Statistical Analysis Plan

Clinical Trial Protocol Identification No.	MS200527-0081
Title:	A Phase IIa Randomized, Double-Blind, Placebo-Controlled Trial to Evaluate the Efficacy and Safety of M2951 in Subjects with Rheumatoid Arthritis on Stable Methotrexate Therapy
Trial Phase	IIa
Investigational Medicinal Product(s)	M2951
Clinical Trial Protocol Version	02 September 2016 / Version 3.0
Statistical Analysis Plan Authors	PPD, Biostatistician PPD, Pharmacokineticist
Statistical Analysis Plan Date and Version	18 April 2017 / Version 1.0
Statistical Analysis Plan Reviewers	Merck Reviewers: PPD, Lead Program Statistician PPD, Lead Statistical Programmer PPD, Statistical Programmer PPD, Global Drug Safety PPD, Medical Writer Lead PPD, Medical Responsible PPD, Lead Data Management PPD, Data Management PPD, Clinical biomarkers PPD, Clinical Pharmacokineticist PPD, PPD PPD, Clinical Trial Lead

PPD Reviewers:

PPD, Statistical Scientist
PPD, Lead Study Statistician
PPD, Global Medical Advisor
PPD, Medical Writer
PPD, Lead Statistical Programmer
PPD, Statistical Scientist
PPD, Pharmacokineticist
PPD, Pharmacokineticist

1 Signature Page

Statistical Analysis Plan: MS200527-0081

A Phase IIa Randomized, Double-Blind, Placebo-Controlled Trial to Evaluate the Efficacy and Safety of M2951 in Subjects with Rheumatoid Arthritis on Stable Methotrexate Therapy

PPD

PPD

PPD

Trial Biostatistician

PPD

PPD

PPD

PPD

Lead Program Statistician
EMD Serono, Inc.
45A Middlesex Turnpike | Billerica MA 01821 | USA

2 Table of Contents

1 Signature Page 3
2 Table of Contents....4
3 List of Abbreviations and Definition of Terms 6
4 Modification History 9
5 Purpose of the Statistical Analysis Plan 9
6 Summary of Clinical Trial Features 10
7 Sample Size/Randomization....17
8 Overview of Planned Analyses....18
8.1 Independent Monitoring Committee review....18
8.2 Interim Analysis....18
8.3 Primary Analysis 18
8.4 Final Analysis 18
9 Changes to the Planned Analyses in the Clinical Trial Protocol 19
10 Protocol Deviations and Analysis Sets 20
10.1 Definition of Protocol Deviations and Analysis Sets 20
10.2 Definition of Analysis Sets and Subgroups 21
11 General Specifications for Statistical Analyses 24
12 Trial Subjects 31
12.1 Disposition of Subjects and Discontinuations 31
12.2 Protocol Deviations 33
12.2.1 Important Protocol Deviations....34
12.2.2 Clinically Important Protocol Deviations 34
13 Demographics and Other Baseline Characteristics....35
13.1 Demographics 35
13.2 Medical History 36
13.3 Other Baseline Characteristics....36
13.3.1 Disease history....36
13.3.2 Other 37
14 Previous or Concomitant Medications/Procedures....38
15 Treatment Compliance and Exposure....41

16 Endpoint Evaluation 46
16.1 Primary Endpoint Analyses 49
16.2 Secondary Endpoint Analyses 51
16.2.1 Key secondary endpoint analysis....51
16.2.2 Other secondary endpoint analyses 51
16.3 Other Endpoint Analyses 54
16.3.1 Open-label Extension Endpoint Analyses 54
16.3.2 Pharmacokinetic Endpoint Analyses 55
16.3.3 Pharmacodynamic Endpoint Analyses 56
16.3.4 Exploratory Endpoints 56
17 Safety Evaluation....57
17.1 Adverse Events 58
17.1.1 All Adverse Events 59
17.2 Deaths, Other Serious Adverse Events, and Other Significant Adverse Events 61
17.2.1 Deaths 61
17.2.2 Serious Adverse Events 62
17.3 Clinical Laboratory Evaluation....62
17.4 Vital Signs 64
17.5 12-Lead Electrocardiogram (ECG)....64
17.6 Physical Examination 65
17.7 Pregnancy test 66
17.8 B cell (CD19+) count....66
17.9 Urine Protein/Creatinine Ratio 66
17.10 Urinalysis Microscopic Evaluation....66
17.11 Chest X-ray evaluations....66
17.12 Other safety evaluations 66
18 References....67
19 Appendices 68
19.1 Appendix 1: HAQ-DI scoring 68
19.2 Appendix 2: DAS28-hsCRP formula 70
19.3 Appendix 3: Important Protocol Deviations identified by programming and Clinically Important Protocol Deviations 71

3 List of Abbreviations and Definition of Terms

Δ	Difference
ACR	American College of Rheumatology
ADaM	Analysis Data Model
AE	adverse event
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
anti-CCP	anti-cyclic citrullinated peptide
AST	aspartate aminotransferase
ATC	anatomical therapeutic class
AUC_{0-6h}	area under the concentration-time curve from time zero to 6 hours
bid	twice daily
BLQ	below limit of quantification
BMI	body mass index
BOA	Biostatistics Outputs Assembly
BTK	Bruton's tyrosine kinase
CFB	change from baseline
CI	confidence interval
C_{max}	maximum observed plasma concentration
C_{trough}	plasma concentration observed immediately before next dosing
CSR	Clinical Study Report
CTP	Clinical Trial Protocol
CV	coefficient of variation
CXR	chest X-ray
DAS28-hsCRP	disease activity score based on 28 joints and high-sensitivity C-reactive protein
DBP	diastolic blood pressure
DMARD	disease-modifying antirheumatic drug
eCRF	Case Report Form, electronic
ECG	electrocardiogram

ESR	erythrocyte sedimentation rate
EULAR	The European League Against Rheumatism
FSH	follicle-stimulating hormone
GGT	gamma-glutamyl transferase
GeoCV	geometric coefficient of variation
GeoMean	geometric mean
HAQ-DI	Health Assessment Questionnaire - Disability Index
HBsAg	hepatitis B surface antigen
HIV	human immunodeficiency virus
hsCRP	high-sensitivity C-reactive protein
IA	interim analysis
IDMC	Independent Data Monitoring Committee
Ig	immunoglobulin
IMP	investigational medical product
IUD	intrauterine device
IUS	intrauterine hormone-releasing system
IWRS	Interactive Web Response System
LOCF	last observation carried forward
LPLD	last patient's last dose
LTBI	latent TB infection
MCMC	Markov chain Monte Carlo
MedDRA	medical dictionary for regulatory activities
mITT	modified intent-to-treat
MTX	methotrexate
N	number of subjects
NCI-CTCAE	National Cancer Institute - Common Terminology Criteria for Adverse Events
NSAID	nonsteroidal anti-inflammatory drug
OLE	open-label extension
PD	pharmacodynamics
PK	pharmacokinetics

