- **Protocol number:** D3461C00007
- **Document title:** A Multicentre, Randomised, Double-blind, Placebo controlled, Phase 2 Study Evaluating the Efficacy and Safety of Anifrolumab in Adult Subjects with Active Proliferative Lupus Nephritis
- **NCT number:** NCT02547922
- **Version number:** 5.0
- **Date of the document:** 24 March 2020

AstraZeneca

Statistical Analysis Plan

Study Code	D3461C00007
Edition Number	5.0
Date	24 March 2020

A Multicentre, Randomised, Double-blind, Placebo-controlled, Phase 2 Study Evaluating the Efficacy and Safety of Anifrolumab in Adult Subjects with Active Proliferative Lupus Nephritis

TABLE OF CONTENTS

CCI [redacted]

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
aCRR	Alternative Complete Renal Response
ACEI	Angiotensin-converting-enzyme inhibitor
ADA	Anti-drug antibodies
AE	Adverse event
AESI	Adverse event of special interest
AIFA	Italian Medicine Agency
ALP	Alkaline phosphatase
ALT	Alanine transaminase
ANA	Antinuclear antibody
ARB	Angiotensin II receptor blockers
AST	Aspartate transaminase
AUC	Area under the curve
BDR	Blind data review/Blind delivery review
BMI	Body mass index
C-SSRS	Columbia–Suicide Severity Rating Scale
C3	Third component of complement
C4	Fourth component of complement
CI	Confidence interval
CMH	Cochran-Mantel-Haenszel
CPP	Comité de Protection des Personnes
CRR	Complete Renal Response
CSR	Clinical Study Report
CVA	Cerebrovascular accident
dsDNA	Double-stranded deoxyribonucleic acid
DMPK	Drug metabolism and pharmacokinetics
ECG	Electrocardiogram
eCRF	Electronic case report form
EDC	Electronic data capture
eGFR	Estimated Glomerular Filtration Rate
GGT	Gamma glutamyl transferase

Abbreviation or special term	Explanation
hpf	high power field
ICF	Informed consent form
IFN	Interferon
Ig	Immunoglobulin
IM	Intra-muscular
ISN	International Society of Nephrology
IV	Intravenous
kg	kilograms
LTB	Latent tuberculosis
LTE	Long-term extension
m	metre
MACE	Major adverse cardiovascular events
MAD	Median absolute deviation
MDRD	Modification of diet in renal disease
MedDRA	Medical Dictionary for Regulatory Activities
mg	milligram
min	minute
mITT	Modified intent-to-treat
mL	millilitres
MMF	Mycophenolate mofetil
mmHg	millimetres of mercury
MMRM	Mixed model for repeated measures
MPA	Mycophenolic acid
ms	milliseconds
nAb	Neutralising antibodies
NSAID	Non-steroidal anti-inflammatory drug
LN	Lupus Nephritis
OAE	Other significant AE
OCS	Oral corticosteroids
PD	Pharmacodynamic
PGA	Physician’s Global Assessment

Abbreviation or special term	Explanation
PHQ-8	Personal Health Questionnaire Depression Scale-8
PK	Pharmacokinetic(s)
PRR	Partial Renal Response
PtGA	Patient Global Assessment
Q	Question
Q-Q	Quartile-quartile
Q4W	Every 4 weeks
QTcB	QTc with Bazett's formula
RBC	Red blood cell
RPS	Renal Pathology Society
SAE	Serious adverse event
SAP	Statistical analysis plan
SAS	Statistical analysis system
SD	Standard deviation
SDI	Systemic Lupus International Collaborating Clinics (SLICC)/American College of Rheumatology (ACR) Damage Index
SI	Internal system of units
SLE	Systemic lupus erythematosus
SLEDAI-2K	Systemic Lupus Erythematosus Disease Activity Index 2000
SOC	Standard of care
TB	Tuberculosis
TELVC	Treatment emergent laboratory/vital signs changes
ULN	Upper limit of normal
UPCR	Urine protein to creatinine ratio
VAS	Visual analogue scale
WHO	World Health Organisation

AMENDMENT HISTORY

Date	Brief description of change
18 Mar 2020 Version 5.0	Updated Statistical Analysis Plan (SAP) as follows: • Clarified that the "while on treatment" estimand includes data up to and including date of IP discontinuation, and if IP has not been reported as prematurely discontinued, available data up to and included Week 104 will be used for continuous endpoint analyses. Otherwise, the date of IP discontinuation will be determined as the date of last IP administration. However, all subjects who have an infusion of IP within the Week 48 analysis visit window will be determined to have completed treatment in the double-blind treatment period, and their Week 52 data will be included in the analyses. • Clarification surrounding the Week 52 exemption of non-responder imputation due to discontinuation of IP was added. • Added table for programmatically identifying prohibited medications to Appendix A. • Added numbering and clarifications, including some rewording, for restricted medication rules in Appendix B. • Specified the calculation to be used for Cochran-Mantel-Haenszel (CMH) p-values. • Added explanatory text for the time to achieve 24-hour UPCR $\leq$ 0.7 mg/mg and time to achieve CRR variables in Section 3.1 and 3.2.1 respectively. • Clarification added that the tipping point analyses for CRR and sustained OCS response will only be statistically interpreted for a treatment comparison and visit if the nominal p-value for the treatment comparison from the main analysis of the endpoint at the specified visit is <0.05 (in favour of anifrolumab). • Added clarification that the "percent suppression of fold change" will be presented as the "percent of baseline IFN21 PD signature" in all summaries, figures and listings. • Updated all time to event variable censoring definitions to remove date of restricted medication, and to specify what date is used for censoring in the case of discontinuation of IP, depending on when the last dose if IP was received, and for subjects completing treatment. Flare exposure time was updated similarly. • Shift tables of changes in renal biopsy ISN/RPS classification from first to second years have been removed due to no renal biopsies being peformed in the second year. • The table summarising disallowed medications has been split into "on IP" and "off IP".

Date	Brief description of change
	• CCI [redacted] • During the blind delivery review (BDR; initiated 05 Feb 2020), it has been identified that the use of body weight-based (mg/kg) maximal daily oral corticosteroid (OCS) dose for classifying patients as non-responders when receiving any OCS dose >0.5 mg/kg/day is considered to be too stringent and clinically inappropriate for the following reasons: o These marginally higher doses per body weight are very unlikely to be clinically meaningful and are not thought to confound the renal efficacy assessments, especially since additional intravenous (IV) corticosteroid pulses up to 0.5 g are allowed early in study and these IV doses are not body weight-based and by far exceed 0.5 mg/kg/day threshold used for OCS. o Body weight often varies over time especially in patients with active renal disease and on corticosteroid treatment. Even minor changes in body weight can result in OCS dose exceeding the threshold at fluctuating time points. o For practical reasons the OCS dose must be rounded and adjusted to available strengths of OCS pills with risk of exceeding the threshold even if only slightly rounded upwards. o Using body weight to guide OCS dosing is commonly used as a rule of thumb in clinical practice, but exact calculations are rarely performed and as mentioned doses are commonly rounded to match available strengths of eg prednisolone pills. Consequently, this specific threshold (>0.5 mg/kg/day) was removed from the following rule: Prednisone-equivalent >0.5 mg/kg/day or >40 mg/day for 2 or more consecutive days after Day 1.

Date	Brief description of change
20 Mar 2019 Version 4.0	SAP to include updates to the restricted medications, non-responder rules and the interim analysis. **Interim analysis** Text was added to explain that the possible interim analysis will be either efficacy or futility, dependent on the results of a pivotal Phase 3 SLE study in anifrolumab. Non-binding stopping rules for the possible futility analysis have been included. **Changes to statistical analysis** The method used to manage strata with low counts for the CMH analyses has been clarified and updated. MedImmune will no longer be calculating the type I IFN 21-gene signature parameters. Instructions on how to calculate these have been added to the SAP **Other changes** Medication beyond the protocol-allowed threshold and responder/non-responder criteria have been clarified in separate appendices (Appendix A and Appendix B, respectively). P-values will be presented to 4 decimal places. The estimand for all continuous endpoints (including the primary endpoint) is changing to the "while on treatment strategy" from the "treatment policy strategy". The "while on treatment strategy" uses data up to Investigational Product (IP) discontinuation. The "treatment policy strategy" uses all collected data, irrespective of whether a subject discontinued IP or not. This change aims to deal with the risk of bias due to rescue medications after IP discontinuation. The sensitivity analysis will be changed to follow the "treatment policy strategy". The definition of oral corticosteroids at baseline has been updated. OCS subgroup thresholds have been changed from (≤20 mg/day, >20 mg/day) to (<20 mg/day, ≥20 mg/day)

Date	Brief description of change
12 Feb 2018 Version 3.0	Updated SAP in line with Global Protocol Amendment 4, as follows: **The renal function and proteinuria components of renal response criteria was modified by changing the estimated glomerular filtration rate (eGFR) and 24-hour urine protein to creatinine ratio (UPCR) cut-off values.** For complete renal response (CRR), the alternative complete renal response (aCRR), the partial renal response (PRR), graded CRR, and graded aCRR, the cut-off values for the renal function (eGFR) was changed to: • eGFR to ≥60 mL/min/1.73m^2 or no confirmed decrease of eGFR from baseline of ≥20% For CRR and aCRR, the cut-off value for proteinuria (24-hour UPCR) was changed to: • 24-hour UPCR to ≤0.7 mg/mg For graded CRR and graded aCRR, the cut-off values for proteinuria was changed to: • For subjects with baseline 24-hour UPCR of >3 mg/mg 24-hour UPCR cut-off was changed to ≤1 mg/mg • For subjects with baseline 24-hour UPCR of ≤3 mg/mg 24-hour UPCR cut-off was changed to ≤0.7 mg/mg The flare definition for proteinuria (UPCR) was changed to >1.5 mg/mg. **Changes to objectives** aCRR and 24-hour UPCR were added as exploratory objectives at Week 104. Clarification was added that International Society of Nephrology (ISN)/Renal Pathology Society (RPS) classification and National Institute of Health (NIH) indices were also summarised based on CRR and PRR at Week 104. **Interim analysis** Text was added that an interim analysis may be performed after approximately 50% of subjects have completed the Week 52 visit. CCI [redacted] The statistical analysis plan (SAP) v1.0 was completed and the last SAP amendment will be completed prior to unblinding of data. CCI [redacted]

Date	Brief description of change
	Changes to statistical analysis Text was added that strong control of the familywise error rate will be performed for the primary and secondary endpoints for the pooled anifrolumab group compared with placebo as well as the respective tests for the individual anifrolumab regimens and the testing strategy to account for multiplicity considerations has been described. The power and minimal detectable difference was updated based on the primary endpoint for the pooled anifrolumab group compared with placebo. The sample size will provide approximately 86% power with a 2-sided alpha of 0.049. • Text was added that the primary and secondary endpoints are based on the pooled anifrolumab group compared with the placebo group. • One additional subgroup; eGFR at baseline (<60 mL/min/1.73m^2, ≥60 mL/min/1.73m^2) has been added. The geographic region subgroup has been renamed placebo response region. • The age subgroups were changed to ≥18 to 64 and ≥65 years. • It has been clarified that all personnel involved in the conduct of the study will remain blinded until database soft lock. It has been clarified that nominal p-values can be presented for endpoints not included in the strategy for preserving type 1 error rate. **Other changes** It has been clarified that no increase in OCS, or the use of IV, intra-articular, tendon sheath or bursal injections is allowed from Week 40 until Week 52 assessment. Criteria for discontinuing IP at Week 52, Week 56 or Week 60 has been included. Assessment for 24-hour UPCR should be completed for subjects who discontinue IP. Randomised subjects who receive anti-malarial therapy within 12 months prior to signing ICF must have an eye exam by a qualified professional within 12 months prior to signing the ICF and that subjects starting anti-malarial therapy during the Screening Period must have an eye exam within 12 weeks after signing the ICF.

Date	Brief description of change
12 Apr 2017 Version 2.0	Updated Statistical Analysis Plan (SAP) in line with Global Protocol Amendments 1 and 2, as follows: • Text was added to clarify that stratification sample for 24-hour urine protein to creatinine ratio (UPCR) can be obtained within 14 days prior to the expected date of randomisation. Without the results of this second (stratification) sample, subjects cannot be randomised. Turn-around time for results from central laboratory is up to 7 days. On rare occasion an extension of the 30-day screening window is allowed if the re-collection of the sample is necessary or the results needed for randomisation are delayed. • An error noted in relation to the duration of administration (first "12 weeks" instead of first "3 doses of the investigational product") has been corrected throughout the SAP. • The abbreviation "SFI" has been deleted throughout the SAP. The term SFI has been replaced by the term "SLEDAI-2K based Flare Assessment Instrument" • The outcome measure to evaluate the safety and tolerability of anifrolumab has been updated as "extra-renal flares using Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K) based Flare Assessment Instrument". • For the exploratory objectives, the criteria to determine improvement in 24-hour UPCR for subjects with a baseline UPCR >3 mg/mg has been updated as >50% improvement from baseline and ≤3.0 mg/mg (previously <3.0). • A note has been added to clarify that MMF will be supplied by the Sponsor to the subjects from the day of randomisation onwards until the end of subject's participation in the study. • The criteria for discontinuing investigational product at any time have been updated to include: receipt >1 methylprednisolone pulse after the day of randomisation, receipt of any methylprednisolone pulse after Week 8, discontinuation of MMF and initiation of another immunosuppressant. In addition, text was added to clarify that decrease in eGFR is to be based on two independent samples. • The criteria for discontinuing investigational product at Week 12 and Week 24 have been revised to clarify the wording on oral corticosteroid and that decrease is eGFR is to be based on two independent samples. Cut-off values for nephrotic range UPCR have been aligned with that of UPCR values for stratification. • The minimum dose of MMF has been updated from ≤1.5 gm/day to ≤1.0 gm/day. • Danazol, dapsone, sulfasalazine, and prednisolone have been removed from the list of restricted medications.

Date	Brief description of change
	• Text was added to clarify the study design of the second year extension period. Eligible subjects enrolled in the original Phase 2 study will continue in their randomised treatment groups. The treatment groups will continue to receive investigational product infusions (anifrolumab 300 mg or placebo) every 4 weeks (Q4W) starting at Week 52 for an additional 48 weeks in addition to standard of care (SOC) until Week 100. Final efficacy assessment will be made at Week 104. Safety follow-up will continue Q4W until Week 112. • Criteria for discontinuation of investigational product at Week 52 were added. • Criteria for discontinuation of investigational product during the second year extension period were added. • CCI [redacted]. • CCI [redacted] • CCI [redacted] • CCI [redacted] • Two new subgroups were added: ADA (positive at any time, negative), and OCS dose at baseline (≤20 mg/day, >20 mg/day)

1 STUDY DETAILS

1.1 Study objectives

1.1.1 Primary objective

Primary Objective:	Outcome Measure:
To evaluate the efficacy of anifrolumab plus standard of care (SOC[a]) compared with placebo plus SOC in subjects with active proliferative lupus nephritis (LN) measured by the relative difference in change from baseline to Week 52 in 24-hour urine protein to creatinine ratio (UPCR)	24-hour UPCR

[a] Required SOC is described in Section 1.2.1

1.1.2 Secondary objectives

Secondary Objectives:	Outcome Measures:
To evaluate the effect of anifrolumab plus SOC compared with placebo plus SOC[a] on the proportion of subjects achieving Complete Renal Response (CRR) at Week 52	CRR is defined as meeting all of the following: • Estimated glomerular filtration rate (eGFR) is – ≥60 mL/min/1.73m^2 or no confirmed decrease of eGFR from baseline of ≥20% • 24-hour UPCR≤0.7 mg/mg • No discontinuation of investigational product (IP) or use of restricted medication[b] beyond the protocol-allowed threshold before assessment eGFR is based on Modification of Diet in Renal Disease (MDRD) formula

[a] Required SOC is described in Section 1.2.1
[b] Allowed medication are described in Appendix A and Appendix B

1.1.3 Safety objective

Safety Objective:	Outcome Measures:
To characterise the safety and tolerability of anifrolumab	Adverse events (AEs) (including AEs of special interest [AESIs]), vital signs, physical examination, baseline and End of Treatment 12-lead electrocardiograms (ECG), and clinical laboratory tests (haematology, clinical chemistry, urinalysis), Columbia-Suicide Severity Rating Scale (C-SSRS), Personal Health Questionnaire Depression Scale-8 (PHQ-8), extra-renal flares using Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K)-based Flare Assessment Instrument .

