Safety analysis will be performed based on the SA_IA for primary analysis and on the SA_IA50 Subset for additional supportive analysis.

Cutoff dates for all efficacy and safety analysis are as follows:

- For subjects in FAS_IA50 Subset or SA_IA50 Subset, the cutoff date is 7JUL2021.
- For the rest of subjects, the cutoff date is 20JUN2021.

This interim analysis will also include PK and Biomarker data. PK analysis and communication plan will be provided separately by CPPM.

11.2.2 Subject Disposition, Demographics and Disease Characteristics, and Treatment Compliance

The following summaries will be provided based on FAS_IA and FAS_IA50 Subset:

- Subject disposition (randomized, took at least one dose, completed study drug at Week 24, completed study drug at Week 48, prematurely discontinued study drug, completed the 24 weeks study period, completed Week 48 Study Period, prematurely discontinued study, enrolled in LTE (M20-186)).
- Demographics and Disease Characteristics:
  - Sex (male, female).
  - Age (years).
  - Age category [18 – < 40 years old, 40 – < 65 years old, ≥ 65 years old].
  - Ethnicity (Hispanic or Latino, Not Hispanic or Latino).
  - Race (White, Black or African American, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Asian, Other).
  - Geographic region (North America, South/Central America, Western Europe, Eastern Europe, Asia Japan, Asia Other, Other).

- Weight (kg).
- Weight Categories (< 60 kg, ≥ 60 kg).
- Height (cm).
- Body Mass Index (BMI) (kg/m^2).
- BMI Category (BMI < 25 kg/m^2, BMI ≥ 25 kg/m^2).
- Study drug duration and compliance.

11.2.3 Efficacy Analyses

Primary Endpoint

The primary efficacy endpoint (achieving SRI-4 and steroid dose ≤ 10 mg prednisone equivalent once a day (QD) at Week 24). Will be analyzed according to the study SAP (Section 8.4) based on the FAS_IA50 Subset.

Additional analyses will also be performed for the primary endpoint at the interim analysis based on FAS_IA50 Subset: for subjects who don't have Week 24 data, all available data will be used to impute Week 24 response using multiple imputation method.

Secondary and Additional Endpoints

The following secondary and additional endpoints will be included in the interim analysis:

- Achievement of SRI-4 at all visits.
- Achievement of BICLA response at all visits.
- Achievement of ≥ 4-point decrease in SLEDAI-2K from Baseline at all visits.
- Change from Baseline in SLEDAI-2K at all visits.
- Change from Baseline in PhGA at all visits.
- Number of mild, moderate, or severe flares per patient-year (respectively and overall) by Safety of Estrogens in Lupus Erythematosus National Assessment (SELENA) SLEDAI Flare Index (SFI), assessed by number and types of flare

(mild/moderate, severe, and any) per subject through Week 24 and through Week 48.

- Achievement of at least a 50% reduction from baseline in prednisone dose at all visits.
- Achievement of SRI-4 and steroid dose ≤ 10 mg prednisone equivalent once a day (QD) at all visits starting from Week 8.
- Steroid burden, assessed as change from Baseline at all visits.
- Change from Baseline in Functional Assessment of Chronic Illness Therapy (FACIT)-Fatigue at all visits.
- Change from Baseline in Pain Numerical Rating Scale (NRS) at all visits.

Analysis specified in Section 8.5 will be performed.

11.2.4 Safety Analyses

Safety Analysis specified in this SAP will be included in the interim analysis except summary of other baseline characteristics listed in Section 7.1:

The following safety results will be summarized in the interim analysis:

- Overview of TEAEs (treatment emergent adverse events).
- TEAEs by SOC (system organ class) and PT (preferred term).
- Treatment-emergent SAEs (serious adverse events) by SOC and PT.
- TEAEs leading to study drug discontinuation by SOC and PT.
- TEAEs by SOC and PT and maximum relationship to study drug.
- TEAEs by SOC and PT and maximum toxicity grade.
- AESIs (adverse events of special interest) by SOC and PT.
- Listing of Pre-treatment SAEs.
- Listing of SAEs.
- Listings of AESIs.
- Listing of Aes leading to death.
- Listing of Aes leading to study drug discontinuation.

- Analysis of mean change from baseline in laboratory and vital sign variables.
- Summary of potentially clinically significant laboratory and vital sign variables.
- Listings of potentially clinically significant laboratory and vital sign variables.
- Shift tables for laboratory and vital sign variables.