PT	preferred term
Q1	25th percentile
Q3	75th percentile
RA	rheumatoid arthritis
$R_{acc(AUC0-6h)}$	accumulation ratio for AUC_{0-6h}
$R_{acc(Cmax)}$	accumulation ratio for C_{max}
RF	rheumatoid factor
SAF	safety analysis set
SAP	Statistical Analysis Plan
SBP	systolic blood pressure
SCR	screening analysis set
SD	standard deviation
SDTM	Study Data Tabulation Model
SEM	standard error of the mean
SJC	swollen joint count
SOC	system organ class
TB	Tuberculosis
TEAE	treatment-emergent adverse event
TJC	tender joint count
t_{max}	time to reach maximum plasma concentration
ULN	upper limit of normal
VAS	visual analog scale
WHO-DD	World Health Organization drug dictionary

4 Modification History

Unique Identifier for SAP Version	Date of SAP Version	Authors	Changes from the Previous Version
Final 1.0	18 April 2017	PPD, PPD	NA – first version

5 Purpose of the Statistical Analysis Plan

The purpose of this Statistical Analysis Plan (SAP) is to document technical and detailed specifications for analyses of data collected for protocol MS200527-0081 dated 02 September 2016 (version 3.0). The SAP is based upon Section 8 (Statistics) of the trial protocol and is prepared in compliance with International Council for Harmonisation E9.

The first version (version 1.0) of the SAP includes details for the primary and final statistical analyses, planned for the end of 12 weeks treatment and for the end of study, respectively.

Results of the analyses described in this SAP will be included in the Clinical Study Report (CSR). Additionally, the planned analyses identified in this SAP will be included in regulatory submissions or future manuscripts. Any post-hoc, or unplanned analyses performed to provide results for inclusion in the CSR but not identified in this prospective SAP will be clearly identified in the CSR.

The final clinical database cannot be locked until the final SAP has been approved and signed.

Another SAP document (MS200527-0081_IA_IDMC_SAP_v3.0 08 Feb2017) detailed specifications for the Independent Data Monitoring Committee (IDMC) and Interim analyses (IA).

6 Summary of Clinical Trial Features

Clinical Trial Protocol Number	MS200527-0081
Title	A Phase IIa Randomized, Double-Blind, Placebo-Controlled Trial to Evaluate the Efficacy and Safety of M2951 in Subjects with Rheumatoid Arthritis on Stable Methotrexate Therapy
Trial Phase	IIa
IND Number	122,963
FDA covered trial	No
EudraCT Number	2016-000064-42
Coordinating Investigator	PPD PPD Phone PPD Fax PPD PPD
Sponsor	Merck KGaA Darmstadt, Germany Medical Responsible: PPD 45A Middlesex Turnpike Billerica, MA 01821, USA Telephone: PPD
Trial centers/countries	Approximately 50 sites in 10 countries
Planned trial period (first subject in-last subject out)	July 2016 to March 2017
Trial Registry	Clinicaltrials.gov

Objectives:

Primary objective: To assess the efficacy of M2951 in subjects with active rheumatoid arthritis (RA) on stable methotrexate (MTX) therapy, as measured by the American College of Rheumatology (ACR) 20% (ACR20) response rate over a duration of 84 days.

Key secondary objective: To assess the level of disease activity at 4 weeks in subjects with active RA as assessed by the percent change in high-sensitivity C-reactive protein (hsCRP) from Baseline to Day 29.

Other secondary objectives: To assess the safety and tolerability of M2951 and to evaluate the pharmacokinetics (PK) and pharmacodynamics (PD) of M2951 in subjects with RA on

stable MTX therapy.

Exploratory objectives: CCI [redacted]

Open-label extension objectives:

To evaluate the on-going safety and efficacy of M2951 in an open-label extension with treatment at 50 mg bid for an additional 6 months.

Methodology: This is a Phase IIa randomized, double-blind, placebo-controlled trial in approximately 64 modified intent-to-treat (mITT) subjects 18 to 75 years of age with RA on stable MTX therapy.

Total duration of subject participation is approximately between 5 months (20 weeks) to 11 months (46 weeks), which includes:

- Screening: 28 days (4 weeks)
- Treatment: 84 days (12 weeks)
- Optional 6-month open-label extension period (26 weeks)
- Safety Follow-Up: 28 days (4 weeks)

Screening will occur over Day -28 to -1 (Visit 1), during which time all Screening assessments must be completed and reviewed.

On Day 1 (Visit 2), after confirmation of eligibility, assessment of RA disease activity and vital signs, and obtaining of required blood and urine samples, subjects will be randomly assigned 1:1 to receive M2951 or placebo in combinations with MTX.

Blood and urine samples will be obtained for safety, markers of disease activity, PK, and PD following the schedule detailed in the Schedule of Assessments.

CCI [redacted]

Subjects will self-administer the Investigational Medical Product (IMP) at a set time each day (± 2 hours). Subjects must take their daily dose more than 1 hour prior to a meal or snack or more than 2 hours after a meal or snack. Clear fluids are allowed at any time. On trial visit days, the IMP will be administered during the trial visit, after trial visit procedures (other than PK/PD sampling) are completed.

Subjects will have trial visits weekly for the first 3 weeks of treatment; bi-weekly up until 8 weeks of treatment; a final On-Treatment Visit on Day 85, and then return for 1 Follow-Up Visit approximately 28 days (4 weeks) after the last dose of IMP.

After completing the 12-week treatment period, subjects will be offered the opportunity to participate in an optional 6-month open-label extension period where all subjects will receive M2951. For subjects who choose to participate in the optional extension period, the duration of participation will last a total of approximately 11 months.

Subjects who continue to receive M2951 during the extension period will also be required to undergo the Safety Follow-up Visit scheduled for approximately 28 days after the last treatment visit.

Planned number of subjects: A total of 64 subjects with RA.

Primary endpoint: The primary endpoint is the ACR20 response at Day 85.

Key secondary endpoints:

The key secondary endpoint is the percent change in hsCRP from Baseline to Day 29.