1.1.4 Exploratory objectives

Exploratory Objectives:	Outcome Measures:
To evaluate the effect of anifrolumab plus SOC[a] compared with placebo plus SOC[a] on:	
CCI [redacted]	[redacted]
CCI [redacted]	[redacted]
CCI [redacted]	[redacted]
Proportion of subjects achieving alternative CRR (aCRR) at Week 52 (and Week 104) The difference between the CRR and the aCRR is the addition of a criterion regarding "inactive urine sediment"	aCRR is defined as meeting all of the following: • eGFR is - ≥60 mL/min/1.73 m^2 or no confirmed decrease of eGFR from baseline of ≥20% • 24-hour UPCR ≤0.7 mg/mg • Inactive urine sediment (defined as <10 red blood cells [RBC]/hpf) • No discontinuation of IP or use of restricted medication[b] beyond the protocol-allowed threshold before assessment eGFR is based on MDRD formula

CCI [redacted]	[redacted]
Proportion of subjects meeting Graded aCRR at Week 52 (and Week 104)	Graded aCRR is defined as meeting both the 24-hour UPCR and eGFR criteria: • A decrease in 24-hour UPCR: – For subjects with baseline UPCR >3 mg/mg: UPCR ≤1 mg/mg – For subjects with baseline UPCR ≤3 mg/mg: UPCR ≤0.7 mg/mg • eGFR: – ≥60 mL/min/1.73 m^2 or no confirmed decrease of eGFR from baseline of ≥20% • Inactive urine sediment (defined as <10 RBC/hpf) • No discontinuation of IP or use of restricted medication[b] beyond the protocol allowed threshold before assessment
Proportion of subjects able to achieve sustained reduction in oral corticosteroids (OCS) dose at Week 52 or Week 104	**Sustained reduction of OCS dose:** • Week 52: Prednisone-equivalent dose ≤7.5 mg/day by Week 24 and not exceeding this dose through Week 52 • Week 104: Prednisone-equivalent dose ≤5.0 mg/day by Week 80 and not exceeding this dose through Week 104 **and** • No discontinuation of IP or use of restricted medication[b] beyond the protocol-allowed threshold before assessment

Proportion of subjects achieving CRR at Week 52 or Week 104 **and** achieving sustained reduction of OCS dose	CRR (see definition of CRR above) Sustained reduction of OCS dose (see definition above)
CCI [redacted]	[redacted]
CCI [redacted]	[redacted]
CCI [redacted]	[redacted]

CCI [redacted]	[redacted]
CCI [redacted]	[redacted]
Mean change in scores for overall disease activity from baseline to Week 52 (and to Week 104)	SLEDAI-2K
Mean change in score measures of non-renal disease activity from baseline to Week 52 (and to Week 104)	Non-renal components of SLEDAI-2K
Mean change in scores for overall disease activity from baseline to Week 52 (and to Week 104)	Physician Global Assessment (PGA)

Mean change in the Systemic Lupus International Collaborating Clinics/American College of Rheumatology (SLICC/ACR) Damage Index (SDI) score from baseline to Week 52 (and to Week 104)	SDI
Mean change in scores for patient-reported health status from baseline to Week 52 (and to Week 104)	Patient Global Assessment (PtGA)
To evaluate the immunogenicity of anifrolumab, pharmacokinetics (PK), pharmacodynamics (the PK and immunogenicity results will be reported in the clinical study report)	Anti-drug antibodies (ADA), anifrolumab concentration and PK parameters, 21-gene type I interferon (IFN) gene signature
To evaluate the pharmacokinetics of mycophenolic acid (MPA)	MPA concentration and PK parameters (if applicable)
Mean change in lupus serology from baseline to Week 52 (and to Week 104)	Anti-dsDNA antibodies, C3 and C4 complement levels
CCI [redacted]	[redacted]

[a] Required SOC is described in Section 1.2.1
[b] Allowed medication are described in Appendix A and Appendix B.
[c] Note: Spot UPCR will be used instead of 24-hour UPCR for the PRR and CRR classification, when evaluating time to achieve renal response modified to include OCS tapering requirement as well as for the PRR and CRR classification for flare assessment.

1.2 Study design

This is a Phase 2, multi-centre, multinational, randomised, double-blind (prior to the primary analysis), placebo-controlled study to evaluate the efficacy and safety of two IV treatment regimens of anifrolumab versus placebo, while taking protocol-specified SOC treatment with mycophenolate mofetil (MMF) and corticosteroids, in subjects with active biopsy-proven proliferative LN. The study will be performed in adult subjects, 18 to 70 years of age. The study duration is approximately 116 weeks (including screening, treatment and follow-up periods) for those subjects who meet the criteria for the second year extension period and approximately 64 weeks for those subjects who do not meet the criteria for the second year extension period.

In this study, approximately 150 subjects will be randomised in a 1:1:1 ratio to receive one of two IV anifrolumab dosing regimens or placebo, which will continue until Week 52, as follows:

- **Basic Regimen:** Anifrolumab 300 mg IV Q4W for 13 doses plus SOC or

- **Intensified Regimen:** Anifrolumab 900 mg IV Q4W for the first 3 doses followed by 300 mg IV Q4W for 10 doses plus SOC or

- **Placebo:** IV Q4W plus SOC

The primary objective will be evaluated at Week 52. After Week 52, eligible subjects will continue with their randomised treatment plus SOC until Week 104. Those subjects not eligible to continue with treatment based on the Week 52 assessments will complete the 8 weeks of additional safety follow-up to complete involvement in the study.

After all randomised subjects have completed the 52-week double-blind period the database will be soft-locked, and the **primary analysis** will be performed, including the assessment of the primary, secondary, exploratory and safety objectives. This analysis will also encompass all available data collected after the 52-week time point. At this stage the Sponsor and Sponsor's delegate who are not directly involved in the management of sites will be unblinded to randomised treatment, but the subjects and investigators will remain blinded (single-blind). The **end of study analysis** will be performed after all subjects have completed the second year extension period, and the database will be locked. One interim analysis may be conducted after approximately 80% of subjects have completed the Week 52 visit or would have completed if they had not discontinued the study. CCI [redacted]

IP will be administered as an IV infusion via an infusion pump over no less than 60 minutes for the first 3 doses (Visit 1, 2 and 3). Starting with Visit 4, IP can be administered as an IV infusion via an infusion pump over no less than 30 minutes.

Randomisation will be stratified using two factors:

- Results of IFN test at screening using a 4-gene type I IFN test (IFN test-high versus IFN test-low)

- 24-hour UPCR ≤3.0 mg/mg versus >3.0 mg/mg (based on 24-hour UPCR performed on a sample obtained within 14 days prior to the expected date of randomisation). Without the results of the second (stratification) sample, subjects cannot be randomised. Turn-around time for results from central laboratory is up to 7 days.

On rare occasion an extension of the 30-day screening window may be allowed if re-collection of the 24-hour UPCR or other required samples is necessary or if the results needed for randomisation are delayed.

Renal biopsy performed within 12 weeks prior to signing the informed consent form (ICF) or during the screening period will be considered as the screening biopsy. Biopsies will be evaluated locally and the local classification will be used to confirm the eligibility criteria.

The biopsy must reveal Class III (±class V) or class IV (±class V) LN according to the World Health Organisation (WHO) or 2003 International Society of Nephrology/Renal Pathology Society (ISN/RPS) classification. An external renal biopsy adjudication group will adjudicate renal biopsies post-randomisation.

The overall study duration is approximately 116 weeks:

A Screening Period: Up to 30 days

Treatment Period:

The total treatment period is up to 100 weeks (26 doses administered Q4W).

- **52 week double-blind treatment period:** IP will be administered Q4W from Week 0 to Week 48 for a total of 13 doses. The primary endpoint will be evaluated at Week 52.
- **Second year extension period:** At Week 52, eligible subjects will continue treatment with randomised IP (anifrolumab 300 mg or placebo) administration Q4W from Week 52 to Week 100 for a total of 13 doses. The last efficacy assessments will be performed at Week 104.

Follow-up: After the completion of the last IP treatment (Week 48 for subjects who are not participating in the second year extension period and Week 100 for subjects who participate in the second year extension period) subjects will continue in the study for another 12 weeks of safety follow-up after the last administration of IP.

1.2.1 Standard of care immunosuppressive treatment for lupus nephritis

Standard of care treatment for LN will consist of the combination of MMF and corticosteroids. From the day of randomisation onwards until the end of subject's participation in the study including the second year extension period, MMF will be supplied by the Sponsor and Sponsor's delegate to subjects. The target dose of MMF will be 2 gm/day during the first year of the study, where the dose is titrated to the target dose between randomisation and not later than Week 8. It is not mandatory for a subject to receive 2 gm/day of MMF if local treatment standards dictate a lower dose to be given (e.g., for constitutionally small subjects). During the 52-week double-blind treatment period, the minimum dose of MMF is 1.0 gm/day between Week 8 and Week 52. Adjustments of the dose due to suboptimal response, toxicity, or intolerability are allowed if needed. A maximum dose of 3.0 gm/day is allowed up to Week 24 for subjects with suboptimal response between Weeks 8 and 24, which needs to be reduced to ≤2.0 gm/day by Week 32. The dose of MMF has to be stable from Week 40 to Week 52.

In the second year, extension period, MMF dose should be ≤2.0 gm/day or ≤Week 52 dose, whichever is lower. If MMF dose is > 2.0 gm/day at Week 52, taper to ≤2 gm/day will be required by Week 60. Failure to do so will lead to withdrawal from IP. A decrease in MMF dose is allowed at the investigator's discretion between Week 52 to Week 92. The dose of MMF must be stable from Week 92 to Week 104.

Subjects may enter the study taking daily OCS at a maximum dose of 0.5 mg/kg/day of prednisone-equivalent not to exceed 40 mg/day. In addition, subjects will receive IV (or oral,

if applicable) methylprednisolone pulse 500 mg on the day of randomisation, prior to receiving the IP, unless they have received a methylprednisolone pulse of ≥500 mg within 10 days prior to randomisation. Subjects may receive one additional (optional) dose of IV (or oral, if applicable) methylprednisolone pulse (≤500 mg) for renal or extra-renal disease activity after the Week 0 (Day 1) visit up to and including the Week 8 visit. Methylprednisolone pulse can be administered on two consecutive days but the cumulative dose must not exceed 500 mg.

Details on other medications permitted during the study are given in Appendix A to Appendix C.

1.2.2 Oral corticosteroid burst and taper

From randomisation to Week 40, subjects may receive 1 burst of OCS for increased extra renal SLE disease activity or for non-SLE activity. Subjects who receive more than 1 burst during this time may continue in the study but will be considered non-responders (based on the conditions in Appendix B) for subsequent responder analyses (such as CRR), regardless of whether the burst was administered for increased SLE or non-SLE activity. No increase in OCS is allowed from Week 40 until Week 52 assessment. One burst and taper is allowed between Week 52 and Week 92. Further details are given in Appendix B.

1.2.3 Protocol-specified oral corticosteroid tapering for the double-blind treatment period

OCS tapering is required during the study with the goal of tapering OCS to prednisone-equivalent of ≤10 mg/day by Week 12 and prednisone-equivalent of ≤7.5 mg/day by Week 24. If subjects experience an increase in SLE disease activity upon tapering of OCS their dose may be returned to a dose less than or equal to the dose prior to the taper. The return to the pre-taper dose will not be considered an OCS "burst and taper" (see Section 1.2.2).

Subjects who are unable to taper OCS to ≤15 mg/day at Week 12 or <15 mg/day at Week 24 will be discontinued from IP treatment. However, subjects who exceed the maximum daily OCS dose at the Week 12 or Week 24 visits may continue to receive IP if the current dose is part of a protocol allowed temporary OCS dose (e.g., burst and taper) increase. Subjects who cannot be returned to their pre-increase dose within 14 days from the start of burst will have their investigational product discontinued at the next visit.

The subjects discontinued from IP treatment before Week 48 will be followed until Week 52 (or for a minimum of 12 weeks after last dose of investigational product for subjects for whom IP was discontinued within 12 weeks prior to Week 52).

Investigators will not be required, but may continue, to taper OCS dose beyond the target of 7.5 mg/day up to Week 40. Steroid tapering will not be permitted between Week 40 and Week 52. After the Week 52 primary endpoint assessments are completed additional OCS tapering can be done.

1.2.4 Management of oral corticosteroids during the second year extension period

The following rules apply to OCS during the second year extension period. Further details are provided in Appendix B.

- If OCS >7.5mg/day at Week 52, tapering to ≤7.5mg/day is required by Week 60. Failure to achieve ≤7.5 mg/day by Week 60 will lead to discontinuation from IP.
- OCS tapering goal ≤5 mg/day at Week 80. Tapering below 5 mg/day is allowed any time until Week 92.
- No change in OCS dose will be allowed from Week 92 to Week 104.
- One burst and taper is allowed between Week 52 and Week 92.

1.2.5 Discontinuation of IP

Criteria for discontinuing IP for worsening of LN or SLE (at any time):

Blinded IP will be discontinued in subjects who meet pre-defined criteria for worsening LN or SLE. Subjects may receive rescue treatment as clinically indicated. Subjects who had their IP discontinued will be followed according to the study schedule and should have a 24-hour UPCR measurement within 4 weeks of the decision to discontinue IP and before the start of a new treatment for LN. These subjects will be followed until Week 52 (or for a minimum of 12 weeks after last dose of IP for subjects for whom IP was discontinued within 12 weeks prior to Week 52). If discontinuation occurs during the second year extension period, subjects will be followed until Week 104 (or for a minimum of 12 weeks after last dose of IP for subjects for whom IP was discontinued within 12 weeks prior to Week 104).

- >30% decrease in eGFR compared to baseline due to LN **and** eGFR <60 mL/min/1.73 m^2 (on two independent samples) **or**
- Increase in renal or extra renal lupus activity requiring prohibited systemic immunosuppressive treatment (e.g. cyclophosphamide, rituximab, belimumab)
- Receipt of >1 methylprednisolone pulse after the day of randomisation
- Receipt of any methylprednisolone pulse after Week 8
- The IP will be discontinued if MMF is discontinued and another immunosuppressant is initiated

Criteria for discontinuation of IP for worsening of LN or SLE at Week 12 and Week 24:

- eGFR <75% of baseline and <60 mL/min/1.73 m^2 (on two independent samples) **or**
- Nephrotic range UPCR (confirmed by a second measurement at least two weeks after the first measurement):

  - Subjects with 24-hour UPCR ≤3 mg/mg at baseline will be withdrawn if 24-hour UPCR increases by >50% from baseline to >3.5 mg/mg
  - Subjects with 24-hour UPCR >3 mg/mg at baseline will be withdrawn if 24-hour UPCR at Week 24 >3.5 mg/mg and there is <60% improvement from baseline **or**
- Inability to adhere to corticosteroids requirements:
  - Inability to reduce OCS to ≤15 mg/day prednisone-equivalent at Week 12
  - Inability to reduce OCS to <15 mg/day by Week 24

Subjects who exceed the maximum daily OCS dose at the Week 12 or Week 24 visits may continue to receive IP if the current dose is part of a temporary increase in OCS dose (e.g., protocol-allowed burst and taper). Subjects who cannot be returned to their pre-increase dose within 14 days from the start of increase will have their IP discontinued at the next visit.

Criteria for discontinuing IP at Week 52:

Subjects not meeting the criteria to continue for the second year extension period will not receive any IP at Week 52 and will complete the study after completing the required follow-up visits.

Subjects meeting the following criteria may continue to receive blinded IP between Weeks 52 and 100.

(i) Meeting all of the following criteria based on the renal portion of the PRR definition:
- eGFR is:
  - ≥60 mL/min/1.73 m^2 or no confirmed decrease of eGFR from baseline of ≥20%
- Improvement in 24-hour UPCR:
  - For subjects with a baseline UPCR ≤3 mg/mg: <1.0 mg/mg
  - For subjects with a baseline UPCR >3 mg/mg: >50% improvement from baseline and ≤3.0 mg/mg

Renal discontinuation criteria must be confirmed in two separate samples. The second measurement should be at least 1 week after the first measurement.

(ii) No discontinuation of IP.