12.0 Overall Type-I Error Control

All tests will be performed at two-sided statistical significance level of $\alpha = 0.1$. No multiplicity adjustment will be conducted.

13.0 Version History

Table 4. SAP Version History Summary

Version	Date	Summary
1.0	12 Apr 2019	Original version
2.0	24 July 2021	• Revise the SAP per the latest SAP template and to maintain consistency with the most recent protocol. • Add definitions of estimand for primary and secondary endpoints. • Clarified the MI method to handle missing data due to COVID-19 and other reasons. • Update Hy's Law rules to be consistent with the most recent Upadacitinib PSSAP. • Add sections to include the planned interim analysis.
3.0	17 Dec 2021	• Update SAP language to be consistent with Protocol Amendment Version 7. • Add one efficacy endpoint in Section 3.3 • Update subgroup analysis for primary endpoint in Section 8.7. • Clarify how efficacy analysis and safety analysis will be performed for discontinued treatment groups and continued treatment groups if the study is partially terminated after the planned interim analysis,
4.0	26 July 2022	• Update AESI searching criteria to be consistent with the most recent Upadacitinib PSSAP. • Add additional efficacy endpoints for exploratory purpose in Section 3.3.

14.0 References

None

Appendix A. Protocol Deviations

The number and percentage of subjects who reported at least one of the following protocol deviation categories will be provided.

- Subject entered into the study even though s/he did not satisfy entry criteria.
- Subject developed withdrawal criteria during the study and was not withdrawn.
- Subject received wrong treatment or incorrect dose of study.
- Subject took prohibited concomitant medication.

Appendix B. Definition of Adverse Events of Special Interest

Adverse Events of Special Interest (AESI) will be identified using the following search criteria:

AESI	Type of MedDRA Query	Broad or Narrow Search	SMQ/CMQ Search Criteria
Serious Infections	CMQ		"Infections" – Subset for SAEs
Opportunistic Infection excluding Tuberculosis and Herpes Zoster	CMQ (code 80000189)		"Opportunistic Infection excluding Tuberculosis and Herpes Zoster"
Herpes Zoster	CMQ (code 80000175)		"Herpes Zoster"
Active Tuberculosis	CMQ (code 80000188)		"Active Tuberculosis"
Possible Malignancy	SMQ	Narrow	"Malignancies"
Malignancy	SMQ	Narrow	"Malignant Tumours"
Non-Melanoma Skin Cancer (NMSC)	SMQ	Narrow	"Skin Malignant Tumours" (SMQ 20000204) removing Melanoma CMQ (code 80000119)
Malignancies excluding NMSC	SMQ	Narrow	"Malignant Tumours" removing NMSC output
Adjudicated GI perforations			Based on adjudicated results (the identification of events to be adjudicated are described in the GI Perforation charter)
Adjudicated cardiovascular events: MACE* Cardiovascular Death Non-fatal Myocardial Infarction Non-fatal Stroke Undetermined/Unknown Cause of Deaths Other Cardiovascular events	Output from CAC		
Anemia	CMQ		"Non-Hemolytic and Non-Aplastic Anemias"

AESI	Type of MedDRA Query	Broad or Narrow Search	SMQ/CMQ Search Criteria
Neutropenia	CMQ		"Hematological Toxicity – Neutropenia"
Lymphopenia	CMQ		"Hematological Toxicity – Lymphopenia"
Renal dysfunction	SMQ	Narrow	"Acute Renal Failure"
Hepatic disorders	SMQ	Narrow	"Drug Related Hepatic Disorders" - Comprehensive Search SMQ
Adjudicated Thrombotic Events: VTE** Deep Vein Thrombosis Pulmonary Embolism Other Venous Thrombosis	Output from CAC		

* MACE: Major Adverse Cardiovascular Events, defined as cardiovascular death, non-fatal myocardial infarction and non-fatal stroke.

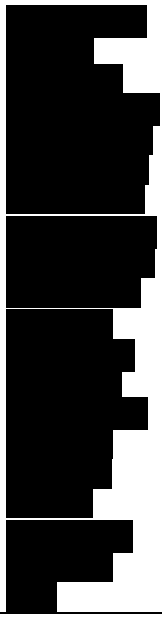
** VTE: Venous thromboembolic events, defined as deep vein thrombosis (DVT) and pulmonary embolism (PE) (fatal and non-fatal).