Other secondary endpoints (during the double-blind treatment period):

Safety endpoints (from signing ICF until Safety Follow-Up Visit):

- Nature, incidence, and severity of treatment-emergent adverse events
- Changes from Baseline in vital signs and 12-lead electrocardiogram (ECG) results
- Changes from Baseline in standard hematology and chemistry laboratory parameters including liver function tests, amylase/lipase, and urinalysis.

Efficacy endpoints:

- ACR50/ACR70 response at each time point assessed
- Percent change in hsCRP from Baseline to Day 85
- Change from Baseline in Disease Activity Score based on 28 joints (DAS28-hsCRP) score at Day 29
- Change from Baseline in DAS28-hsCRP score at Day 85
- Percentage of subjects with DAS28-hsCRP < 3.2 (well-controlled disease) at Day 85
- Percentage of subjects with DAS28-hsCRP < 2.6 at Day 85
- Change from Baseline in erythrocyte sedimentation rate (ESR), anti- cyclic citrullinated peptide (CCP), and rheumatoid factor ([RF] at Days 29 and 85)
- Change from Baseline to Day 85 in subject's self-assessments including:
  - Global assessment of disease activity
  - Self-assessment of pain
  - Self-assessment of disability (Health Assessment Questionnaire – Disability Index; HAQ-DI)
- Change from Baseline to Day 85 in Physician's Global Assessment of Disease Activity.

Pharmacokinetics: Pharmacokinetic parameters calculated for Days 1 and 29 include: plasma concentrations of M2951, area under the plasma concentration-time curve from time zero to 6 hours ($AUC_{0\text{-}6h}$) after administration, maximum observed plasma concentration (C_{max}), concentration observed immediately before next dosing (C_{pre}) (Day 29), time to reach maximum plasma concentration (t_{max}), accumulation ratio for $AUC_{0\text{-}6h}$ ($R_{acc[AUC0\text{-}6h]}$), and

accumulation ratio for C_{max} ($R_{acc[Cmax]}$).

Other assessments:

Pharmacodynamics:

- Absolute concentrations and change from Baseline in immunoglobulin (Ig) levels (IgA, IgM, and IgG, and subclasses) on Days 29 and 85
- Absolute numbers and change from Baseline in B cell numbers and subsets on Day 85.

Exploratory:

[redacted]

Endpoints for open-label extension:

Safety:

- Nature, incidence and severity of TEAEs
- Change from Baseline in vital signs and 12-lead ECG data
- Change from Baseline in standard hematology and chemistry laboratory parameters including liver function tests, amylase/lipase, and urinalysis.

Efficacy:

- ACR50/ACR70 response at Month 6
- Percent change in hsCRP from Baseline to Month 6
- Change from Baseline in DAS28-hsCRP score at Month 6
- Percentage of subjects with DAS28-hsCRP < 3.2 (well-controlled disease) at Month 6
- Percentage of subjects with DAS28-hsCRP < 2.6 at Month 6
- Change from Baseline in ESR, anti-CCP, rheumatoid factor (at Month 6)
- Change from Baseline to Month 6 in subject's self-assessments including:
  - Global assessment of disease activity (VAS)
  - Self-assessment of pain (VAS)
  - Self-assessment of disability (HAQ-DI)

Change from Baseline to Month 6 in Physician's Global Assessment of Disease Activity (VAS).

Diagnosis and key inclusion and exclusion criteria:

The current trial will enroll subjects who fulfill the following key inclusion criteria:

- Men or women 18 to 75 years of age at the time of informed consent signature
- Confirmed diagnosis of RA according to 2010 American College of Rheumatology (ACR)/ The European League Against Rheumatism (EULAR) RA classification criteria of at least 6 months duration
- Positive RF and/or anti-CCP (anti-cyclic citrullinated peptide)
- Persistently active disease defined as ≥ 6 swollen joints (of 66 counted) and ≥ 6 tender joints (of 68 counted)
- hsCRP ≥ 3.6 mg/L
- Treatment for ≥ 12 weeks with 10 to 25 mg/week MTX at a stable dose for at least 4 weeks prior to dosing with the IMP and maintained throughout the trial
- Women of childbearing potential must use a highly effective method of contraception combined with one supplementary barrier method for 4 weeks prior to randomization, throughout the trial, and for 90 days after the last dose of IMP. For the purposes of this trial:
    - Females who are postmenopausal (age-related amenorrhea ≥ 12 consecutive months and increased follicle-stimulating hormone [FSH] > 40 mIU/mL), or who have undergone hysterectomy or bilateral oophorectomy are exempt from pregnancy testing. If necessary to confirm postmenopausal status, an FSH will be drawn at Screening
    - The definition of highly effective contraception includes:
        - combined (estrogen and progestogen containing) hormonal contraception associated with inhibition of ovulation; oral, intravaginal or transdermal
        - progestogen-only hormonal contraception associated with inhibition of ovulation; oral, injectable, implantable
        - intrauterine device (IUD)
        - intrauterine hormone-releasing system (IUS)
        - bilateral tubal occlusion
        - vasectomized partner
        - sexual abstinence
    - Supplementary barrier methods include:
        - male or female condom with or without spermicide
        - cap, diaphragm or sponge with spermicide

- Women of childbearing potential must have a negative serum pregnancy test at the Screening Visit and a negative urine pregnancy test at Day 1/randomization before dosing.

Subjects who fulfill any of the following key exclusion criteria should not be enrolled into this trial:

- Use of oral corticosteroids > 10 mg daily prednisone equivalent, use of injectable corticosteroids, or change in dose of corticosteroids within 2 weeks prior to Screening or during Screening
- Initiation or change in dose for nonsteroidal anti-inflammatory drugs (NSAIDs) within 2 weeks prior to Screening
- Treatment with tofacitinib, other BTK inhibitors, or a biologic disease-modifying antirheumatic drug (DMARD; eg, anti-tumor necrosis factor alpha [anti-TNF-α], tocilizumab [anti-interleukin-6 receptor], abatacept [CTLA4-Fc]), or other immunosuppressive drugs (sulfasalazine would be acceptable at a stable dose) other than methotrexate within 3 months prior to Screening or during Screening
- Treatment with anti-CD20 therapy (eg, rituximab) within 12 months prior to Screening or during Screening
- Immunologic disorder other than RA, with the exception of secondary Sjogren's syndrome associated with RA, and well-controlled diabetes or thyroid disorder, or any other condition requiring oral, intravenous, intramuscular, or intra-articular corticosteroid therapy
- Vaccination with live or live-attenuated virus vaccine within 1 month prior to Screening
- Active, clinically significant, viral, bacterial, or fungal infection, or any major episode of infection requiring hospitalization or treatment with parenteral anti-infectives within 4 weeks of Screening or during Screening, or completion of oral anti-infectives within 2 weeks before or during Screening, or a history of recurrent infections (ie, 3 or more of the same type of infection in a 12-month rolling period). Vaginal candidiasis, onychomycosis, and genital or oral herpes simplex virus considered by the Investigator to be sufficiently controlled would not be exclusionary.
- History of or positive testing for human immunodeficiency virus (HIV), hepatitis C antibody and/or polymerase chain reaction, hepatitis B surface antigen (HBsAg) (+) and/or hepatitis B core total, and/or IgM antibody (+) at Screening.
- History of or current diagnosis of active tuberculosis (TB); undergoing treatment for latent TB infection (LTBI); untreated LTBI (as determined by documented results within 3 months of the Screening Visit of a positive TB skin test with purified protein derivative with induration ≥ 5 mm, a positive QuantiFERON®-TB test or positive or borderline T-SPOT [Elispot] test); or positive QuantiFERON-TB test at Screening. Subjects with documented completed appropriate LTBI treatment would not be excluded and are not required to be tested.
- Subjects with current household contacts with active TB will also be excluded.