(iii) Negative HIV test after signed the Main ICF

Criteria for discontinuing IP at Week 52, Week 56 or Week 60:

- Failure to obtain Pap smear after Week 48 with result available at Week 60 at latest

- Pap smear result not meeting the eligibility criteria (see Appendix K of the Protocol) at Week 52, 56, or 60 of a Pap smear obtained after Week 48
- Failure to obtain Week 52 QuantiFERON test result by Week 56

Criteria for discontinuing IP at Week 60 and later:

- OCS >7.5 mg/day (except one OCS burst and taper)
- MMF >2 gm/day or >Week 52 dose, if Week 52 dose <2 gm/day

Subjects who discontinue the IP during the second year extension period will be followed until Week 104 (or for a minimum of 12 weeks after last dose of IP for subjects for whom IP was discontinued within 12 weeks of Week 104).

Criteria for discontinuation of IP for reasons other than lack of efficacy at any time during the study:

1. Subject decision: The subject is at any time free to discontinue treatment, without prejudice to further treatment. The primary reason should be documented as one of the following:
   (a) Subject is unable to comply with protocol-specified visits and/or procedures due to conflicts not related to clinical study

   (b) An AE or laboratory abnormality is of concern to the subject, but not clinically significant to physician

   (c) The subject is concerned about possibility of receiving placebo or ineffective treatment

   (d) Subject wishes to participate in another clinical study

   (e) Subject is interested in taking a treatment that is not allowed in this study

   (f) Subject perceives logistics at the clinical site to be unacceptable

   (g) Other reason

2. Lost to follow-up

3. Adverse event that, in the opinion of the Investigator of the Sponsor and Sponsor's delegate/Designee Medical Monitor, contraindicates further dosing with IP

4. Severe non-compliance with the study protocol

5. The Investigator or Sponsor and Sponsor's delegate/Designee Medical Monitor deems withdrawal as being in the subject's best interest

6. Pregnancy, positive pregnancy test, or subject expresses an interest in becoming pregnant

7. Isolated HBcAb positivity with HBV DNA above the LLOQ confirmed by the central laboratory

8. Positive HIV test

9. Receipt of any excluded medications

10. The use of restricted medications if the Sponsor and Sponsor's delegate/Designee Medical Monitor, in consultation with the Sponsor and Sponsor's delegate, determines the subject must be discontinued

11. A diagnosis of active TB, premature discontinuation of treatment for latent TB, or noncompliance with latent TB therapy. Note: Duration of treatment for latent TB should follow the local practice. If local practice is not defined, then Centers for Disease Control and Prevention (CDC) guidance for immunocompromised patients should be used.

12. Subject is unblinded by the Investigator

Additional restrictions related to concomitant medications are discussed in Appendix A and Appendix B.

Further details can be found in Section 3.9.4 of the clinical study protocol.

1.3 Number of subjects

A total of approximately 150 subjects will be randomised 1:1:1 to receive one of the two IV anifrolumab dosing regimens (Basic Regimen or Intensified Regimen) or placebo.

The sample size is based on the following assumptions:

- Reductions from baseline to Week 52 in 24-hour UPCR of 65% and 46% for the anifrolumab and placebo arms, respectively (i.e., ratios of 24-hours UPCR from Week 52 to baseline of 0.35 and 0.54, respectively), based on data presented in (Furie R, 2014).
- The change from baseline in the log-transformed 24-hour UPCR values follow a normal distribution with a standard deviation (SD) of 0.8, based on data from the anifrolumab Phase 2b study (CD1013)

Based on these assumptions, a sample size of 50 subjects per arm would result in an observed relative difference in the change from baseline to Week 52 in 24-hour UPCR of 0.65 (here expressed as the ratio, comparing anifrolumab to placebo), and a corresponding 95% confidence interval (CI) of (0.50, 0.85) comparing the pooled anifrolumab group (Basic and Intensified) with placebo group. Under the assumption that the interim analysis is performed

with 120 subjects, a 0.001 2-sided alpha will be spent, and the final analysis will be based on a 2-sided alpha of 0.0499 (East Version 6.4). If the study is not stopped for futility, this sample size provides approximately 87% power for the primary analysis with a 2-sided alpha of 0.0499 to reject the hypothesis of no effect (relative difference =1) for comparing the pooled anifrolumab treatment group with placebo. The minimal detectable relative difference in the change from baseline to Week 52 in 24-hour UPCR between the pooled anifrolumab treatment group versus placebo is approximately 0.76, corresponding to a reduction from baseline to Week 52 in 24-hour UPCR of 59% in the pooled anifrolumab group (ratio of 24-hour UPCR from Week 52 to baseline of 0.41).

Based on the above assumptions, if a futility analysis is performed with a boundary for the predictive power of 20% and supposing that 120 subjects will be included in the interim analysis and 150 subjects at the final analysis, the corresponding treatment effect boundary for relative difference in the change from baseline to Week 52 in 24-hour UPCR between the pooled anifrolumab treatment group versus placebo will be 0.81. This would result in an overall power of 86% (to continue the study after the futility analysis and reject the hypothesis of no effect with a 2-sided alpha of 0.0499).

2 ANALYSIS SETS

2.1 Definition of analysis sets

2.1.1 All subjects analysis set

This analysis set will comprise all subjects screened for the study and will be used for reporting of disposition and screening failures.

2.1.2 Full analysis set

The full analysis set will be used as the primary population for reporting efficacy and safety data. This comprises all subjects randomised into the study who receive at least one dose of IP and will be analysed according to randomised treatment (modified intent-to-treat [mITT] principle). Any major deviations from randomised treatment will be listed and considered when interpreting the safety data.

2.1.3 Pharmacokinetic analysis set

All subjects who received anifrolumab and who had at least one quantifiable serum PK observation post first dose will be included in the PK analysis dataset. All PK summaries will be based on this analysis set.

2.2 Violations and deviations

Protocol deviations that may greatly impact the completeness, accuracy, and/or reliability of the study data or that may significantly affect a subject's rights, safety, or well-being will be summarised.

The final list of protocol deviations will be finalised and documented prior to unblinding the study data and will be based on deviations as defined in the study Protocol Deviation Specification.

During the blind data review (BDR) meetings, protocol deviations will be classified as important or other depending on the impact of the deviation to the analysis and/or the impact to subjects. Only important protocol deviations will be listed and tabulated in the Clinical Study Report (CSR). All protocol deviations are regularly reviewed per the protocol deviation specification document.

3 PRIMARY AND SECONDARY VARIABLES

The baseline is defined as the last measurement prior to Day 1 dose administration. If the Day 1 value is missing or is invalid, the latest assessment prior to dose administration on Day 1 will serve as baseline. If only dates (not times) are recorded then if an assessment occurs on the same day as Day 1 dose administration, it will be considered to have occurred prior to dose administration.

Baseline OCS dose will be defined as the maximum daily OCS dose taken between Day 1 and Day 7, inclusive. If no OCS are taken by the subject on Day 1 or on or before Day 7, the baseline OCS dose will be 0 mg/day.

If not stated otherwise, change from baseline will be calculated as value at the respective time point minus value at baseline.

MMF at randomisation is defined to be the Day 1 MMF dose. If the Day 1 MMF dose is missing, or no MMF is taken by the subject on Day 1, the MMF dose at randomisation will be defined as the Day 2 MMF dose.

Renal biopsies will be adjudicated, however any changes in WHO or ISN/RPS classification values due to adjudication will only be used for sensitivity analyses to evaluate the relationship between biopsy findings and baseline disease activity or response to therapy. For all other analyses and summaries, the local classification of WHO or ISN/RPS will be used.

3.1 Primary outcome variable

The primary endpoint used to evaluate the effect of anifrolumab compared with placebo on LN disease activity is the relative difference in change from baseline to Week 52 in the 24-hour UPCR.

The time to 24-hour UPCR $\leq$ 0.7 mg/mg will be analysed as a supportive analysis and will be derived for subjects with baseline 24-hour UPCR > 0.7 mg/mg as the date of first 24-hour UPCR $\leq$ 0.7 mg/mg minus the date of first dose of IP. If the subject did not achieve 24-hour UPCR $\leq$ 0.7 mg/mg, the time to 24-hour UPCR $\leq$ 0.7 mg/mg will be censored at the earliest of the following:

- Date of withdrawal from the study;

- Discontinuation of IP in case of premature discontinuation of IP where the last IP dose is not in the Week 48 analysis visit window;
- Date of the Week 52 observation used in analysis (or the date of the last observation before Week 52 if the Week 52 observation is missing) in case of premature discontinuation of IP where the last IP dose is in the Week 48 analysis visit window;
- Date of the Week 104 observation used in analysis (or the date of the last observation before Week 104 if the Week 104 observation is missing);
- Database extract date.

3.2 Secondary outcome variables

3.2.1 Complete renal response

To evaluate the effect of anifrolumab compared with placebo on renal response in LN subjects, the proportion of subjects achieving CRR at Week 52 will be used. A subject achieves CRR if **all** of the following criteria are met:

- Estimated glomerular filtration rate (eGFR):
  - ≥60 mL/min/1.73 m^2 or no confirmed decrease of eGFR from baseline of ≥20%
- 24-hour UPCR ≤0.7 mg/mg
- No discontinuation of IP or use of restricted medication beyond the protocol-allowed threshold before assessment. Allowed restricted medications are defined in Appendix A and Appendix B.

The time to achieve CRR will be analysed as supportive analysis and will be derived as the date of first CRR minus the date of first dose of IP. If the subject did not achieve CRR, the time to achieve CRR will be censored at the earliest of the following:

- Date of withdrawal from the study;
- Discontinuation of IP in case of premature discontinuation of IP where the last IP dose is not in the Week 48 analysis visit window;
- Date of the latest of the Week 52 24-hour UPCR and eGFR observations used in analysis (or the date of the latest 24-hour UPCR or eGFR observation before Week 52 if the Week 52 observations are missing) in case of premature discontinuation of IP where the last IP dose is in the Week 48 analysis visit window;
- Date of the Week 104 24-hour UPCR and eGFR observations used in analysis (or the date of the latest 24-hour UPCR or eGFR observation before Week 104 if the Week 104 observations are missing);
- Database extract date.

3.3 Exploratory outcome variables

3.3.1 CCI

[redacted]

3.3.2 CCI

[redacted]

3.3.3 CCI

[redacted]

3.3.4 CCI

[redacted]

- CCI [redacted]

3.3.5 Alternative complete renal response

To evaluate the effect of anifrolumab compared with placebo on renal response in LN subjects when urine sediment is considered in the definition, the proportion of subjects achieving aCRR at Week 52 and Week 104 will be used. A subject achieves aCRR if all of the following criteria are met:

- eGFR:
  - ≥60 mL/min/1.73 m^2 or no confirmed decrease of eGFR from baseline of ≥20%
- 24-hour UPCR ≤0.7 mg/mg
- Inactive urine sediment (defined as <10 RBC/hpf)
- No discontinuation of IP or use of restricted medication beyond the protocol-allowed threshold before assessment. Allowed restricted medications are defined in Appendix A and Appendix B.

3.3.6 Graded alternative complete renal response

The effect of anifrolumab compared with placebo on renal response in LN subjects will also be further assessed using the difference in proportions of subjects achieving graded aCRR at Week 52 and of subjects achieving graded aCRR at Week 104. A subject achieves graded aCRR only if all following criteria are met:

- A decrease in 24-hour UPCR:
  - For subjects with baseline UPCR >3 mg/mg: UPCR ≤1 mg/mg
  - For subjects with baseline UPCR ≤3 mg/mg: UPCR ≤0.7 mg/mg
- eGFR:
  - ≥60 mL/min/1.73 m^2 or no confirmed decrease of eGFR from baseline of ≥20%
- Inactive urine sediment (defined as <10 RBC/hpf)
- No discontinuation of IP or use of restricted medication beyond the protocol allowed threshold before assessment. Restricted medications are defined in Appendix A and Appendix B.

3.3.7 Sustained reduction of oral corticosteroid dose

The effect of anifrolumab compared with placebo on the ability to reduce the OCS dose will be assessed in subjects with baseline OCS ≥20mg/day prednisone-equivalent. Sustained reduction of OCS dose is assessed using the difference in proportions of subjects meeting all the following criteria:

- Achieve an OCS dose of ≤7.5 mg/day prednisone-equivalent by Week 24
- Maintain an OCS dose ≤7.5 mg/day prednisone-equivalent from Week 24 to Week 52
- No discontinuation of IP or use of restricted medication beyond the protocol-allowed threshold before assessment. Allowed restricted medications are defined in Appendix A and Appendix B.

The proportion of subjects achieving OCS tapering requirement will also be assessed at Week 104:

- Achieve an OCS dose ≤5.0 mg/day prednisone-equivalent by Week 80
- Maintain an OCS dose ≤5.0 mg/day prednisone-equivalent from Week 80 to Week 104
- No discontinuation of IP or use of restricted medication beyond the protocol allowed threshold before assessment. Allowed restricted medications are defined in Appendix A and Appendix B.

The standardised area under the curve (AUC) of OCS dose up to Week 52, up to Week 104, and from Week 52 to Week 104 will be calculated as follows:

For each single daily dose, the duration of the single dose will be calculated as end date – start date + 1. If the start date is before Day 1, Day 1 will be used instead for calculating the AUC up to Week 52 and up to Week 104. If the start date is before the date of Week 52 (date of Visit 14), the date of Visit 14 will be used instead for calculating the AUC from Week 52 to Week 104. If the end date is after the date of Week 52 (date of Visit 14) or after the date of Week 104 (date of visit 29), the date of Visit 14 and Visit 29 respectively will be used instead for calculating the AUC up to Week 52 and Week 104. The AUC for each single dose will be derived by the daily dose (mg/day) multiplied with the duration (days). The AUC is the sum of the AUCs of the single doses. The AUC of OCS dose will only be calculated if all necessary data are available up to Visit 14 (i.e., daily prednisone equivalent OCS dose can be calculated) or Visit 29 respectively. For subjects who discontinued IP before Visit 14 or Visit 29 respectively, the AUC will be calculated up to the date of discontinuation of IP. If a subject does not receive any OCS dose (i.e., no corresponding medication documented) the AUC for this subject will be set to 0. The standardised AUC will be derived as:

- For standardised AUC up to Week 52: AUC divided by the available days (date of Visit 14 / date of discontinuation of IP – date of Day 1 + 1) multiplied by 364 (52 weeks).
- For standardised AUC from Week 52 to Week 104: AUC divided by the available days (date of Visit 29 / date of discontinuation of IP – date of Visit 14 + 1) multiplied by 364 (52 weeks).
- For standardised AUC up to Week 104: AUC divided by the available days (date of Visit 29 / date of discontinuation of IP – date of Day 1 + 1) multiplied by 728 (104 weeks).

3.3.8 CCI

[redacted]

3.3.9 CCI

[redacted]

3.3.10 CCI

[redacted]

CCI

[redacted]

3.3.11 CCI [redacted]

[redacted]

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

[redacted]

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

[redacted].

CCI [redacted]

3.3.12 CCI [redacted]

[redacted]

CCI [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]		

3.3.13 CCI [redacted]

[redacted]

3.3.14 SLEDAI-2K

The effect of anifrolumab compared with placebo on SLEDAI-2K score will be evaluated using the difference in mean change from baseline to Week 52 and to Week 104.

The SLEDAI-2K assessment (see Appendix F in the clinical study protocol) consists of 24 lupus-related items. It is a weighted instrument, in which descriptors are multiplied by a particular organ's "weight" and then these are totalled into the final score. The SLEDAI-2K score range is 0 to 105 points with 0 indicating inactive disease. The SLEDAI-2K will be calculated using a timeframe of 30 days in this study.

3.3.15 Non-renal SLEDAI-2K

Extra-renal disease activity will be assessed by using the SLEDAI-2K score (as defined in Appendix E in the clinical study protocol) without the renal components ("non-renal SLEDAI-2K"). The renal components that will not be included for this assessment are proteinuria, hematuria, pyuria and urinary casts. The score range for the non-renal SLEDAI-2K is therefore 0 to 89 with 0 indicating inactive disease. The effect of anifrolumab compared with placebo on non-renal SLEDAI-2K will be evaluated using the difference in mean change from baseline to Week 52 and Week 104.

3.3.16 Physician Global Assessment

The effect of anifrolumab compared with placebo on PGA will be evaluated using the difference in mean change in PGA from baseline to Week 52 and to Week 104.

PGA is a visual analogue scale (VAS) measuring the physician's overall assessment of average disease severity over the last 4 weeks and ranges from 0 (none) to 3 (severe).

PGA may be completed after the visit, to allow for laboratory results, as long as this is done prior to the next scheduled visit. Therefore, the date of the associated scheduled visit (Week 0 etc.) will be used to determine baseline (and subsequent visits), and not the actual date of the assessment.