Appendix C. Potentially Clinically Important Criteria for Safety Endpoints

Vital Sign	Category	Criteria for Potential Clinically Important Vital Signs
Systolic blood pressure	Low	Value ≤ 90 mmHg and decrease ≥ 20 mmHg from Baseline
	High	Value ≥ 160 mmHg and increase ≥ 20 mmHg from Baseline
Diastolic blood pressure	Low	Value ≤ 50 mmHg and decrease ≥ 10 mmHg from Baseline
	High	Value ≥ 100 mmHg and increase ≥ 10 mmHg from Baseline
Weight	High	> 7% increase from baseline
	Low	> 7% decrease from baseline

Appendix D. Attributes of the Estimand for Primary and Secondary Endpoints

	Attributes of the Estimand				
	Population	**Endpoint**	**Treatment**	**Intercurrent Events**	**Statistical Summary***
Primary	FAS	Achievement of SRI-4 and steroid dose ≤ 10 mg prednisone equivalent once a day (QD) at Week 24.	Elsubrutinib 60 mg QD and upadacitinib 30 mg QD Elsubrutinib 60 mg QD and upadacitinib 15 mg QD Elsubrutinib 60 mg QD and upadacitinib placebo QD Elsubrutinib placebo QD and upadacitinib 30 mg QD Elsubrutinib placebo QD and upadacitinib placebo QD	[redacted]	Difference in the percentage of subjects achieving SRI-4 and steroid dose ≤ 10 mg prednisone equivalent once a day (QD) between groups.
Binary Secondary Endpoint	FAS	Achievement of SRI-4 at Week 24. Achievement of BILAG-Based Combined Lupus Assessment (BICLA) response at Week 24. Achievement of Lupus Low Disease Activity State (LLDAS) at Week 24.	Elsubrutinib 60 mg QD and upadacitinib 30 mg QD Elsubrutinib 60 mg QD and upadacitinib 15 mg QD Elsubrutinib 60 mg QD and upadacitinib placebo QD Elsubrutinib placebo QD and upadacitinib 30 mg QD Elsubrutinib placebo QD and upadacitinib placebo QD	[redacted]	Difference in the percentage of subjects achieving each binary secondary endpoint between groups.

	Attributes of the Estimand				
	Population	**Endpoint**	**Treatment**	**Intercurrent Events**	**Statistical Summary***
Continuous Secondary Endpoint	FAS	Steroid burden, assessed as change from Baseline at Week 24.	Elsubrutinib 60 mg QD and upadacitinib 30 mg QD Elsubrutinib 60 mg QD and upadacitinib 15 mg QD Elsubrutinib 60 mg QD and upadacitinib placebo QD Elsubrutinib placebo QD and upadacitinib 30 mg QD Elsubrutinib placebo QD and upadacitinib placebo QD		Difference in the mean change from Baseline in Continuous endpoint between groups.
Count Secondary Endpoint		Number of mild, moderate, or severe flares by Safety of Estrogens in Lupus Erythematosus National Assessment (SELENA) SLEDAI Flare Index (SFI), assessed by number and types of flare (mild/moderate, severe, and any) per subject through Week 24 and through Week 48.	Elsubrutinib 60 mg QD and upadacitinib 30 mg QD Elsubrutinib 60 mg QD and upadacitinib 15 mg QD Elsubrutinib 60 mg QD and upadacitinib placebo QD Elsubrutinib placebo QD and upadacitinib 30 mg QD Elsubrutinib placebo QD and upadacitinib placebo QD		Ratio in the mean incidence rate of mild, moderate, or severe flares per patient-year between groups.

* Note: In the event the study is partially terminated, terminated groups will be removed from the comparisons above.

Appendix E. Random Seeds

In case of non-convergence, the random seed will be updated by adding 100000 at each attempt until convergence of model happens.

Variables	Random Seed	
	MCMC Procedure	**PROC MI**
SLEDAI-2K Total Score*	20001	80001
PhGA*	20002	80002
TJC28*	20003	80003
SJC28*	20004	80004
CLASI*	20005	80005
FACIT-F	20006	80006
PtGA	20007	80007
Pain NRS	20008	80008
No new BILAG A or more than 1 new BILAG B disease activity scores	20009	80009

* This only applies when these variables are a component to derive a binary endpoint.