- Indeterminate QuantiFERON-TB or T-SPOT tests may be repeated once, and will be considered positive if retest results are positive or indeterminate.
- History of cancer, except adequately treated basal cell or squamous cell carcinomas of the skin (no more than 3 lesions requiring treatment in lifetime) or carcinoma in situ/cervical intraepithelial neoplasia of the uterine cervix, unless considered cured > 5 years.
- Clinically significant abnormality on ECG, or an active infective process or any other clinically significant abnormality on Screening chest X-ray (CXR) taken within 4 weeks of the first dose, per Investigator opinion. If a CXR has been taken within the previous 3 months and results are available and normal, the CXR does not need to be carried out
- B cell (CD19) count < 50% of the lower limit of normal at Screening
- Significant cytopenia including absolute neutrophil count < 1,500/ mm^3, platelet count < 100,000/mm^3, or absolute lymphocyte count < 1,000/mm^3

Investigational Medicinal Product: dose/mode of administration/dosing schedule:
M2951 (50 mg twice daily [bid]) will be administered as 2 x 25 mg capsules bid during the 12-week treatment period. Identical capsules filled with mannitol for placebo will be provided. M2951 will be administered as 2 x 25 mg tablets bid during the optional extension period.

Reference therapy: dose/mode of administration/dosing schedule: Not applicable.

Planned trial and treatment duration per subject: Total duration of subject participation is approximately between 5 months (20 weeks) to 11 months (46 weeks), which includes:

- Screening: 28 days (4 weeks)
- Treatment: 84 days (12 weeks)
- Optional 6-month open-label extension period (26 weeks)
- Safety Follow-Up: 28 days (4 weeks)

After completing the 12-week treatment period, subjects will be offered the opportunity to participate in an optional 6-month open-label extension period with M2951. For subjects who choose to participate in the optional extension period, the duration of participation will last a total of approximately 11 months.

Statistical methods:

The sample size of 60 evaluable subjects was chosen to provide 80% power to demonstrate superiority of M2951+MTX to placebo + MTX for the primary endpoint at the 1-sided 10% level. Eligible subjects will be randomized 1:1 to treatment with M2951 or placebo through a central randomization process by an Interactive Web Response System (IWRS). Approximately 64 subjects will be randomized under the assumption of a 6.25% drop-out rate.

There will be 3 analyses: (1) an interim analysis (IA), triggered when subjects from the first 50% of the planned enrollment have reached Day 29 of treatment and have available Day 29 hsCRP data, or have discontinued treatment prior to Day 29, (2) a primary analysis, triggered

when 100% of subjects enrolled have reached Day 85 of treatment or have discontinued treatment prior to Day 85, and (3) a final analysis, triggered when 100% of subjects enrolled complete all study parts, or discontinue prematurely from study.

The percent change in hsCRP from Baseline to Day 29 will be the key endpoint evaluated at the IA. Point estimate and confidence interval (CI) estimates of the Day 29 hsCRP change from Baseline difference, comparing the 2 treatment groups, will be estimated for use in internal decision making. In addition, the change in DAS28-hsCRP from Baseline to Day 29 will be evaluated using the same approach as used for change in hsCRP. The IA will not be used to alter the design of the trial. No adjustment will be made to the 1-sided α of 10% to be used in the primary analysis at the end of the first treatment period (12 weeks).

The Day 85 ACR20 proportion and a 2-sided 80% CI will be provided for each treatment group. The difference (Δ) in 12-week ACR20 proportions between M2951+MTX and placebo + MTX, and a 2 sided 80% CI for Δ, based on the "corrected Miettinen-Nurminen" method will be provided. The primary analysis will be based on the mITT Analysis Set using Last Observation Carried Forward for imputation of missing values and will reject the null hypothesis (H0): $\Delta \leq 0$ if the lower limit of the CI exceeds zero.

The Day 85 ACR50 and ACR70 proportions will be analyzed in the same way as the Day 85 ACR20 proportions.

Safety data will be listed and summarized using descriptive statistics.

In addition, efficacy and safety data collected during the open-label extension period will be summarized.

Pharmacokinetics

M2951 concentrations and PK parameters will be listed and summarized by trial day.

Pharmacodynamics

Data on immunoglobulin levels, CCI [redacted], and B cell numbers and subsets will be summarized in tabular and/or graphic format, as appropriate to the data. Potential PK/PD correlations may be explored and reported separately.

7 Sample Size/Randomization

The trial is powered to demonstrate superiority of M2951 + MTX to placebo + MTX in ACR20 response proportion after 84 days of treatment, using a test of the null hypothesis H0: $\Delta \leq 0$ based on the 2-sided 80% confidence interval (CI) for Δ, where Δ is the difference in Week 12 ACR20 response proportion comparing the 2 treatment groups. The null hypothesis H0 will be rejected if the lower limit of the 2-sided 80% CI exceeds zero. It is assumed that the Week 12 ACR20 response proportion in the placebo + MTX treatment group is 0.21 (1), and that the difference in Week 12 ACR20 response proportions between the 2 treatment groups under the alternative hypothesis H1: $\Delta > 0$ is $\Delta = 0.25$ (1, 2).

If the response proportions in the 2 treatment groups are 0.46 and 0.21, respectively, under the alternative hypothesis, then a sample size of approximately 60 mITT subjects randomized 1:1, M2951+MTX:placebo+MTX, provides 80.2% power to demonstrate superiority at the 1-sided

10% significance level. The power calculation is based on a simulation of the 2-sided 80% CI for Δ estimated using the "corrected Miettinen-Nurminen" method (3) conducted in R Cran software version 2.15.2 using the function PropCIs::diffscoreci (version 0.2 5, 2015).