3.3.17 SLICC/ACR Damage Index

The endpoint used to evaluate the effect of anifrolumab compared with placebo on irreversible damage in SLE subjects is the mean change in SDI global score (as defined in Appendix M in the clinical study protocol) from baseline to Week 52 and to Week 104. SDI will be listed by organ system.

The SDI is defined for 12 organ systems (possible scores): peripheral vascular (0 to 5), ocular (0 to 2), neuropsychiatric (0 to 6), renal (0 to 3), pulmonary (0 to 5), cardiovascular (0 to 6), gastrointestinal (0 to 6), musculoskeletal (0 to 7), skin (0 to 3), endocrine (diabetes) (0 to 1), gonadal (0 to 1) and malignancies (0 to 2). The SDI global score is the sum of the damage scores for all 12 organ systems. Damage over time can be stable or increase (but not decrease from the baseline value), theoretically to a maximum of 47 points.

3.4 Subject reported outcome variables

3.4.1 Patient Global Assessment

The difference between anifrolumab and placebo in the mean change from baseline in PtGA to Week 52 and to Week 104 will be assessed.

PtGA is a VAS completed by subjects asking a single-item question "Considering all the ways in which illness and health conditions may affect you at this time, please make a mark below to show how you are doing". The VAS measures from 0 mm (Very Well) to 100 mm (Very Poorly).

3.5 Assessment of study population

3.5.1 Demographic and baseline characteristic variables

Demographic characteristics (including geographic region [Asia Pacific: Australia, Republic of Korea, Taiwan; Europe: Belgium, France, Germany, Hungary, Italy, Poland, Russian Federation, Serbia, Spain, South Africa, United Kingdom; Latin America: Argentina, Mexico, Peru; North America: United States], age, sex, ethnicity and race) and baseline characteristics (including height, weight, body mass index [BMI] and disease characteristics [including SLEDAI-2K, PGA, SLICC/ACR, OCS use, OCS dose, time from LN diagnosis to randomisation, 4-gene type I IFN status, antinuclear antibody [ANA], anti-dsDNA, C3, C4, 24-hour UPCR, spot UPCR, eGFR, inactive urine sediment and MMF dose at randomisation]) will be assessed.

3.5.2 Stratification factors at randomisation

Stratification factors are:

- Results of IFN test at screening using a 4-gene type I IFN test (high versus low)
- 24-hour UPCR ≤3.0 mg/mg versus >3.0 mg/mg (based on 24-hour UPCR performed on a sample obtained within 14 days prior to the expected date of randomisation). Without the results of the second (stratification) sample, subjects cannot be randomised. Turn-around time for results from central laboratory is up to 7 days.

On rare occasion an extension of 30-day screening window may be allowed if the re-collection of the 24-hour UPCR or other required samples is necessary or if the results needed for randomisation are delayed. The number of subjects in each strata will be summarised.

3.5.3 Surgical and medical history

Surgical and medical histories will be coded using the latest version of the Medical Dictionary for Regulatory Activities (MedDRA). Medical history will be evaluated separately for past and current conditions as given in the electronic case report form (eCRF).

3.5.4 Prior and concomitant medications

All medications will be coded using the latest version of the WHO Drug Dictionary.

Any medications taken by the subject prior to the first dose date of IP will be considered prior medication. Any medication taken by the subject at any time between the date of the first dose (including the date of the first dose) of IP up to Week 104, inclusive, will be considered concomitant medication. Any medication that started prior to the first dose of IP and ended after the first dose up to Week 104 or is ongoing will be considered as both prior and

concomitant medication. Medications that start during the follow-up period will only be displayed in listings.

In the case that the end date of a concomitant medication (including sponsor provided MMF) is the same as the start date of another record of the same ATC code and same administration route, the end date of the record with the earlier start date will be considered to be one day earlier than the recorded end date, to avoid double counting of medication use in one day.
If a medication cannot be classified as prior or concomitant due to missing or partial dates, the imputation rules as given in Appendix D will be applied.

3.5.5 Exposure of IP

The duration of exposure to the IP per subject is defined as the number of days between the start and the end dates of IP plus the dosing frequency time:

Duration of exposure (days) = (Last dosing date + 28 days) – first dosing date + 1.

The total number of infusions will be counted per subject. Furthermore, the number of subjects with an infusion will be assessed in 4-weekly categories (i.e. 4 weeks, 8 weeks, 12 weeks, …, 104 weeks).

The time to discontinuation of IP is the same as the duration of exposure.

3.6 Safety variables

The following safety data will be collected: vital signs, physical examination, 12-lead ECG, haematology, clinical chemistry, urinalysis, C-SSRS (up to Week 52), PHQ-8 (up to Week 52), SLEDAI-2K based Flare Assessment Instrument, tuberculosis questionnaire, assessment of Cushingoid features, assessment of cardiovascular risk (fasting lipid profile) and reported AEs (including AESIs).

Any overdose report information collected will be listed only.

Change from baseline to each post-treatment time point where scheduled assessments were made will be calculated for relevant measurements.

If not stated otherwise, on-treatment values are defined as values with an assessment date after the first administration of IP and on or before the subject's Week 104 visit (Week 52 for C-SSRS and PHQ-8), or last visit if they withdraw early from the study.

3.6.1 Adverse events

Adverse events experienced by the subjects will be collected throughout the entire study, including the second year extension period, and will be coded using the latest version of MedDRA.

Adverse event data will be categorised according to their onset date into the following study periods:

- AEs occurring during screening:
  An AE during screening is defined as an AE with a date of onset ≥ date of the screening visit and date and time of onset < date and time of the first dose of IP.
  AEs occurring during screening will only be listed.
- AEs occurring during treatment:
  An AE during treatment is defined as an AE with a date and time of onset ≥ date and time of first dose of IP and ≤ date of last dose of IP + 28 days.
- AEs occurring during follow-up:
  An AE during follow-up is defined as an AE with a date of onset > date of last dose of IP + 28 days and ≤ date of last dose of IP + 84 days.
- AEs occurring after follow-up:
  An AE after follow-up is defined as an AE with a date of onset > date of last dose of IP + 84 days.
  AEs occurring after follow-up will only be listed.

If an AE has a missing onset date, then unless the stop date of the AE indicates otherwise, this will be considered as an AE during treatment. Similarly, if an AE has a partial onset date, then unless the partial onset date or the stop date indicates otherwise, this will be considered an AE during treatment.

Adverse events with missing intensity will be assumed to be severe. Events with missing relationship to study medication per the investigator will be assumed to be related. If no information about seriousness is available, the AE will be considered serious.

An acute AE is defined as an AE with a date of onset on any day of administration of IP and a time of onset on or after the start of administration of IP or a date of onset occurring on the subsequent day. If an AE has a missing onset date, then unless the stop date of the AE indicates otherwise, this will be considered as an acute AE. Similarly, if an AE has a partial onset date, then unless the partial onset date or the stop date indicates otherwise, this will be considered acute AE.

Adverse events during treatment will also be presented by time intervals of the first onset of the event. For this analysis, repeated events with the same preferred term will not be considered (i.e. if a subject has more than one event with the same preferred term, only the event with the earliest date of onset will be used). For partial or missing dates, the rules as described above will be used. The following time intervals (defined as exact weeks, independent of the study visits and with no visit windowing) are defined:

- Day 1 to < Week 12:
  AEs with date of onset ≥ date of first administration of IP and < date of first administration of IP plus 84 days

- Week 12 to < Week 24:
  AEs with date of onset ≥ date of first administration of IP plus 84 days and < date of first administration of IP plus 168 days
- Week 24 to < Week 36:
  AEs with date of onset ≥ date of first administration of IP plus 168 days and < date of first administration of IP plus 252 days
- Week 36 to < Week 48:
  AEs with date of onset ≥ date of first administration of IP plus 252 days and < date of first administration of IP plus 336 days
- Week 48 to < Week 60:
  AEs with date of onset ≥ date of first administration of IP plus 336 days and < date of first administration of IP plus 420 days
- Week 60 to < Week 72:
  AEs with date of onset ≥ date of first administration of IP plus 420 days and < date of first administration of IP plus 504 days
- Week 72 to < Week 84:
  AEs with date of onset ≥ date of first administration of IP plus 504 days and < date of first administration of IP plus 588 days
- Week 84 to < Week 96:
  AEs with date of onset ≥ date of first administration of IP plus 588 days and < date of first administration of IP plus 672 days
- Week 96 to < Week 100:
  AEs with date of onset ≥ date of first administration of IP plus 672 days and < date of first administration of IP plus 756 days
- ≥ Week 100:
  AEs with date of onset ≥ date of first administration of IP plus 756 days

The event rate per 100 subject years is defined as the number of subjects with an event divided by the sum of exposure time in days (e.g. date of last dose of IP + 28 days - day of first dose of IP + 1 day for summary of AEs during treatment) for all subjects in the analysis set multiplied by 365.25 days/year multiplied by 100.

The time to first onset of herpes zoster during treatment will be derived as date of first onset of herpes zoster – date of first administration of IP + 1. AEs with an onset date before the date of first administration of IP and AEs with an onset after 28 days after the date of last administration of IP will not be considered for the time to first onset of herpes zoster during treatment. If a subject has no herpes zoster during treatment, the time to first onset will be censored at the date of last administration of IP + 28 days.

For herpes zoster, an alternative event rate per 100 subject years will be derived as the number of subjects with herpes zoster divided by the sum of time at risk in days for all subjects in the analysis set multiplied by 365.25 days/year multiplied by 100. The time at risk is defined as time (including start and end date) from start of period (e.g. date of first administration of IP for events during treatment) to the date of first event, discontinuation or completion of the

study, or the database extract date, whatever comes first. This alternative event rate may also be calculated for other AESIs if suggested by data. This will be discussed during the BDR meeting and the decision will be made before unblinding the data.

3.6.1.1 Other significant adverse events

During the evaluation of the AE data, a medically qualified expert will review the list of AEs that were not reported as serious AE (SAEs) or AEs leading to discontinuation.

Based on the expert's judgment, significant AEs of particular clinical importance may, after consultation with the AstraZeneca Global Patient Safety Physician, be considered other significant AEs (OAEs) and reported as such in the CSR.

Examples of these are marked haematological and other laboratory abnormalities, and certain events that led to intervention (other than those already classified as serious), dose reduction, or significant additional treatment.

3.6.1.2 Adverse events of special interest

Adverse events of special interest will be recorded as such in the eCRF. MACE events will be classified according to the assessments of the Cardiovascular Event Adjudication Committee. AESIs are defined as follows:

- Non-opportunistic serious infections
- Opportunistic infections
- Anaphylaxis
- Malignancy
- Herpes zoster
- Tuberculosis (TB), including latent TB (LTB)
- Influenza
- Vasculitis (non-SLE)
- MACE (as adjudicated)

AESIs are further defined in the clinical study protocol. An AESI that meets one of the seriousness outcomes listed in Section 6.2 of the clinical study protocol will also be categorised as an SAE.

3.6.2 Laboratory variables

Parameters of haematology, clinical chemistry, urinalysis (outlined in Table 4, Section 5.4 of the clinical study protocol) and of fasting lipid profile (high density lipoprotein cholesterol, low density lipoprotein cholesterol, and triglycerides) will be examined.

Laboratory data will be reported in SI units. Changes from baseline in haematology, clinical chemistry and lipid profile variables will be calculated.

Absolute values will be compared to the reference range as given in Appendix F and classified as low (below range), normal (within range or on limits) or high (above range). All values (absolute and change) falling outside the reference ranges will be flagged.

Treatment emergent laboratory/vital signs changes (TELVC) will be defined for on-treatment values according to the reference ranges given in Appendix F.

Urinalysis data that is continuous will be reported and summarised as for haematology and clinical chemistry data. Categorical urinalysis data will be categorised as negative (0), trace (trace), or positive (+, ++, +++, or >+++) at each time-point. Treatment-emergent changes will also be assessed. Treatment-emergent changes of urinalysis data are defined as

- Negative/+ at baseline to ++, +++, ++++ at any on-treatment value OR
- Increase of from baseline of at least ++ at any on-treatment value.

For the liver function tests: aspartate transaminase (AST), alanine transaminase (ALT), alkaline phosphatase (ALP), gamma glutamyl transferase (GGT) and total bilirubin, the multiple of the upper limit of the normal (ULN) range (see Appendix F) will be calculated for each data point as follows:

Multiple=Value/ULN (i.e. if the ALT value was 72 IU/L (ULN = 36) then the multiple would be 2)

Subjects meeting any of the following biochemical criteria for Hy's law (potential Hy's Law) at any point during the study will be flagged:

- AST ≥3x ULN
- ALT ≥3x ULN
- Total bilirubin ≥2x ULN

3.6.3 Electrocardiograms

A 12-lead ECG tracing will be performed locally at screening, Week 52 and Week 104. The Investigator or qualified designee will review the ECG for clinically significant acute changes. The following categories used for analysis:

- Normal,
- Abnormal, not clinically significant
- Abnormal, clinically significant.

The number of subjects in each category will be presented.

3.6.4 SLEDAI-2K-based flare assessment instrument

The SLEDAI-2K based Flare Assessment Instrument will be used to assess the extra-renal flares. Flares will be categorised as mild/moderate or severe and defined as any one criterion

in either the mild/moderate and/or severe flare categories. New or worsened manifestation should only be reported for manifestations of SLE. The distinction between mild and moderate flare will be based on investigators' judgement. The SLEDAI-2K based Flare Assessment Instrument consists of the following parts:

- Mild/moderate flare
    - Change in non-renal components of the SLEDAI-2K
    - New/worse manifestations of SLE
    - PGA
- Severe flare
    - Change in non-renal components of the SLEDAI-2K
    - New/worse manifestations of SLE
    - Hospitalisation
    - PGA

See Appendix M in the clinical study protocol for further details. Flares will be scored in comparison to the previous study visit (i.e. over the past 30 days) and will only include findings which, in the opinion of the investigator, are due to SLE disease activity within that timeframe.

3.6.5 Physical examination

Complete physical examinations will be performed and will include an assessment of the following components: general appearance, head and neck, skin, lymph nodes, thyroid, musculoskeletal/extremities, cardiovascular, respiratory, abdomen, neurological, genital/rectal and breast.

A focused physical examination will also be performed at specific on-treatment visits and this will include an assessment of the organ systems required to complete protocol-specified SLEDAI-2K assessment tool. Additional assessments will be done as clinically indicated. Abnormal findings will be recorded as part of medical history, AE, SAE, AESI, or lupus activity, as appropriate.

3.6.6 Vital signs

Vital signs will be obtained at each visit and measured several times around each infusion of IP as follows:

- Subjects will be monitored during the administration of the IP and for at least two hours after the first 4 infusions (Weeks 0, 4, 8, and 12). If there are no safety concerns, for subsequent infusions subjects will be monitored during administration of the IP and for a minimum of one hour after completion of the IV infusion thereafter (Week 16 to Week 100).
- Vital signs in a sitting position will be taken:

- Shortly before the infusion (up to 20 minutes before the beginning of the infusion)
- Every 15±5 minutes during infusion
- Immediately after completion of administration of IP, including post-dose saline flush (within 20 minutes after completion of the post-dose saline flush)
- Every 30±5 minutes after completion of IP administration (not including saline flush) for at least 2 hours after the first 3 doses (Week 0 to Week 8) of IP are administered, and for at least 1 hour, thereafter (Week 12 to 100)

Vital signs may also be taken more frequently if the investigator judges this to be appropriate.

The following vital signs will be examined:

- Pulse (beats per minute)
- Systolic blood pressure (mmHg)
- Diastolic blood pressure (mmHg)
- Respiration rate (breaths per minute)
- Body temperature (°C)

Changes from baseline will be calculated.

Where applicable, absolute values will be compared to the reference ranges given in Appendix G and classified as low (below range), normal (within range or on limits) or high (above range). All values (absolute and change) falling outside the reference ranges will be flagged.

On-treatment values will be classified as TELVC according to reference ranges given in Appendix G.

Weight (kg) and BMI (kg/m^2) will be measured at all study visits and examined using the difference in mean change from baseline longitudinally over time. BMI will be calculated at all visits using the associated visit weight measurement and the height measured at screening.

3.6.7 Cushingoid features

Subjects will be assessed for Cushingoid features such as moon face, buffalo hump, purple or violaceous striae, central obesity, hirsutism, acne, easy bruising, and fragile skin which will be captured separately within the study eCRF to evaluate whether resolution of same can occur overtime with OCS reduction. These features will be presented by visit.

3.6.8 Tuberculosis monitoring and questionnaire

Regular testing for TB/LTB will be assessed during the study. In addition, to aid in the early detection of new or reactivated TB, a TB questionnaire will be used to evaluate subjects for signs and symptoms of TB at every visit prior to receiving IP. Any confirmed incidence of TB/LTB will be recorded as an AESI (see Section 3.6.1.2). All TB test results will be included as part of the laboratory outputs.

3.6.9 C-SSRS

The C-SSRS is an assessment tool that evaluates suicidal ideation and behaviour. It assesses the lethality of attempts and other features of ideation (frequency, duration, controllability, reasons for ideation, and deterrents), all of which are significantly predictive of completed suicide (see Appendix Q in the clinical study protocol for further details).

The following outcomes are C-SSRS categories and have binary responses (yes/no). The categories have been re-ordered from the actual scale in an increasing order of severity from 1 to 10 to facilitate the definitions of the comparative variables.

- Category 1 – Wish to be Dead
- Category 2 – Non-specific Active Suicidal Thoughts
- Category 3 – Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
- Category 4 – Active Suicidal Ideation with Some Intent to Act, without Specific Plan
- Category 5 – Active Suicidal Ideation with Specific Plan and Intent
- Category 6 – Preparatory Acts or Behaviour
- Category 7 – Aborted Attempt
- Category 8 – Interrupted Attempt
- Category 9 – Actual Attempt (non-fatal)
- Category 10 – Completed Suicide

The Suicidal Ideation or Behaviour Score will be derived from the C-SSRS categories as the maximum suicidal ideation or behaviour category (1-10 on the C-SSRS) present at the assessment. The score will be derived at each assessment for each subject. Non-suicidal self-injurious behaviour will be assigned if no ideation or behaviour is present.

Composite variables based on the above re-ordered categories are defined for assessments during screening, during treatment, and during follow-up, respectively (with the same definitions for the study periods as given for AEs in Section 3.6.1), as follows:

Suicidal ideation: A “yes” answer at any time in the respective study period to any one of the 5 (re-ordered) suicidal ideation questions (Categories 1-5) on the C-SSRS.

Suicidal behaviour: A “yes” answer at any time in the respective study period, to any one of the 5 (re-ordered) suicidal behaviour questions (Categories 6-10) on the C-SSRS.

No suicidal ideation or behaviour: No “yes” answer at any time in the respective study period to any one of the 10 (re-ordered) suicidal ideation and behaviour questions (Categories 1-10) on the C-SSRS.

The total number of attempts, total number of interrupted attempts, and total number of aborted attempts will be derived for the different study periods by summing up all respective attempts during screening, during treatment and during follow-up, respectively.

Occurrence of suicidal behaviour and ideation, based on the C-SSRS, from baseline up to Week 52 will be presented.

3.6.10 Personal Health Questionnaire Depression Scale-8

The PHQ-8 assesses symptoms of depression over the last 2 weeks by asking eight questions related to depression (see Appendix R in the clinical study protocol for more details).

The mean change from baseline in PHQ-8 total score will be assessed by visit up to Week 52 as well as at Week 60. The PHQ-8 total score will be derived as the sum of the 8 single item scores, each ranging from 0 (not at all) to 3 (nearly every day). If more than one item is missing, then the value of the scale is set to missing.

3.7 Pharmacokinetics, pharmacodynamics, and immunogenicity

3.7.1 Immunogenicity variables

Anti-drug antibodies (ADA) assessments will be conducted utilising a tiered approach (screen, confirm, titre).

The proportion of subjects with ADA responses during the treatment and follow-up periods will be summarised. Anti-drug antibody response categories summarised will include (but are not limited to):

- ADA prevalence: a positive ADA result at any visit including study baseline and all post-baseline measurements
- ADA negative: does not meet criteria for ADA positive
- Subjects who are ADA positive at baseline only
- Subjects who are ADA positive post-baseline only
- Subjects who are ADA positive at both baseline and post-baseline
- Subjects who are ADA positive by visit
- Subjects who are ADA positive at a post-baseline measurement for the first time (for subjects who are ADA negative at baseline) by visit
- ADA persistently (4 months) positive: positive at ≥2 post-baseline assessments, with ≥16 weeks between first and last positive or positive at last post-baseline assessment (for subjects who are ADA negative at baseline)
- ADA prolonged (6 months) persistently positive: positive at ≥ 2 post-baseline assessments, with ≥24 weeks between first and last positive or positive at last post-baseline assessment (for subjects who are ADA negative at baseline)
- ADA prolonged (12 months) persistently positive: positive at ≥ 2 post-baseline assessments, with ≥48 weeks between first and last positive or positive at last post-baseline assessment (for subjects who are ADA negative at baseline)
- ADA transiently (<4 months) positive: at least 1 post-baseline ADA positive assessment and not fulfilling the conditions of ADA persistently positive (for subjects who are ADA negative at baseline)

- ADA positive with titre ≤median of maximum titre. To calculate the median of maximum titre, take the maximum value for each subject, then calculate the median of these values across all subjects (for subjects who are ADA negative at baseline and ≥ 1 post-baseline ADA result available).
- ADA positive with titre > median of maximum titre (for subjects who are ADA negative at baseline and ≥ 1 post-baseline ADA result available).
- Neutralising antibody (nAb) incidence: post-baseline nAb positive result only (not positive at baseline)
- Prolonged (6 months) nAb positive: nAb positive at ≥2 post-baseline assessments, with ≥24 weeks between the first and last positive or positive at last post-baseline assessment (for subjects who are nAb negative at baseline).
- Prolonged (12 months) nAb positive: nAb positive at ≥2 post-baseline assessments, with ≥48 weeks between the first and last positive or positive at last post-baseline assessment (for subjects who are nAb negative at baseline).

For the presentation of ADA results at a single time point (e.g. baseline or by visit summaries), the corresponding titre summary will be based on the titres of the positive samples for that particular visit. For the presentation of ADA results across visits (e.g. any post-baseline), the corresponding titre summaries will be based on the maximum titre of all positive samples for each subject. Titres of positive measurements reported as ≤30 (minimum required dilution) will be imputed as 30.

The presence of nAb will be tested in all ADA-positive samples using a ligand binding assay. The following variables will be evaluated:

- Proportion of subjects with at least one non-missing ADA result (at any time) who are nAb positive at any time
- Proportion of subjects with at least one non-missing ADA result (at any time) who are nAb positive for the first time by visit
- Proportion of subjects with a non-missing ADA result (at a specific visit) who are nAb positive by visit

3.7.2 Pharmacokinetic variables

PK samples will be collected both pre and post-dose for specific infusions of investigational treatment. Due to the limited sampling schedule, the PK assessment for both anifrolumab and mycophenolate will be primarily based on the observed steady-state serum trough (pre-dose) concentrations, C_{trough}. For anifrolumab, maximum concentrations after the first and last dose of the first year will also be evaluated.

Individual concentrations will be reviewed for exclusion from descriptive statistics by identifying outliers and reviewing dosing information and sample collection times. Analysis to determine if the identified concentrations should be excluded includes visual inspection of PK-time profiles and comparison of descriptive stats with identified concentrations excluded and included.

In the second year extension period, only C_{trough} will be assessed.

3.7.3 Pharmacodynamic outcome variables

The IFN21 gene PD signature is calculated as the mean expression level of 21 IFN-inducible genes in lupus nephritis subjects compared to the expression level in control samples from healthy individuals. The IFN21 gene PD signature following treatment is compared to the IFN21 signature at baseline for each subject at multiple timepoints after treatment. The "percent suppression of fold change" will be presented as the "percent of baseline IFN21 PD signature" in all summaries, figures and listings.

This variable will be derived as defined in Appendix H.

3.7.4 Lupus Serology variables

The outcome variable for anti-dsDNA, C3 and C4 complement levels will be the mean change from baseline longitudinally up to Week 104 in subjects with abnormal complement level at baseline, defined as complement level below lower limit of normal, or elevated anti-dsDNA at baseline.

Elevated anti-dsDNA is defined as an anti-dsDNA value above the lower limit for a positive value.

4 ANALYSIS METHODS

4.1 General principles

4.1.1 Visit windows

For visit based analyses, the variables will be summarised based on the scheduled days with adjusted analysis-defined visit windows. The adjusted analysis-defined windows are summarised below:

Table 2 Visit windows

Adjusted Defined Windows Visit	Scheduled Study Day	Maximum Windows
Day 1	1	Study Day 1
Week 4	29	2≤ Study Day ≤42
Week 8	57	43≤ Study Day ≤70
Week 12	85	71≤ Study Day ≤98
Week 16	113	99≤ Study Day ≤126
Week 20	141	127≤ Study Day ≤154
Week 24	169	155≤ Study Day ≤182
Week 28	197	183≤ Study Day ≤ 210

Adjusted Defined Windows Visit	Scheduled Study Day	Maximum Windows
Week 32	225	211≤ Study Day ≤238
Week 36	253	239≤ Study Day ≤266
Week 40	281	267≤ Study Day ≤294
Week 44	309	295≤ Study Day ≤322
Week 48	337	323≤ Study Day ≤350
Week 52	365	351≤ Study Day ≤378
Week 56	393	379≤ Study Day ≤406
Week 60	421	407≤ Study Day ≤434
Week 64	449	435≤ Study Day ≤462
Week 68	477	463≤ Study Day ≤490
Week 72	505	491≤ Study Day ≤518
Week 76	533	519≤ Study Day ≤546
Week 80	561	547≤ Study Day ≤574
Week 84	589	575≤ Study Day ≤602
Week 88	617	603≤ Study Day ≤630
Week 92	645	631≤ Study Day ≤658
Week 96	673	659≤ Study Day ≤686
Week 100	701	687≤ Study Day ≤714
Week 104	729	715≤ Study Day ≤742
Week 108	756	743≤ Study Day ≤770
Week 112	784	771≤ Study Day ≤798

For assignment of data to time points using the visit windows, study day will be defined as (Date of assessment – date of first administration of IP) +1. Using this definition, the day of first dose of IP will be Day 1 and the scheduled visit date of Week 4 will be study day 29 (=28+1) for example.

If multiple readings are recorded within a single visit window, the following rules will be followed.

- If there are two or more observations within the same visit window, then the non-missing one closest to the scheduled visit will be used in the analysis.
- If two observations are equidistant from the scheduled visit, then the non-missing observation with the earlier collection date will be used in the analysis.
- If two observations are collected on the same day, then the non-missing one with the earlier collection time will be included in the analysis.

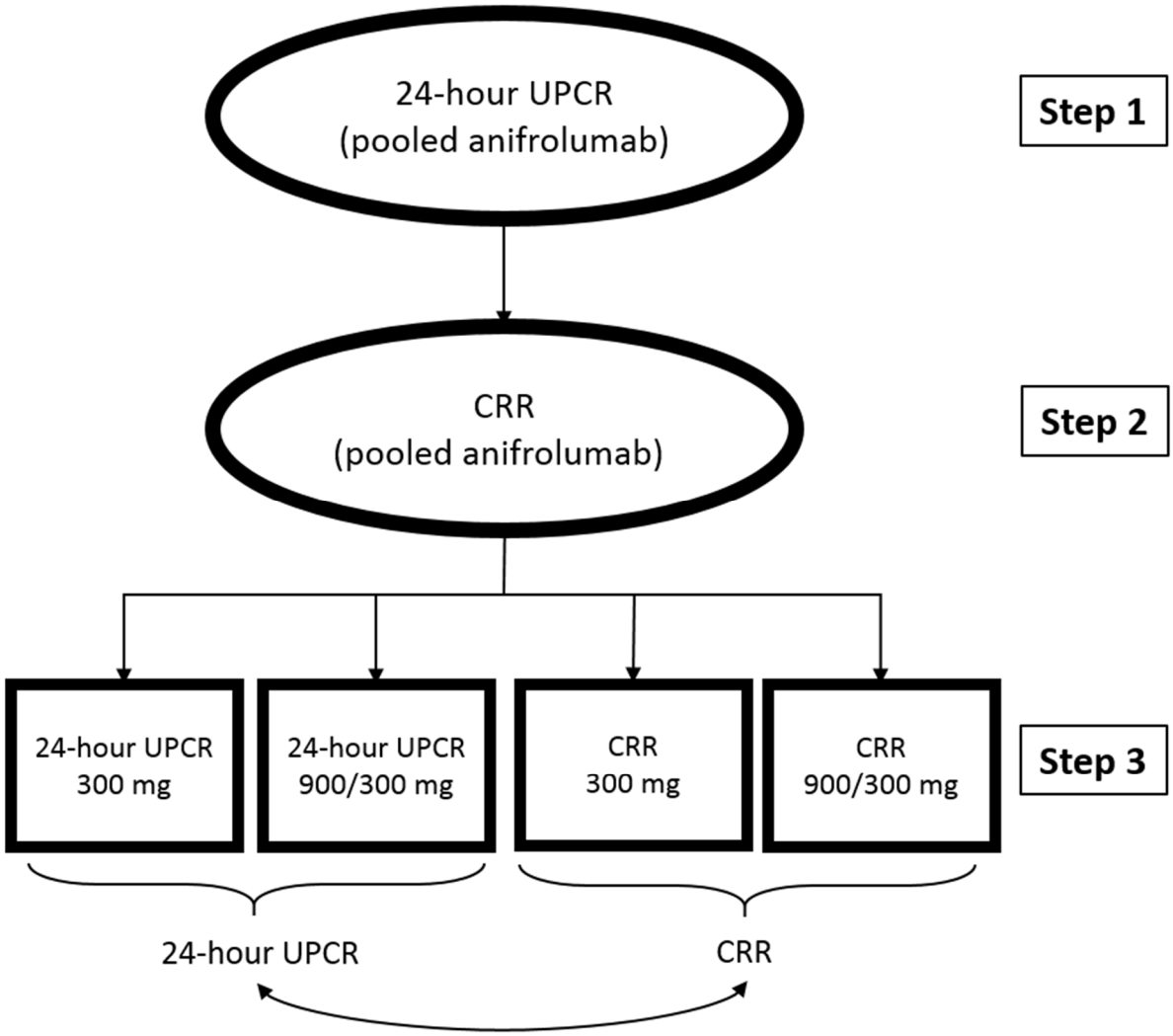
If a visit window does not contain any observations, then the data will remain missing.

For overall analyses not based on any particular study visit, all data will be listed and/or analysed, including any repeat or unscheduled visits, unless otherwise specified.

4.1.2 Statistical testing

While the primary and secondary objectives are defined based on the comparison between the pooled anifrolumab and placebo groups, the respective endpoints will also be tested for the individual anifrolumab regimens. The following hierarchical testing strategy (presented in Figure 1) will be used to provide strong control of the familywise error rate (FWER). As described in Section 5, if the interim analysis is performed, some alpha will be spent, and the alpha at final analysis will be calculated based on a Peto-Haybittle spending function and the number of subjects at each analysis. The description below assumes that a two-sided alpha of 0.0499 is being used at the final analysis. If the interim analysis is not performed, the hierarchical testing strategy at the final analysis will use a two-sided alpha of 0.05.

Figure 1 Multiplicity correction at final analysis



- Step 1: The pooled anifrolumab regimens will be tested versus placebo at a 2-sided 4.99% significance level with regards to 24-hour UPCR (primary endpoint).
- Step 2: If the null hypothesis is rejected at Step 1, then the pooled anifrolumab regimens will be tested versus placebo at a 2-sided 4.99% significance level with regards to CRR (secondary endpoint).
- Step 3: If the null hypothesis is rejected at Step 2, then both anifrolumab regimens are individually tested against placebo, each at 2.50% significance level, for both endpoints (24-hour UPCR and CRR). If both regimens can be rejected for either endpoint, then alpha can be recycled and both regimens can be tested at 4.99% significance level for the remaining endpoint.

At Steps 1 and 2, strong control is due to the sequence testing. At Step 3, alpha is controlled given that the pooled hypothesis is tested first and if false, the two hypotheses corresponding to the two regimens for that endpoint cannot both be true.

4.1.3 Presentation of results

All analyses will use SAS® version 9.3 or higher. Summary tables will be presented by treatment group, and will include a pooled group comprising of both anifrolumab regimens. Treatment groups will be presented in the following order: anifrolumab basic regimen, anifrolumab intensified regimen, all anifrolumab and placebo. For continuous endpoints, data only up to IP discontinuation will be analysed. If IP has not been reported as prematurely discontinued, available data up to and including Week 104 will be used for continuous endpoints analyses. Otherwise, the date of IP discontinuation will be determined as the date of last IP administration. However, all subjects who have an infusion of IP within the Week 48 analysis visit window will be determined to have completed treatment in the double-blind treatment period, and their Week 52 data will be included in the analyses. Additionally, subjects who are determined to have completed treatment and have a Week 52 visit will not be imputed as non-responders at Week 52 for discontinuation of IP. The treatment policy sensitivity analysis of 24-hour UPCR will be analysed using all available data. Data (including derived variables) will be presented in listings sorted by treatment group and subject number.