A total of approximately 64 subjects will be randomized to provide 60 mITT subjects in the presence of a 6.25% dropout rate.

Eligible subjects will be randomized in a blinded fashion by the IWRS in a ratio of 1:1 to M2951 dose of 50 mg bid + MTX or placebo + MTX, CCI

8 Overview of Planned Analyses

8.1 Independent Monitoring Committee review

An IDMC will be set up to continually review available safety and tolerability data and will be mandated to make immediate decisions regarding the conduct of the trial; no additional risk minimization measures are proposed. IDMC meetings will take place every 4 months or as requested by the IDMC members and at the IA. Details of the statistical analyses of the IDMC are provided in the IDMC/IA SAP and the IDMC charter.

8.2 Interim Analysis

An IA will be conducted with the percent change in hsCRP from Baseline to Day 29 as key endpoint evaluated at the IA. Details of the statistical analyses of the IA are provided in the IDMC/IA SAP.

8.3 Primary Analysis

When the last subject has been administered the last dose of IMP during the 12-Week treatment period, the protocol violations are determined, the SAP is signed and the Week 12 database is partially locked for the primary analysis, the drug code will be broken and made available for the primary data analysis. The definition and details of partial database lock for the primary analysis will be documented in the Data Handling Report which will be finalized before the database partial lock for the primary analysis.

8.4 Final Analysis

The final analysis will occur only when the last subject completes all study parts (or discontinues study prematurely), the protocol violations are determined and the database is locked for the final analysis.

9 Changes to the Planned Analyses in the Clinical Trial Protocol

Computation of point estimates and CIs of differences from baseline in hsCRP and Disease Activity Score based on 28 joints and high-sensitivity C-Reactive Protein (DAS28-hsCRP)
Details were added to complete the planned analysis of hsCRP and DAS28-hsCRP.
In the SAP, it has been further specified that:

- the differences in mean changes between M2951+MTX and Placebo+MTX will be considered for hsCRP and DAS28-hsCRP, along with the 2-sided 80% CI for the difference
- the estimates and the CI limits will be derived using an Analysis of Covariance (ANCOVA) model with baseline as a covariate and treatment group as an independent variable
- Percent Change From Baseline (CFB) of hsCRP will be provided as summary statistics

Missing hsCRP and DAS28-hsCRP imputation
Missing hsCRP values and DAS28-hsCRP scores will be imputed for the primary and final analyses:

- Missing hsCRP values will be imputed using the same rule used for the analysis of the primary endpoint, ie, post-baseline Last Observation Carried Forward (LOCF)
- Missing components of DAS28-hsCRP will be imputed using the same rule as the primary endpoint, i.e., post-baseline Last Observation Carried Forward (LOCF)

Imputation of hsCRP and DAS28-hsCRP is not applicable for open-label extension (OLE) period, only for 12-week treatment period.

CCI [redacted]

OLE PD endpoints
The Schedule of Assessments table in protocol specifies immunoglobulin and B cells and subtypes to be collected during the optional OLE period. A description of the analysis of these data is missing from the "Endpoints for Open-label Extension" section of the protocol. A separate analysis set has been added here in Section 10.2.

CTCAE Grades for Relevant Laboratory Parameters
The Grade 1 definitions for the laboratory parameters AST, ALT, bilirubin, amylase, lipase, and creatinine in Table 3 in the protocol were cited incorrectly. For these parameters, Grade 1 will be defined as in CTCAE, version 4.03. In fact, all laboratory abnormalities will be graded as specified in CTCAE, version 4.03; Table 3 of the protocol will not be used.

PK statistical presentation
The protocol specified that M2951 concentrations and PK parameters would be described with standard error of mean and 25th and 75th percentile. For this study, M2951 concentrations and PK parameters will be summarized as described in Specific rules for PK parameters, Section 11. Also specified in the protocol was an additional descriptive statistics table for BLQ PK values set to missing, this will not be presented for this study.

Change of IA trigger
The protocol mentioned that IA would be conducted when subjects from the first 50% of the planned enrollment (ie 32 subjects) have reached Day 29 of treatment and have Day 29 data, or have discontinued treatment prior to Day 29. The trigger has been updated and the change from protocol is detailed in IDMC/IA SAP.

OLE extension endpoints
For the final analysis, in addition to the ACR50 and ACR70, the ACR20 will be also analyzed as an open-label extension endpoint.

OLE mITT definition
The definition of OLE mITT set has been updated. It will include all subjects who received at least 1 dose of M2951 during open-label extension period and have at least 1 available efficacy evaluation at a time point after first dose in open-label extension period, instead of 1 available ACR20 evaluation.

10 Protocol Deviations and Analysis Sets

10.1 Definition of Protocol Deviations and Analysis Sets

No per Protocol analysis is planned.

Important protocol deviations are protocol deviations that might significantly affect the completeness, accuracy, and/or reliability of the study data or that might significantly affect a subject's rights, safety, or well-being.

The following important deviations will be identified and confirmed prior to or at the Data Review Meeting, which will occur before the database lock. The outcome of the Data Review Meeting will document the important protocol deviations as well as the finalization of the analysis population in a memo.

- Subjects that are dosed on the study despite not satisfying the inclusion criteria
- Subjects that develop withdrawal criteria whilst on the study but are not withdrawn
- Subjects that receive the wrong treatment or an incorrect dose
- Subjects that receive an excluded concomitant medication
- Deviation from Good Clinical Practice

- Inclusion and exclusion criteria violations
- Other violations/events that may have a relevant influence on the pharmacokinetic/pharmacodynamic (PK/PD) analysis (eg, adverse events (AEs), vomiting, sample processing errors, inaccurate dosing, etc.)

Any important protocol deviations that lead to the exclusion of a subject from an analysis set will be considered clinically important (see Section 10.2).

All important protocol deviations should be documented in Clinical Data Interchange Standards Consortium Study Data Tabulation Model (SDTM) (or Analysis Data Model (ADaM) for those derived by programming) whether identified through sites monitoring, medical review and/or programming based on the inclusion/exclusion criteria presented in the protocol. Important protocol deviations to be identified by programming as well as all clinically important protocol deviations are listed and described in Appendix 3.

10.2 Definition of Analysis Sets and Subgroups

Screening Analysis Set (SCR)
The SCR includes all subjects who signed the informed consent.