Unless otherwise noted, categorical data will be presented using counts and percentages with the denominator for percentages being the number of subjects in the analysis set by treatment group. Percentages will be rounded to one decimal place; except 100% which will be displayed without any decimal places. Percentages will not be displayed for zero counts.

Unless otherwise noted, continuous variables will be summarised using the number of observations (n), mean, standard deviation (SD), median, minimum, and maximum. The minimum and maximum values will be displayed to the same level of precision as the raw data, the mean and median to a further decimal place and the SD to two additional decimal places. Any CI's, CVs, geometric means and any quartile values will also be presented to one more decimal place than the raw data.

95% CIs will be presented for treatment comparisons. If a model is used to estimate the treatment difference, the corresponding CI according to the model will be presented. Otherwise, the unadjusted CI will be used. Nominal p-values may be presented for endpoints not included in the strategy for preserving the type I error rate, but these cannot be interpreted in terms of statistical significance. The presentation of p-values will be to four decimal places unless a p-value is less than 0.0001, in which case "<0.0001" will be displayed.

4.1.4 Missing data

The study was designed to reduce the risk for missing data as much as possible through the following measures:

- From randomisation to Week 40, the study design allows for one burst and taper of corticosteroid for increased SLE disease activity or non-SLE causes.

- The dose of MMF may be increased up to 3 gm/day in case of suboptimal response.
- Subjects will receive IV methylprednisolone 500 mg on Week 0 (Day 1), prior to receiving the IP.
- Subjects may also receive up to one additional (optional) dose of IV methylprednisolone (≤500 mg) for renal or extra-renal disease activity between Week 0 (Day 1) visit and Week 8 visit. This optional dose can be given over two consecutive days.
- Subjects who require additional bursts of OCS will still be allowed to remain in the study, but will be considered non-responders for subsequent responder analyses (such as CRR).

Subjects who discontinue IP will be followed up and asked to come to each visit for the scheduled assessments through the Week 52 visit (or for a minimum of 12 weeks after last dose of IP for subjects for whom IP was discontinued within 12 weeks prior to Week 52).

The following measures will be applied to the second year extension period to also reduce the risk for missing data:

- From Week 52 to Week 92, the study design allows for one burst and taper of corticosteroid.
- Subjects who require additional bursts of OCS during the second year extension period will still be allowed to remain in the study, but may be considered non-responders for subsequent responder analyses (such as CRR at Week 104).

Subjects who discontinue IP in the second year extension period will be followed up until Week 104 (or for a minimum of 12 weeks after last dose of IP for subjects for whom IP was discontinued within 12 weeks prior to Week 104).

Missing safety data will generally not be imputed. However, safety assessment values of the form of <x (i.e. below the lower limit of quantification) or >x (i.e. above the upper limit of quantification) will be imputed as x in the calculation of summary statistics but displayed as <x or >x in the listings.

Details about imputation of partial or missing dates are given in Appendix D. The imputation of single missing items for the derivation of a total score is described in Appendix E.

4.1.4.1 Observed case

A mixed model for repeated measures (MMRM) will be used for all continuous endpoints based on the observed data. MMRM allows the inclusion of repeated effects and the specification of an appropriate variance-covariance structure, which will model any missing data.

4.1.4.2 Last observation carried forward

For secondary and exploratory responder efficacy endpoints, any component with missing value will be imputed using last observation carried forward (LOCF) if only single (non-consecutive) visits are missing data for that component. In the event of two or more consecutive visits with missing data for the same component, the last observed post-baseline value prior to the first missing value will be carried forward only for the first missing value of each sequence, and subjects will be imputed as non-responders for the specific responder endpoint for the second visit on, where data is missing. The responder endpoint is derived based on the imputed values.

If one or more components of SLEDAI 2K and SDI are missing, then the LOCF approach will be used to impute the missing component score. However, if less than 50% of the respective components are available, the data will not be imputed. Additionally, in the event of two or more consecutive visits with missing data for the same component, the last observed post-baseline value prior to the first missing value will be carried forward only for the first missing value of each sequence, and the component score will remain missing for the second visit onwards.

4.1.4.3 Non-responder imputation

For the analysis of the secondary and exploratory efficacy responder endpoints (excluding time to modified renal response), subjects will be imputed as non-responders from time at which the criteria are met, in the following cases:

a) Medication criteria

Subjects who meet any of the following medication criteria:

- Subjects who receive any restricted medication (see Appendix A and Appendix B) after randomisation or beyond any protocol allowed thresholds will be considered a non-responder from the time the respective medication was administered up to and including Week 104.
- Subjects who meet any of the corticosteroid criteria (see Appendix B) will be considered a non-responder from the time the condition is met, up to and including Week 24, Week 52, Week 80 or Week 104, depending on the condition and when it is met.
- Subjects who meet any of the MMF criteria (see Appendix B) will be considered a non-responder from the time the condition is met, up to and including to Week 24, Week 52, Week 80 or Week 104, depending on the condition and when it is met.

b) Discontinued IP

Subjects who prematurely discontinue from IP, including subjects who withdraw from the study. All subjects who have an infusion of IP within the Week 48 analysis visit window will be determined to have completed treatment in the double-blind treatment period.

Non-responder imputation will be applied at all scheduled visits, including missing visits. If a subject does not have a response at a visit, they will be imputed as a non-responder (note that unless IP was discontinued or any of the medication criteria above are met, LOCF will be applied to missing values at the first missed visit in case of consecutive missed visits and the non-responder imputation will be applied to the response values determined after LOCF imputation. See Section 4.1.4.2 for more details on LOCF).

4.1.5 Sensitivity analysis

Sensitivity analyses where missing data is handled in different ways will be carried out for the primary endpoint. To examine the impact of missing data, including the impact of non-responder imputation due to permanent discontinuation of IP, tipping point analyses will be performed for endpoints specified in Section 4.2. These analyses will vary the assumptions about outcomes among the subsets on the treatment arms that discontinue IP early. Details of the sensitivity analyses are provided in Section 4.2 with each associated endpoint.

4.1.6 Examination of model assumptions

Underlying model assumptions for all statistical models will be checked with graphical displays (e.g. plots of residuals versus predicted values, histograms with normal density overlaid, quartile-quartile [Q-Q] plots showing the residual quantiles versus quantiles of a normal distribution) as appropriate. If any model assumptions are not met, appropriate (or alternative) data transformations or the use of non-parametric approaches will be discussed as soon as is appropriate. The decision about an appropriate approach will be made before unblinding the data.

4.1.7 CCI

[redacted]

4.2 Analysis methods

4.2.1 Analysis of the primary variable

The primary estimand is evaluating the efficacy on disease activity of the pooled anifrolumab groups relative to placebo in subjects with active proliferative LN. This is measured by the

relative difference in change from baseline to Week 52 in the 24-hour UPCR. The full analysis set (as defined in Section 2.1.2) will be used, and any data collected after IP discontinuation will be excluded from the analysis, in order to reflect the effect of the initially assigned and dosed IP (anifrolumab or placebo), prior to the occurrence of the intercurrent event of IP discontinuation, applying a "while on treatment" estimand strategy. The individual anifrolumab regimens will also be assessed but not as part of the primary analysis.

The estimand for all continuous endpoints (including the primary endpoint) follows the "while on treatment" strategy, which includes data only up to IP discontinuation. If IP has not been reported as prematurely discontinued, available data up to and including Week 104 will be used for continuous endpoints analyses. Otherwise, the date of IP discontinuation will be determined as the date of last IP administration. However, all subjects who have an infusion of IP within the Week 48 analysis visit window will be determined to have completed treatment in the double-blind treatment period, and their Week 52 data will be included in the analyses.

The primary analysis will be performed using an MMRM fitted to log-transformed data comparing the pooled anifrolumab group with the placebo group, with fixed effects for treatment group, visit, stratification factors and log-transformed 24-hour UPCR at baseline. An interaction term for visit and treatment will also be included in the model to allow the relationship to differ across treatment groups. Note that visit will be fitted as a repeated variable in the model. Denominator degrees of freedom will be estimated using the Kenward-Roger approximation. An unstructured correlation pattern will be used to estimate the variance-covariance of the within-subject repeated measures. In the case that the model does not converge using an unstructured covariance matrix, the heterogeneous Toeplitz covariance structure will be evaluated, followed by the heterogeneous autoregressive of first order, heterogeneous compound symmetry, and homogeneous compound symmetry covariance structures in the case of further non-convergence. The restricted maximum likelihood method will be used.

Separate MMRMs will be fitted to all data from baseline to Week 52 for the analyses of Week 12 to Week 52, and to all data from baseline to Week 104 for the analyses of Week 64 to Week 104.

Stratification factors are:

- Results of IFN test at screening using a 4-gene type I IFN test (high versus low)
- 24-hour UPCR ≤3.0 mg/mg versus >3.0 mg/mg (based on 24-hour UPCR performed on a sample obtained within 14 days prior to the expected date of randomisation). Without the results of the second (stratification) sample, subjects cannot be randomised. Turn-around time for results from central laboratory is up to 7 days. On rare occasion an extension of 30-day screening window is allowed if re-collection of the sample is necessary or the results needed for randomisation are delayed.

The 24-hour UPCR values will be log-transformed before analysis and the results will be back-transformed for presentation. Model assumptions will be checked as defined in Section 4.1.6. The estimated treatment effect, corresponding 95% CI, and two-sided unadjusted

p-value for the difference at Week 52 and nominal p-values for the difference at all other post-baseline visits will be presented.

The following sensitivity analyses will also be performed to assess the robustness of the primary analysis:

- Treatment policy analysis: Includes all data collected, irrespective of whether a subject discontinued IP.
- Biopsy analysis: Excludes any subjects whose adjudicated kidney biopsy results show them to be ineligible for study entry (i.e. subjects whose adjudicated kidney biopsy results are not class III or class IV)

Subgroup analyses will be performed as described in Section 4.3.

Summary statistics for 24-hour UPCR and change from baseline will be presented by week. Longitudinal presentations of results over time (i.e. for each post-baseline visit up to Week 104 where 24-hour UPCR is recorded) based on the same main analysis, with the corresponding 95% CI, will be created. Mean changes from baseline for 24-hour UPCR will be displayed in a longitudinal plot by treatment group only for the first year, where the estimated adjusted means with corresponding 95% CI at each visit will be displayed. A Kaplan-Meier plot will also be produced for time to 24-hour UPCR $\leq$ 0.7 mg/mg by treatment group for the whole study.

4.2.2 Analysis of the secondary variable

As a secondary endpoint, a composite endpoint is used to evaluate the effect of anifrolumab on renal response. For this analysis the estimand of interest is the difference in change from baseline at Week 52 in renal response between anifrolumab and placebo, to reflect the effect of the initially assigned IP (full analysis set). This estimand answers a clinically relevant question comparing the number of subjects able to both tolerate therapy sufficiently to remain on treatment to a given point in time and to achieve adequate response without further medication being required. Any missing data will be imputed as defined in Section 4.1.4.

The proportion of subjects achieving CRR in the anifrolumab pooled group will be compared with the placebo group using a CMH approach (Stokes ME, 2012) stratified by:

- Results of IFN test at screening using a 4-gene type I IFN test (IFN test-high versus test-low)
- 24-hour UPCR $\leq$3.0 mg/mg versus >3.0 mg/mg at screening (within 14 days prior to the expected date of randomisation). Without the results of the second (stratification) sample, subjects cannot be randomised. Turn-around time for results from central laboratory is up to 7 days. On rare occasion an extension of 30-day screening window is allowed if re-collection of the sample is necessary or the results needed for randomisation are delayed.

The individual anifrolumab regimens will also be assessed but not as part of the secondary analysis. The analysis can be described as follows:

- There are *nij* subjects in each stratum, where *i* is the stratum, and *j* is the treatment group. The number of responders is x_{ij}. The proportion of responders is denoted as $p_{ij} = x_{ij} / n_{ij}$.

  Strata with low counts will be pooled with adjacent strata. If a sub-stratum within the IFN-low stratum has less than 16 subjects (in the pooled treatment group, i.e. placebo pooled with the respective anifrolumab group that is tested), then the two IFN-low sub-strata will be pooled together. If a sub-stratum within the IFN-high stratum has less than 16 subjects (in the pooled treatment group), then the two IFN-high sub-strata will be pooled together. If the two IFN-low sub-strata are pooled together and the two IFN-high sub-strata are pooled together, and any of these strata has less than 16 subjects (in the pooled treatment group) then all strata will be pooled together.
- For each stratum, the difference in proportion of subjects achieving CRR is calculated as $d_i = p_{iA} - p_{iP}$, where *A* and *P* denote the different treatment groups (anifrolumab and placebo, respectively).
- Weights for each stratum, w_i, are calculated as $n_{iP} \times n_{iA} / (n_{iA} + n_{iF})$. The weighted difference is calculated as

$$WD = \frac{\sum w_i d_i}{\sum w_i}$$

- The SE of the weighted difference under the null hypothesis is given by

$$SE = \sqrt{\frac{\sum[w_i^2 Var(d_i)]}{(\sum w_i)^2}}$$

$$Var(d_i) = \frac{p_{i.}(1 - p_{i.})n_i}{w_i(n_i - 1)}$$

$$p_{i.} = \frac{x_{i.}}{n_{i.}} = \frac{x_{iA} + x_{iP}}{n_{iA} + n_{iP}}$$

- For deriving the CI for the weighted difference in proportions, a correction will be applied to the variance, providing a CI with more accurate coverage. This will be applied to all strata and is derived as below. The 95% confidence interval can be generated using the weighted difference ± $Z_{0.975}$ × SE.

$$Var(d_i) = \frac{p^*_{iA}(1 - p^*_{iA})}{n_{iA}} + \frac{p^*_{iP}(1 - p^*_{iP})}{n_{iP}}$$

$$p^*_{ij} = \frac{x_{ij} + 2}{n_{ij} + 4}$$

- The 95% confidence interval for the weighted proportion ($\sum_i w_i\, p_{ij}/w$) in a treatment group *j* can be generated using a normal approximation and independence between strata, where p_{ij}* are used as above

$$s_{ij}^2 = Var(p_{ij}) = p_{ij}^*(1 - p_{ij}^*)/n_{ij}$$

$$s_j^2 = \sum_i w_i^2\, s_{ij}^2/w^2$$

$$w = \sum_i w_i$$

- The value of the test statistic is calculated as $\frac{WD}{SE}$. The p-value from the two-sided test of no difference in treatment groups is calculated as $2\left(1 - Prob\left(\left|\frac{WD}{SE}\right|\right)\right)$, where *Prob(.)* is the distribution function of the standard normal distribution.

The estimated treatment effect (the difference in response rate for each anifrolumab regimen versus placebo) and corresponding 95% CI for the difference at Week 52 and Week 104 will be presented. In addition, the response rate and the corresponding 95% CI within each treatment group will be presented.

The following sensitivity analyses will also be performed to assess the robustness of this secondary outcome variable at Week 52 and Week 104:

- A tipping point analysis to examine the impact of missing data, including the impact of non-responder imputation due to permanent discontinuation of IP. The analysis for each treatment group, and the pooled anifrolumab group, will only be statistically interpreted if the nominal p-value from the analysis of CRR detailed above, for the respective comparison with placebo at the specified visit, is <0.05 (in favour of anifrolumab). This analysis will vary the assumptions about outcomes among the subsets of subjects on the treatment arms who permanently discontinued IP early, and can be described as:
    - The proportions of responders will be analysed using a Pearson's chi-squared test, i.e., disregarding the stratification factors that are used in the main analysis (CMH).
    - Subjects who permanently discontinued IP early having not received restricted medications prior to discontinuation are by definition imputed as non-responders. For this sensitivity analysis, these subjects will be considered as having had the potential of being responders if they continued into the study until Week 52.
    - At each step of the analysis one of these subjects switches from non-responder to responder, and the Pearson's chi-squared test is re-run. The results (statistical significance) will be presented for each anifrolumab treatment arm as well as the pooled anifrolumab group in a grid where the x-axis and the

y-axis represent the number of subjects that have had their values imputed as non-response values and are assumed as responders for the placebo arm and the active arm respectively. The region where the conclusion changes, will be considered as the tipping point.