Safety Analysis Set (SAF)
The SAF will include all subjects who received at least 1 dose of M2951 or placebo. The SAF will be used for all demographic/baseline characteristics and safety analyses.

Analyses performed on the SAF will consider subjects as treated. Subjects will be allocated to active arm if at least one dose of M2951 was taken during first 12-week treatment period.

The OLE SAF will include all subjects who received at least 1 dose of M2951 during the OLE period. The OLE SAF will be used for all safety analyses at the time of the final analysis.

mITT Analysis Set
The mITT Analysis Set will include all subjects who received at least 1 dose of M2951 or placebo and have at least 1 available ACR20 evaluation at a time point postdose. The mITT Analysis Set will be used for the analysis of all efficacy endpoints. Subjects will be analyzed as randomized.

The OLE mITT Analysis Set will include all subjects who received at least 1 dose of M2951 during OLE period and have at least 1 available efficacy evaluation (ACR components, ESR, anti-CCP, RF) at a time point after first dose in OLE period. The OLE mITT Analysis Set will be used for the analysis of efficacy endpoints at the time of the final analysis.

Pharmacokinetic Analysis Set
The PK Analysis Set will include all subjects who receive at least 1 dose of M2951 and have at least 1 quantifiable M2951 plasma concentration at a scheduled PK time point postdose without any important deviations or events that may impact the quality of the data or alter the evaluation of the PK. The PK Analysis Set will be used for all PK analyses. Subjects who receive placebo will not be included.

Pharmacodynamic Analysis Set
The PD Analysis Set will include all subjects who receive at least 1 dose of M2951 or placebo and have at least 1 measured PD endpoint, not including CCI, at a scheduled PD time point after first dose without deviations or important events affecting PD, and who provide evaluable PD data. All PD analyses, not including CCI, will be based on this analysis set.

Deviations, or changes to the procedures or events that may have a relevant influence on the PK may also influence PD analysis or impact the quality of the PD data and may lead to exclusion from the PD Analysis Set. Please see examples of deviations or events listed in Section 12.2.2.

In the case subjects or data are excluded from the PD analysis set, their PD data will be listed in the CSR and flagged.

The OLE PD Analysis Set will include all subjects who receive at least 1 dose of M2951 during OLE period and have at least 1 available PD endpoint at a time point after first dose in OLE period. All PD analyses, not including CCI, will be based on this OLE PD Analysis Set at time of final analysis.

CCI

CCI

Randomized Analysis Set
The randomized analysis set will include all randomized subjects. This set will be used to provide a listing for checking whether randomization occurred as planned.

The "as-treated" principle will be applied to all evaluations, except those based on the mITT Analysis Set, ie, subjects who receive a treatment other than the one assigned by randomization will be analyzed by the actual treatment received. The use of the analysis sets in the different analyses is summarized in the following table:

Table 1: Summary of analyses by sets

Analyses	SCR	SAF	OLE SAF	mITT Analysis Set	OLE mITT Analysis Set	PK Analysis Set	PD Analysis Set	OLE PD Analysis Set	CCI	Randomized Analysis Set
Subject Disposition status	✓									
Analysis sets	✓									
Subject disposition by visit										✓
Important Protocol Deviations										✓
Demographic and Baseline Assessments		✓								
Prior Medications		✓								
Concomitant Medications		✓	✓							
Concurrent procedures		✓	✓							
Compliance and Exposure		✓	✓							
Safety Analysis		✓	✓							
Primary Efficacy Analysis				✓						
Secondary Efficacy Analysis				✓						
OLE Efficacy Analysis					✓					
PK Analysis						✓				
PD Analysis							✓	✓		
CCI										
Randomization listing										✓

Subgroup analysis:

Subgroup analyses will be performed on randomized subjects with either:

- hsCRP at screening ≥ 3.6 mg/L
- hsCRP at baseline ≥ 3.6 mg/L

- hsCRP at screening ≥ 5 mg/L
- hsCRP at baseline ≥ 5 mg/L

Details of these subgroup analyses can be found in Table 14 of Section 16.

11 General Specifications for Statistical Analyses

All statistical analyses will be performed by PPD .

Treatment groups and IMP:
Treatment groups are defined as placebo + MTX and M2951 + MTX groups, except for the analyses including both the first 12-week treatment period and OLE period data, which will be summarized according to treatment groups 'Placebo/M2951 + MTX' and 'M2951 + MTX'. Unless otherwise indicated, all analyses will be presented separately for the two treatment groups. The IMPs are placebo and M2951.

Presentation of Tables/Figures
For the primary analysis at the end of first 12-week treatment period, drug code will be broken and tables and figures will be produced using true treatment groups and their true labels (ie, 'Placebo', 'M2951 50 mg BID'). For the final analysis, tables and figures will be produced using true treatment groups and the following labels: 'Placebo/M2951 50 mg BID' and 'M2951 50 mg BID/M2951 50 mg BID'.

Tables and figures will be sorted by treatment group (in order of 'Placebo' and 'M2951 50 mg BID', or 'Placebo/M2951 50 mg BID' and 'M2951 50 mg BID/M2951 50 mg BID' for final analysis) and chronological scheduled time point (where applicable).

Presentation of Listings:
All data recorded during the trial will be presented in individual data listings. All listings are sorted by treatment group (in order of 'Placebo' and 'M2951 50 mg BID', or 'Placebo/M2951 50 mg BID' and 'M2951 50 mg BID/M2951 50 mg BID' for final analysis), subject, and scheduled time point (where applicable), if not otherwise stated.

Data handling for the planned analysis:
For the primary analysis, data will be handled as follows:

- Data from OLE, or safety follow-up posterior to OLE, will be excluded from the primary analysis. This exclusion will be performed at the ADaM level; no selection will be performed, neither at the data management level, nor at the SDTM level.
- As the database lock for the primary analysis will be based on data up to and including the last patient's last dose of IMP (LPLD) during 12-Week treatment period, all data from screening to LPLD during the 12-Week treatment period for all subjects will be locked and available in the database for the primary analysis.
- A cut-off date will be applied at the SDTM level: all data posterior to LPLD during 12-Week treatment period will be removed from the database. Details on the implementation of this cut-off date are available in the document named *SDTM Cut-off date implementation rules_BTKi(M2951)-Compound_v2.0.docx*
- For all subjects, data during the 12-week treatment period up to and including the LPLD will be included for the primary analysis except for the safety follow-up data. All safety follow-up data will be reported for the mITT population at the final analysis stage.
- Data selection for the primary and final analyses will be identified at the ADaM level first before any statistical calculations can start. Therefore, the data handlings described in all tables in the following sections of this SAP will depend on data selected in the ADaM level.