- The grid will be divided into three regions, limited in the top by the expected number of subjects that have had their values imputed as non-response values in the active arm which would have been responders if not permanently discontinued IP early (based on the active response rate in the observed cases):
  - Likely – its right limit is the expected number of subjects that have had their values imputed as non-response values in the placebo arm which would have been responders if not permanently discontinued IP early (based on the placebo response rate in the observed cases);
  - Uncertain – its right limit is the expected number of subjects that have had their values imputed as non-response values in the placebo arm which would have been responders if not permanently discontinued IP early (based on the active response rate in the observed cases);
  - Unlikely – it is the region to the right of the uncertain region.

- Biopsy sensitivity analysis: Based on the main analysis it excludes any subjects whose adjudicated kidney biopsy results show them to be ineligible for study entry.

Subgroup analyses will be performed as described in Section 4.3.

Longitudinal presentations of results over time (i.e. for each post-baseline visit up to Week 104 where CRR based on 24-hour UPCR can be derived) based on the same analysis, with the corresponding 95% CI, will be created. Bar plots showing the proportion of subjects who are responders/non-responders for CRR at Week 24 and Week 52 by treatment group (including CIs and number of subjects included in the analysis) will be provided by 24-hour UPCR at screening (≤ 3 mg/mg and > 3 mg/mg). Additionally, the individual components of CRR will be summarised over time by treatment group.

The proportion of subjects achieving CRR based on 24-hour UPCR at Week 52 and at Week 104 will also be presented in a summary table alongside the proportion of subjects achieving CRR based on spot UPCR at Week 52 and at Week 104 together with a summary of the individual components.

In addition, the time to achieve CRR by treatment group will be presented in a Kaplan-Meier plot including the number of subjects at risk at each visit.

4.2.3 Analysis methods for exploratory outcome variables

All analyses of the exploratory outcome variables will be conducted with the full analysis set.

4.2.3.1 CCI [redacted]

[redacted]

[redacted]

4.2.3.2 CCI [redacted]

[redacted]

4.2.3.3 Oral corticosteroid management

The exploratory endpoint used to evaluate the effect of the anifrolumab regimens versus placebo on the ability to achieve a sustained reduction of OCS dose is the difference in proportions of subjects able to taper OCS to ≤7.5 mg/day prednisone-equivalent by Week 24 and maintain this dose through Week 52. The effect of the anifrolumab regimens versus placebo will also be evaluated on the ability to achieve a sustained reduction of OCS dose using the difference in proportions of subjects able to taper OCS to ≤5.0 mg/day prednisone-equivalent by Week 80 and maintain this dose through Week 104.

The same CMH approach as described in Section 4.2.2 will be used to estimate the treatment differences between the anifrolumab regimens and placebo at Week 52 and at Week 104. Any missing data will be imputed as defined in Section 4.1.4.

Subgroup analyses will be performed as described in Section 4.3.

Reduction of OCS dose will also be assessed at Week 24 by examining the proportions of subjects able to taper OCS to ≤7.5 mg/day prednisone-equivalent by Week 24, using the same analysis approach. Reduction of OCS dose will also be assessed at Week 80 by examining the proportions of subjects able to taper OCS to ≤5.0 mg/day prednisone-equivalent by Week 80, using the same analysis approach.

In order to examine the impact of missing data, a tipping point analysis as described in Section 4.2.2 will also be performed for the sustained reduction of OCS dose at Week 52 and Week 104. The analysis for each treatment group, and the pooled anifrolumab group, will only be statistically interpreted if the nominal p-value from the main analysis of sustained reduction of OCS dose detailed above, for the respective comparison with placebo at the specified visit, is <0.05 (in favour of anifrolumab).

Daily OCS doses and changes from baseline will be presented by descriptive statistics by visit for all subjects.

Additionally, the daily OCS dose will be displayed graphically by a longitudinal plot presenting the means and corresponding SDs. The number of subjects will also be included in the graph.

Shift plots (scatter plot) for the daily OCS dose presenting baseline values versus Week 24 values, baseline values versus Week 52 values and Week 52 values versus Week 104 values will also be provided. A diagonal line indicating no change will be included in this shift plot.

Furthermore, the standardised AUC up to Week 52, up to Week 104 and from Week 52 to Week 104 will be summarised by treatment group for all subjects.

4.2.3.4 Renal response modified to include oral corticosteroid tapering requirement

The renal response endpoints CRR and PRR will be modified to include an OCS tapering requirement and used to evaluate the effect of anifrolumab on renal response.

The same CMH approach as described in Section 4.2.2 will be used to estimate the treatment differences between the anifrolumab regimens and placebo at Week 52 and at Week 104. Any missing data will be imputed as defined in Section 4.1.4. Modified CRR and PRR will also be assessed at Week 24 by examining the proportions of subjects who achieve CRR/PRR and are also able to taper OCS to ≤7.5 mg/day prednisone-equivalent by Week 24, using the same analysis approach.

A bar plot showing the proportion of subjects who achieve modified CRR or PRR by treatment group (including CIs and number of subjects included in the analysis) will be provided for both Weeks 24 and 52.

4.2.3.5 CCI

[redacted]

4.2.3.6 CCI

[redacted]

4.2.3.7 CCI

[redacted]

4.2.3.8 CCI

[redacted]

[redacted]

[redacted]

4.2.3.9 CCI [redacted]

[redacted]

[redacted]

4.2.3.10 SLEDAI-2K

Change from baseline in SLEDAI-2K will be evaluated using the same summaries and analysis approach except SLEDAI-2K data will not be log transformed, as specified in Section 4.2.3.5, with the appropriate substitution for the baseline value.

4.2.3.11 Non-renal SLEDAI-2K

Change from baseline in non-renal SLEDAI-2K will be evaluated using the same summaries and analysis approach except non-renal SLEDAI-2K data will not be log transformed, as specified in Section 4.2.3.5, with the appropriate substitution for the baseline value.

4.2.3.12 Physician global assessment

Change from baseline in PGA will be evaluated using the same summaries and analysis approach except PGA data will not be log transformed, as specified in Section 4.2.3.5, with the appropriate substitution for the baseline value.

4.2.3.13 SLICC/ACR Damage Index

Change from baseline in SDI will be evaluated using the same analysis approach and summaries (except that SDI data will not be log-transformed) as specified in Section 4.2.3.5, with the appropriate substitution for the baseline value. SDI will also be analysed by baseline SDI (≥ 1 vs 0).

4.2.4 Subject reported outcome variables

4.2.4.1 Patient Global Assessment

Change from baseline in PtGA will be evaluated using the same summaries and analysis approach except PtGA data will not be log transformed, as specified in Section 4.2.3.5, with the appropriate substitution for the baseline value.

4.2.5 Presentation of study population

4.2.5.1 Subject disposition

The number and percentage of subjects randomised, randomised but not treated with IP, and not randomised will be summarised by treatment group for the all subjects analysis set.

The number and percentage of subjects completing the double-blind treatment period of the study up to and including Week 52 and the number and percentage of subjects withdrawing from the study during the double-blind treatment period including reason for withdrawal, will be summarised by treatment group for the full analysis set .

The number and percentage of subjects completing treatment with IP during the double-blind treatment period up to and including Week 48 (i.e. with last administration of IP at Visit 13), the number and percentage of subjects permanently discontinuing the treatment with IP during the double-blind treatment period including reason for withdrawal, and the number and percentage of subjects who discontinued the treatment with IP after completing treatment with IP in the double-blind treatment period and did not enter the second year extension period including reason for withdrawal, will be summarised by treatment group for the full analysis set.

The number and percentage of subjects continuing into the second year extension period of the study from Week 52 up to and including Week 112 will be summarised by treatment group for the full analysis set.

The number and percentage of subjects completing the second year extension period of the study from Week 52 up to and including Week 112 and the number and percentage of subjects withdrawing from the study during the second year extension period including reason for withdrawal, will be summarised by treatment group for subjects who continued in the second year extension period.

The number and percentage of subjects completing treatment with IP during the second year extension period from Week 48 up to and including Week 100 (i.e. with last administration of IP at Visit 26) and the number of subjects permanently discontinuing the treatment with IP during the second year extension period including reason for withdrawal, will be summarised by treatment group for the subjects who continued into the second year extension period.

The number and percentage of subjects not continuing into the second year extension period of the study including reason for not continuing, will be summarised by treatment group for the full analysis set.

The number and percentage of subjects completing the study up to and including Week 112 and the number and percentage of subjects withdrawing from the study including reason for withdrawal, will be summarised by treatment group for the full analysis set.

The number and percentage of subjects completing treatment with IP up to and including Week 100 (i.e. with last administration of IP at Visit 28) and the number of subjects

permanently discontinuing the treatment with IP including reason for withdrawal, will be summarised by treatment group for the full analysis set.

A summary of number and percentage of subjects in each country and each site by treatment group and overall will be provided for the full analysis set.

Summaries of important protocol deviations and analysis sets will be provided.

4.2.5.2 Demographic and baseline characteristics

Demography and baseline characteristics will be presented by descriptive statistics by treatment group as well as overall for the full analysis set.

Stratification factors at randomisation will be presented by descriptive statistics by treatment group as well as overall for the full analysis set.

Past and current medical history will be summarised separately by MedDRA system organ class and preferred term. Surgical history will be listed.

Prior and concomitant medications will be summarised separately by WHO Drug Dictionary ATC classification (ATC level 1), and preferred term. A summary of disallowed concomitant medication taken on or after the date of first dose of study treatment which would lead to the classification of a subject as a non-responder will be provided, splitting into medications received while on IP or while off IP, based on medication start date. For subjects having prematurely discontinued IP, medications received after IP initiation and before or on the date of last IP administration will be considered "on IP". However, for subjects who have a last infusion of IP within the Week 48 analysis visit window, medications received after IP initiation and before or on the date of Week 52 (or the date of Week 48 + 28 days in case that Week 52 was missed) will be considered "on IP". Medications received after the date of last IP administration, in case of premature IP discontinuation will be considered to be "off IP".

A summary of ACE inhibitors and ARBs taken during the study will be presented separately. A list of preferred terms for inclusion in this summary will be finalised and agreed prior to study database lock and unblinding.

4.2.5.3 Exposure

Exposure will be summarised by treatment group for the full analysis set.

Summary statistics will be provided for the cumulative duration of exposure. The number and percentage of subjects treated ≥4 weeks, ≥8 weeks, and up to 104 weeks in 4-weekly intervals will be provided.

The number and percentage of subjects with infusions will be presented by total number of infusion (i.e. 1, 2, …, 13) and in 4-weekly intervals (i.e. 4 weeks, 8 weeks, 12 weeks, …, 100 weeks).

Furthermore, the time to discontinuation of IP will be presented as Kaplan-Meier plot including the number of subjects at risk (i.e. still on IP).

Subjects with complications (e.g. an interruption) in any infusion will be presented in a listing.

4.2.6 Analysis methods for safety variables

Safety variables will be summarised by treatment group for the full analysis set.

4.2.6.1 Adverse Events

If not stated otherwise, all summaries described below will be presented separately for

- AEs during treatment
- AEs during follow-up

For summaries during follow-up, only subjects with any study documentation after the date of last dose of IP + 28 days will be considered.

An overall summary of subjects with at least one AE in the following categories will be presented, and will include the event rate per 100 subject years, and total exposure in years for each treatment group:

- Any AE
- Any acute AEs
  This category will not be included in the summary during follow-up.
- Any AE with outcome = death
- Any SAE (including events with outcome = death)
- Any AE leading to discontinuation of IP
  - This category will not be included in the summary during follow-up.
- Any AE related to IP by investigator's judgement
- Any AE of severe intensity
- Any AESI
  - Any AESI of non-opportunistic serious infections
  - Any AESI of opportunistic infections
  - Any AESI of anaphylaxis
  - Any AESI of malignancy
  - Any AESI of herpes zoster
  - Any AESI of tuberculosis
  - Any AESI of influenza
  - Any AESI of vasculitis (non-SLE)
  - Any AESI of major acute cardiovascular events (MACE as adjudicated)
- Any other significant AE

The number and percentage of subjects with at least one AE (i.e. multiple occurrences of an AE in 1 subject will only be counted once) will be summarised by MedDRA system organ class and preferred term (unless specified otherwise) for the following AE categories. These summaries will also include the event rate per 100 subject years and total exposure in years for each treatment group (unless specified otherwise).

- Any AE
- Any AE above reporting threshold of 2% of subjects in any treatment group in the full analysis set.

  This summary will be presented by preferred term for AEs during treatment only.
- Any non-serious AE above reporting threshold of 5% of subjects in any treatment group in the full analysis set

  This summary will not include the event rate per 100 subject years and total exposure in years.
- Any SAE with outcome = death

  This summary will be presented for SAEs with an outcome of death during the treatment and follow-up periods combined.
- Any SAE (including events with outcome = death)
- Any AE leading to discontinuation of IP
  This summary will be presented for AEs during treatment only.
- Any AE by relationship to IP (yes, no)
  (multiple occurrences of an AE in 1 subject will only be counted once as related if at least one AE is related and as not related if all occurrences are not related)
  This summary will be presented for AEs during treatment only.
- Any AE by maximum intensity (mild, moderate, severe)
  (i.e. multiple occurrences of an AE in 1 subject will only be counted once with the maximum intensity in this AE)
  This summary will be presented for AEs during treatment only.
- Any AESI
  This summary will not be presented by system organ class but by AESI category (non-opportunistic serious infections, opportunistic infections, anaphylaxis, malignancy, herpes zoster, tuberculosis, influenza and vasculitis [non- SLE]).
- Any other significant AE
  This summary will be presented for AEs during treatment only.
- Any AE by time interval (as detailed in Section 3.6.1) for the first onset of event
  This summary will be presented for AEs during treatment only.

Cardiovascular outcome events as determined by the Cardiovascular Event Adjudication Committee will be presented separately, summarising the number of AEs submitted for adjudication and the outcomes of the adjudication including MACE classification. Site reported cardiovascular AEs, cardiovascular SAEs and their corresponding adjudicated outcomes will be listed.

The time to first onset of herpes zoster during treatment will be presented as a Kaplan-Meier plot including the number of subjects at risk at each visit.

Furthermore, the alternative event rates per 100 subject years for herpes zoster (and possible other AESIs) will be summarised for events during screening, during treatment, and during follow-up. The total time at risk will also be presented.

An overall summary of subjects with AE's in pre-specified categories will also be presented by ADA category.

Key subject information for subjects with an AE with outcome of death, subjects with serious AEs, subjects with an AE leading to discontinuation of IP, subjects with cardiovascular events and subjects with AESIs, respectively, will be provided.

All AEs and AEs in ADA positive (at any time) subjects will be listed.

4.2.6.2 Laboratory variables

Observed values and changes from baseline of laboratory data for haematology, clinical chemistry and fasting lipid profile will be summarised by visit. The summary statistics presented will be minimum, 1st quartile, median, 3rd quartile, maximum, mean, and SD.

Shift plots (scatter plots) for each of the parameters presenting baseline values versus minimum post-baseline values and maximum post-baseline values, respectively, will be provided. A diagonal line indicating no change and horizontal and vertical reference lines indicating the limits of the reference ranges will also be displayed on the shift plots.

If any laboratory variables show any unusual features (high or low values or a general shift in the data points) at specific visits, then shift plots of these data may also be produced. This will be discussed and agreed upon during the BDR meeting.

For each laboratory parameter with available criteria, the number and percentage of subjects with TELVC values will be summarised by visit. Additionally, the number and percentage of subjects with at least one TELVC value will be presented.

The number and percentage of subjects with laboratory values below, within or above the corresponding normal range will be presented by visit. Urinalysis will be summarised as shift tables from baseline to the last on-treatment value for each parameter. Furthermore, the number and percentage of subjects with treatment-emergent changes will be summarised by parameter. In addition, observed values and change from baseline and pre-specified categories will be summarised by visit for any continuous variables.

In order to identify potential Hy's Law cases, the number and percentage of subjects in categories based on multiples of ULN for maximum on-treatment ALT and AST by maximum total bilirubin will be presented in a table. The table will include the following categories: < and ≥ 2xULN for total bilirubin and < 3xULN, ≥ 3 - < 5xULN, ≥ 5 - < 10xULN and ≥ 10xULN for ALT/AST.