For the final analysis, data will be presented on subjects who agreed to participate in OLE.

Presentation of continuous and qualitative variables:
Continuous variables will be summarized using descriptive statistics, ie

- number of subjects (N)
- number and percentage of non-missing values
- number and percentage of missing values
- mean, standard deviation (SD)
- median, 25th Percentile - 75th Percentile (Q1-Q3)
- minimum and maximum

For both continuous and qualitative variables, percentages such as 0% or 100% should be reported with the same format used for the column, together with the count of observations. For example, if the count of observations is zero, then display '0 (0.0)'; if the count of observations represents 100%, then display 'xx (100.0)'.

Qualitative variables will be summarized by counts and percentages. Unless otherwise stated, the calculation of proportions will be based on the number of subjects in the analysis set of interest. Therefore counts of missing observations will be included in the denominator and presented as a separate category.

In case the analysis refers only to certain visits, percentages will be based on the number of subjects in the analysis set of interest, unless otherwise specified. Total of missing and non-missing observations at each time-point will reflect the population still in the trial at that time. For example, if a subject is still in the trial at the time-point but with missing data, it should be counted in the number of missing observations.

Pharmacodynamic endpoints will be reported as described above for presentation of continuous and qualitative variables.

Specific rules for PK concentrations and PK parameters

Plasma concentrations of M2951 will be presented descriptively by study day and nominal time point for the PK analysis set including number of subjects (N), number of observations (n), mean, SD, coefficient of variation (CV%) median, minimum, and maximum. The below limit of quantification (BLQ) PK concentration values will be set to zero prior to computing descriptive statistics of PK concentrations. Missing concentrations will be omitted from the calculation of descriptive statistics.

Individual plasma and mean concentration-time plots will be provided using a linear and semi-logarithmic scale.

Descriptive statistics for PK parameters (see Section 16.3.2 for calculations) will be summarized by study day (N, n, arithmetic mean, SD of the arithmetic mean, CV%, median, minimum, maximum geometric mean [GeoMean], geometric coefficient of variation [GeoCV%], and the 95% CI for the GeoMean). For t_{max}, only n, minimum, median, and maximum will be reported. For calculation and presentation of PK descriptive statistics, and the inferential statistical analysis, values as presented in the data listings will be used as the source data. Minimum, maximum, mean, GeoMean, median, Q1, Q3, and 95% CI for mean or GeoMean will be presented to the same significant figure precision as listed data. Standard deviation will be presented to 1 significant figure more than the precision of the data listed. The CV% and GeoCV% will always be reported to 1 decimal place. Derived parameters will be reported to 3 significant figures, except for parameters that are taken directly from concentration data which will be reported using similar precision to those from which they were derived or elapsed sampling times which will be reported as 2 decimal places.

Definition of baseline:
For the purpose of statistical analysis, baseline is defined as the last non-missing measurement prior to the first dose of study drug as described in the following table:

Table 2: Definition of baseline

Randomized	Time Period	Baseline
Placebo+MTX	First 12-week treatment period	Last non-missing measurement prior to first dose of study drug of the 12-week treatment period
	OLE Period	Last non-missing measurement prior to the first dose of M2951 of the OLE period
M2951+MTX	First 12-week treatment period	Last non-missing measurement prior to first dose of study drug of the 12-week treatment period
	OLE Period	Last non-missing measurement prior to first dose of study drug of the 12-week treatment period

Definition of CFB:
CFB and percent CFB will be computed as follows:

- CFB = visit value – baseline value
- Percent CFB = 100 * (visit value – baseline value) / baseline value

Definition of duration:
Unless otherwise specified, duration will be calculated as the difference between start and stop date + 1 (e.g. duration of AE (days) = AE start date – AE stop date + 1). Unless otherwise specified, missing dates will not be imputed.

The time since an event will be calculated as:

- date of event minus reference date +1 (eg days in study at onset of AE = AE start date - date of randomization + 1) if date of event is equal or greater than reference date
- date of event minus reference date (eg days in study at onset of AE = AE start date - date of randomization) otherwise

Conversion factors:
The following conversion factors will be used to convert days into months or years: 1 month = 30.4375 days, 1 year = 365.25 days. For time windows calculation, 1 month is expressed as 30 days.

Definition of Body Mass Index (BMI) (kg/m^2):
BMI will be computed as weight / ($height^2$), where weight is expressed in kg and height in m.

Handling of missing data:
The Table 3 summarizes how missing data will be handled for each ACR component and DAS28-hsCRP:

Table 3: Missing data handling of ACR and DAS28-hsCRP

Period covered	Parameter	Methods of imputation
12-Week treatment period	ACR	1. LOCF-NR (Primary analysis): a. Imputation of each component using LOCF b. Derivation of ACR (responder/non-responder) c. If the subject discontinues from treatment, the subject will be considered as non-responder for all visits after discontinuation 2. Non-responder imputation (sensitivity analysis) a. Derivation of ACR (responder/non-responder) using observed ACR components b. If the ACR is missing, then the ACR will be assigned as non-responder 3. Other sensitivity analysis (see Table 13)
	DAS28-hsCRP	LOCF: a. Imputation on each component using LOCF b. Derivation of DAS28-hsCRP at each timepoint
OLE period	ACR	No imputation
	DAS28-hsCRP	No imputation

Unless otherwise specified, missing data will not be replaced.

In all subject data listings, imputed values will be presented. In all listings imputed information will be flagged. Non imputed partial dates will be presented in a format such as "_____YYYY". Where imputation is defined, an imputed date with flag (ie, D for day, M for month) will be reported.

Missing statistics, eg when they cannot be calculated, should be presented as 'nd' (denoting "not done"). For example, if n=1, the measure of variability (SD) cannot be computed and should be presented as 'nd'.

In case of zero records, an empty output with 0 occurrences, or a sentence mentioning that there are no data, will be presented. For tables of AEs and Deaths (outputs required for EudraCT and/or clinicaltrial.gov), if there are no observations, the output must contain the first line 'Subject with…' or 'Subject who died' displayed with 0 occurrence.

If the System Organ Class (SOC) or Anatomical Therapeutic Class (ATC) term is missing/not coded yet, then 'Uncoded SOC' (or 'Uncoded ATC') will be indicated at the ADaM level. When a Preferred Term (PT) is missing, it will be set to 'Uncoded PT: verbatim text'.