Individual subject data will be presented f or all subjects who meet the criteria for potential Hy's law. The relevant laboratory parameters (ALT, AST, total bilirubin and ALP) will be tabulated showing all visits for these subjects. Subjects with elevated ALT or AST (ALT or AST ≥ 3xULN) and elevated total bilirubin (≥ 2xULN) at any time during the study (not necessarily at the same time) will be included.

4.2.6.3 ECGs

The number and percentage of subjects with normal, abnormal, not clinically significant and abnormal, clinically significant ECG results will be presented as a shift table from baseline to last observation on treatment.

Individual ECG parameters will not be summarised or analysed as the data is not collected.

4.2.6.4 SLEDAI-2K-based flare assessment instrument

The flare rate per of subject years for extra-renal flares will be presented overall, and by timing ("on-treatment" and "off-treatment"), for mild/moderate flares, severe flares and any flares, respectively. An "on-treatment" flare is defined as a flare with a date of onset after the date of first dose of IP and on or prior to the date of last dose of IP + 28 days. An "off-treatment" flare is defined as a flare with a date of onset after the date of last dose of IP + 28 days. Where a flare has been marked as both mild/moderate and severe, it will be considered to be severe for the purposes of this summary.

4.2.6.5 Vital signs

Observed values and changes from baseline of pulse, systolic blood pressure, diastolic blood pressure, respiration rate and body temperature, respectively, will be summarised by pre, during and post infusion, time, visit and treatment group. As multiple values are recorded during infusion and post-infusion, the average value during infusion and the average value post-infusion will also be calculated for each subject at each visit (along with the change from baseline for these average values) and these will be included in the summary tables along with the individual values.

For each parameter with available criteria, the number and percentage of subjects with TELVC values will be summarised by visit. Additionally, the number and percentage of subjects with at least one TELVC value will be presented.

For each parameter, the number and percentage of subjects with values below, within or above the corresponding normal range will be presented by visit. In addition, this information will be presented as shift tables from baseline to each post-baseline visit.

Observed values and changes from baseline of body weight and BMI will be summarised.

4.2.6.6 Cushingoid features

The number and percentage of subjects will be summarised for each feature by visit and treatment group. For subjects with baseline OCS ≥10 mg/day prednisone or equivalent, the summary will be repeated by OCS status at Week 52 (target OCS dose of ≤7.5 mg/day

reached versus target OCS dose of ≤7.5 mg/day not reached) and at Week 104 (target OCS dose of ≤5.0 mg/day reached versus target OCS dose of ≤5.0 mg/day not reached).

4.2.6.7 C-SSRS

The number and percentage of subjects with suicidal ideation (overall and by maximum category), suicidal behaviour (overall and by maximum category), and no suicidal ideation or behaviour will be presented by treatment group for assessments during screening, during treatment, and during follow-up, respectively.

Furthermore, descriptive statistics on the total number of attempts, total number of interrupted attempts and total number of aborted attempts will be summarised by treatment group for attempts during screening, during treatment and during follow-up, respectively.

Subjects with a suicidal ideation or behaviour at any time up until Week 52 will be presented in a listing.

4.2.6.8 Personal Health Questionnaire Depression Scale-8

Observed values and changes from baseline in PHQ-8 total score will be presented with descriptive statistics by visit and treatment group.

4.2.7 Analysis methods for pharmacokinetics, pharmacodynamics, and immunogenicity

If not stated otherwise, all analyses of pharmacodynamics and immunogenicity will be performed for the full analysis set by treatment group. If not stated otherwise, all analyses of pharmacokinetics will be performed for the PK analysis set by treatment group.

4.2.7.1 Analysis methods for immunogenicity variables

The proportion of subjects with ADA results as described in Section 3.7.1 will be presented together with descriptive statistics of the available titres. Descriptive statistics will include n (number of reportable titres), minimum, quartiles (1st quartile, median and 3rd quartile) and maximum.

ADA response and the corresponding titre will be summarised at baseline and at all scheduled post-baseline visits in the study period by treatment group for the full analysis set. In addition, the proportion of ADA positive subjects will be presented cumulatively.

A summary of key subject information will be provided for ADA positive subjects.

The analyses of impact of ADA on PK and PD endpoints are described in Sections 4.2.7.2 and 4.2.7.3.

Pre-specified AE categories and hypersensitivity events during the on-treatment period will be summarised by ADA status (positive, negative).

Additional summaries of immunogenicity data may be generated based on the results of the above analyses.

4.2.7.2 Analysis methods for pharmacokinetic variables

Anifrolumab serum concentrations will be summarised using descriptive statistics by visit reporting n (number of non-missing values), m (number of non-zero values), geometric mean, geometric CV%, arithmetic mean, SD, arithmetic CV%, median, minimum and maximum. If applicable, this summary will be repeated including individual concentrations excluded from descriptive statistics.

Serum concentration-time profiles of anifrolumab by treatment group will be generated as plots of mean values (including SD) by time point in a semi-log scale and a linear scale. This will be presented for all values and additionally for values measured before administration (C_{trough}).

The potential influence of demographic covariates such as body weight, race, sex and age will be explored. Furthermore, potential correlation between anifrolumab concentrations and 21-gene type IFN PD signature will be explored. These analyses will be performed by MedImmune's Clinical Pharmacology and Drug Metabolism and Pharmacokinetics (DMPK) group.

Impact of ADA on PK will be explored by spaghetti plots of individual concentration data over time using different styles by ADA result (positive at any time/negative). This will be presented for all values and additionally for values measured before administration (C_{trough}). Summaries of anifrolumab serum concentrations by visit and ADA status (positive, negative) will be presented.

Mycophenolic acid serum concentrations will be summarised using descriptive statistics as specified above at each visit by treatment group. Serum concentration-time profiles of MPA by treatment group will also be generated. Analyses of mycophenolic acid serum concentrations will be performed on the full analysis set.

4.2.7.3 Analysis methods for pharmacodynamic and lupus serology variables

IFN21 gene PD signature as percent of baseline IFN21 PD signature will be summarised by time point for subjects who are IFN test-high at randomisation. The median absolute deviation (MAD) will also be included in this summary. Furthermore, the median percent of baseline IFN21 PD signature (including MAD) over time will be presented as a line plot for the same population.

The impact of ADA on IFN21 gene PD signature will be explored. For IFN test-high subjects, IFN21 gene PD signatures as percent of baseline IFN21 PD signature will be summarised descriptively by ADA status (ADA negative/positive at any time) and ADA positive subgroups (ADA persistently/prolonged persistently positive, ADA titre > median, ADA nAb positive) by timepoint and treatment group.

For subjects who are positive for anti-dsDNA at baseline and abnormal complement level at baseline, respectively, observed values and changes from baseline of anti-dsDNA, C3 and C4 complement levels will be summarised by visit, providing the frequency of positive data allows.

4.3 Subgroup analysis

To explore the uniformity of the detected overall treatment effect, subgroup analyses will be performed for the following factors:

- Type I IFN 4-gene test at screening (IFN test-high, IFN test-low)
- 24-hour UPCR at screening ≤3 mg/mg, >3 mg/mg
- Age (≥18 to 64, ≥65 years)
- Placebo response region (low placebo response rate versus high placebo response rate) Countries with a low placebo response rate versus high placebo response rate are defined as [Low: Australia, Belgium, France, Germany, Italy, South Africa, Spain, United Kingdom, United States], [High: Argentina, Hungary, Republic of Korea, Mexico, Peru, Poland, Russian Federation, Serbia, Taiwan]
- Race (white; black or African American; Asian, native Hawaiian or other Pacific Islander; American Indian or Alaska native; other)
- Screening biopsy classification (class III, class IV, mixed [class III or IV + class V])
- Number of methylprednisolone pulses (1, >1). Pulses are counted during the whole study.
- ADA result (Positive at any time, Negative)
- OCS at baseline (<20 mg/day, ≥20 mg/day)
- eGFR at baseline (<60 mL/min/1.73m^2, ≥60 mL/min/1.73m^2)

Table 3, Table 4 and Table 5 give an overview of subgroup analyses to be performed at Week 24, 52 and Week 104, respectively. Where necessary (e.g. for subgroup analysis based on stratification factors), the model factors will be reduced. If not stated otherwise, the subgroup analysis will not be performed for a sub-population if any treatment group consists of less than 10 subjects within that sub-population. Where a sub-population contains less than 10 subjects in any treatment group, only ns and treatment estimates will be presented (no treatment comparisons will be presented).

Forest plots will be used to summarise the estimates of the treatment effect for the applicable subgroups.

Table 3 Overview of subgroup analyses at Week 24

Analysis	IFN test	24-hour UPCR	OCS at baseline
24-hour UPCR	X	X	X
CRR	X	X	X

Table 4 Overview of subgroup analyses at Week 52

Analysis	IFN test	24-hour UPCR	Age	Placebo response	Race	Number of methylprednisolone pulses	Screening biopsy	ADA result	OCS at baseline	eGFR at baseline
24-hour UPCR	X	X	X	X	X	X	X	X	X	X
CRR	X	X	X	X	X	X	X		X	X
aCRR	X	X							X	
PRR	X	X				X	X			
Graded CRR	X	X				X	X			
OCS tapering	X	X		X		X	X			

Table 5 Overview of subgroup analyses at Week 104

Analysis	IFN test	24-hour UPCR	OCS at baseline	eGFR at baseline
24-hour UPCR	X	X	X	X
CRR	X	X	X	X

Analysis	IFN test	24-hour UPCR	OCS at baseline	eGFR at baseline
PRR	X	X		
Graded CRR	X	X		
OCS tapering	X	X		

5 INTERIM ANALYSES

One interim analysis either for efficacy or futility may be conducted after approximately 80% of subjects have completed the Week 52 visit or would have completed Week 52 if they had not discontinued the study. Primary efficacy and safety analyses will be performed once all data from the main study (up to and including Week 52) is available, cleaned and soft-locked. The study will also be unblinded at this stage for the team performing the analysis.

5.1 Potential Interim analysis for futility

The objective of a potential interim analysis for futility is to give consideration for stopping the study early, should the chance of detecting a statistically significant difference with placebo for the primary endpoint at the final analysis be sufficiently small given the interim data. Non-binding futility rule for stopping the study will be used if a futility analysis is performed. Even though non-binding futility boundary will be used for stopping the study, if interim analysis for futility is performed, a portion of the 2-sided alpha will still be spent. A Peto-Haybittle spending function would be used with a fixed 2-sided p-value threshold of 0.1% at the interim. The p-value threshold at the final analysis will be adjusted to account for the interim analysis, based on the actual observed information fraction for the primary endpoint. Under the assumption that there are 120 subjects included at the interim and 150 subjects at the final analysis, the critical threshold at the final analysis will be 4.99% (East Version 6.4).

Considerations for stopping the study may be given if based on the observed effect at the interim analysis, the probability of detecting a statistically significant difference with placebo for the primary endpoint at the final analysis as measured by the predictive power (Herson, 1979), is no higher than 20%. The predictive power will be derived from the following expression:

$$Predictive\ Power = \Phi\left((1-\Pi)^{-\frac{1}{2}}\left(z_{0.025}\cdot\sqrt{\Pi}-\widehat{\theta}_{fut}\sqrt{I_{fut}}\right)\right),$$

where

- Φ denotes the cumulative distribution function of a standard normal distribution;
- $z_{0.025}$ is the Z-score at the lower tail corresponding to a 2-sided alpha level 0.0499 assuming 120 patients in the interim analysis (lower tail is used instead of upper due to using a logarithmic scale);
- $\hat{\theta}_{fut}$ is the observed at the interim analysis log-transformed relative difference in the change from baseline to Week 52 in 24-hour UPCR between the pooled anifrolumab treatment group versus placebo;
- I_{fut} is the information at the futility analysis $I_{fut} = \{Var(\hat{\theta}_{fut})\}^{-1}$, which can be estimated as $1/SE^2_{MMRM}$, where SE^2_{MMRM} is the standard error of the relative difference in change from baseline in 24-hour UPCR on a logarithmic scale from the MMRM model for the primary endpoint at the interim analysis;
- Π is the fraction of information $\Pi = \frac{I_{fut}}{I_{fin}} = \frac{number\ of\ patients\ in\ the\ interim\ analysis}{expected\ number\ of\ patients\ in\ the\ final\ analysis}$, if $I_{fin} = \{Var(\hat{\theta}_{fin})\}^{-1}$ is the information at the final analysis.

The above derivation is based on the weighted average of the conditional power over values of the difference with placebo θ, drawn from the posterior distribution below, which is obtained using an uninformative prior (Jennison & Turnbull, 2000):

$$\theta | Z_{fut} \sim N\left(\hat{\theta}_{fut}, \frac{1}{I_{fut}}\right),$$

The conditional power is the conditional probability that the final test statistic will exceed a critical value given the data observed thus far and an assumption about the pattern of the data to be observed in the remainder of the study, expressed as:

$$Conditional\ Power = 1 - \Phi\left((1-\Pi)^{-\frac{1}{2}}\left(\hat{\theta}_{fut}\sqrt{I_{fut}} \cdot \sqrt{\Pi} - z_{0.0245}\right) + \theta\sqrt{\frac{I_{fut}}{\Pi}}\sqrt{1-\Pi}\right).$$

Based on the above, the corresponding treatment effect boundary providing 20% predictive power can be calculated as:

$$\delta_{fut} = exp\left\{\left(z_{0.025} + z_{1-20\%}\sqrt{\frac{1}{\Pi} - 1}\right)\sqrt{\frac{\Pi}{I_{fut}}}\right\}$$

An observed effect $\hat{\theta}_{fut}$ worse than δ_{fut} corresponds to a predictive power less than 20%.

5.2 Potential Interim analysis for efficacy

CCI [redacted] Even though there is no provision to stop the study early due to an efficacy claim, if an interim analysis for efficacy is performed, a portion of the 2-sided alpha

will be spent. A Peto-Haybittle spending function would therefore be used with a fixed 2-sided p-value threshold of 0.1% at the interim. The p-value threshold at the final analysis will be adjusted to account for the interim analysis, based on the actual observed information fraction for the primary endpoint. Under the assumption that there are 120 subjects included at the interim and 150 subjects at the final analysis, the critical threshold at the final analysis will be 4.99% (East Version 6.4).

6 CHANGES OF ANALYSIS FROM PROTOCOL

A possible interim analysis for futility may be performed, following results from a pivotal Phase 3 SLE study for anifrolumab in subjects with SLE. In the case that there is sufficient evidence to suggest that anifrolumab in addition to SOC is not beneficial over placebo in addition to SOC for subjects with SLE, consideration can be giving for stopping this study early for ethical reasons. The presence of such evidence can be evaluated only after obtaining results from the second anifrolumab pivotal study in SLE.

The threshold for the analysis of sustained reduction in OCS has been updated from 10 to 20.

The OCS at baseline subgroup categories are changed from (≤20 mg/day, >20 mg/day) in the protocol to (<20 mg/day, ≥20 mg/day) in the SAP for consistency with the sustained reduction in OCS endpoint definition.

7 REFERENCES

Austin H, M. L. (1983). Prognostic factors in Lupus Nephritis. *American Journal of Medicine, 75*, 382-391.

Furie R, N. K.-V. (2014). Efficacy and safety of abatacept in lupus nephritis: a twelve-month, randomized, double-blind study. *Arthritis Rhematology, 66*(2), 379-389.

Herson, J. (1979). Predictive probability early termination plans for phase II clinical trials. *Biometrics, 35*, 775-783.

Jennison, C., & Turnbull, B. (2000). *Group sequential methods with applications in clinical trials.* Chapman & Hall.

Levey AS, C. J. (2006). Using Standardized Serum Creatinine Values in the Modification of Diet in Renal DIsease Study Equation for Estimating Glomerular Filtration Rate. *Annals of Internal Medicine, 145*(4), 247-254.

Stokes ME, D. C. (2012). *Categorical Data Analysis using SAS* (Third ed.). USA: SAS Institute.

Weening J, D. V. (2004). The classification of glomerulonephritis in systemic lupus erythematosus revisited. *Kidney International, 65*, 521-530.

SIGNATURE PAGE

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature

Document Name: d3461c00007-sap-ed-5		
Document Title:	D3461C00007 Statistical Analysis Plan Edition 5	
Document ID:	CCI [redacted]	
Version Label:	2.0 CURRENT LATEST APPROVED	
Server Date **(dd-MMM-yyyy HH:mm 'UTC'Z)**	**Signed by**	**Meaning of Signature**
25-Mar-2020 13:35 UTC	PPD	Content Approval
25-Mar-2020 13:37 UTC	PPD	Author Approval

Notes: (1) Document details as stored in ANGEL, an AstraZeneca document management system.