Treatment day:
Treatment day is defined as relative day to the start date of treatment. Treatment day will be calculated accordingly, ie:

- For primary analysis, treatment day 1 is defined as the date of first administration of any IMP.
- For final analysis, treatment day 1 is defined as the first administration of M2951:
    - For subjects who switch from placebo to M2951, the treatment day 1 will be defined as the first administration of M2951 in OLE period
    - For subjects who have taken M2951 during 12-Week treatment period and continue with the same treatment during OLE, the treatment day 1 will be defined as the first administration of M2951 during 12-Week treatment period

The day before is defined as Treatment day –1 (no Treatment day 0 is defined).

Repeated and unscheduled measurements:
Data collected at unscheduled visits will be included and analyzed for safety analyses in the same fashion as the data collected at scheduled visits. Unscheduled measurements will also be used for shift tables.

Gap between 12-Week treatment period and OLE period
Subjects with a gap strictly greater than 4 weeks between 12-Week treatment period and OLE period will be analyzed the same way as subjects without a gap. Specific data listings on safety data (AEs and all laboratory data) will be produced for these subjects for the final analysis.

Time windows
As a measurement may have been performed out of window, data will be re-allocated according to time windows defined in Table 4, so that each measurement is allocated to the closest timepoint. In case of a gap between 12-Week treatment period and OLE, time windows allocation will be handled case by case at the time of statistical analysis.

For efficacy and safety analyses, each measurement will be assigned an analysis visit number first according to Table 4. The analysis visit will then be used for all missing data imputations, analysis variable derivations, statistical calculations and presentations. For patient data listings by timepoint, the visit as collected in the database as well as the analysis visit will be displayed.

Table 4: Time windows

Analysis	Treatment groups	Period covered	Interval (days) [Date of assessment – Date of first dose + 1]	Analysis Visit
Primary analysis	Placebo or M2951 50 mg BID	12-Week treatment period	1	Day 1
			[2 ; 11)	Week 1 – Day 7
			[11 ; 21)	Week 2 – Day 14
			[21 ; 35)	Week 4 – Day 28
			[35 ; 49)	Week 6 – Day 42
			[49 ; 70)	Week 8 – Day 56
			[70 ; 91)	Week 12 – Day 84
Final analysis	Placebo/M2951 50 mg BID	12-Week treatment period	NA	NA
		OLE period	1	D1 – OLE
			[2 ; 22)	D14 – OLE
			[22 ; 60)	Month 1 – OLE
			[60 ; 135)	Month 3 – OLE
			[135 ; 187)	Month 6 – OLE
	M2951 50 mg BID/ M2951 50 mg BID	12-Week treatment period	1	Day 1
			[2 ; 11)	Week 1 – Day 7
			[11 ; 21)	Week 2 – Day 14
			[21 ; 35)	Week 4 – Day 28
			[35 ; 49)	Week 6 – Day 42
			[49 ; 70)	Week 8 – Day 56
			[70 ; 91)	Week 12 – Day 84
		OLE period	[70 ; 91)	D1 – OLE
			[91 ; 106)	D14 – OLE
			[106 ; 144)	Month 1 – OLE
			[144 ; 219)	Month 3 – OLE
			[219 ; 271)	Month 6 – OLE

Software:

Non-compartmental computation of PK parameters will be performed using Phoenix® WinNonlin® Version 6.4, or higher (PPD). Graphics may be prepared with SAS Version 9.4, or higher; SigmaPlot® 12.5, or higher (PPD); or Phoenix WinNonlin 6.4, or higher.

All other statistical analyses will be performed using SAS® (PPD).

12 Trial Subjects

The subsections in this section include specifications for reporting subject disposition and treatment/trial discontinuations. Additionally procedures for reporting protocol deviations are provided.

12.1 Disposition of Subjects and Discontinuations

A table on screened subjects describing the number and percent of subjects in each of following disposition categories will be produced by treatment group:

Table 5: Disposition

	Primary analysis	Final analysis
Total number of screened subjects, ie subjects that gave informed consent (overall summary only)	✓	✓
Number of subjects who discontinued prior to randomization and reason (overall summary only)	✓	✓
Number of randomized subjects	✓	✓
Number of randomized subjects who did not start treatment	✓	✓
Number of randomized subjects who completed 12-week double-blind treatment period	✓	✓
Number of randomized subjects who permanently discontinued treatment during 12-week double-blind treatment period and reason	✓	✓
Number of randomized subjects who discontinued from trial after randomization during 12-week double-blind treatment period and reason	✓	✓
Number of randomized subjects who declined to participate in OLE and discontinued from trial after completion of 12-week double-blind treatment period during safety follow-up and reason	✓	✓
Number of randomized subjects who declined to participate in OLE and completed safety follow-up period according to protocol	✓	✓
Number of randomized subjects who agreed to participate in OLE period	✓	✓
Number of randomized subjects with a gap > 4 weeks between double-blind period and OLE period		✓
Number of randomized subjects who completed the OLE period		✓
Number of randomized subjects who permanently discontinued treatment during OLE period and reason		✓
Number of randomized subjects who discontinued from trial after randomization during OLE period and reason		✓
Number of randomized subjects who completed safety follow-up after OLE according to protocol		✓
Number of randomized subjects who discontinued from trial after completion of OLE period during safety follow-up and reason		✓

For the primary analysis, a table of randomized subjects not treated as randomized, including reason for not being treated as randomized will be produced by randomized treatment group.

A table based on screened subjects, describing the number and percent of subjects in each analysis set by treatment group, will be produced:

Table 6: Analysis sets

	Primary analysis	Final analysis
Number of screened subjects	✓	✓
Number of randomized subjects	✓	✓
Number of subjects included in the SAF	✓	✓
Number of subjects included in the OLE SAF		✓
Number of subjects included in the PK Analysis Set	✓	✓
Number of subjects included in the PD Analysis Set	✓	✓
Number of subjects included in the OLE PD Analysis Set		✓
Number of subjects included in the CCI [redacted]	✓	✓
Number of subjects included in the mITT Analysis Set	✓	✓
Number of subjects included in the OLE mITT Analysis Set		✓

A table based on screened subjects, describing the number and percent of subjects in each region (eg Europe, European Economic Area (as required by EudraCT) and North America), country within region and site, by analysis set, will be produced.

A table based on randomized subjects with patients still on trial at each visit will be also produced. This table will also include the number of patients who discontinued before each visit.

Corresponding individual listings will be prepared. Discontinued subjects (from treatment or study) will be listed with their reason for withdrawal (from treatment or study). Additionally, a listing of the subjects screened but not randomized will be produced with the reason for not being randomized. A listing of randomized subjects with subject number, randomization date, and randomized treatment group, ordered by randomization number within randomization stratum, will be produced for the purpose of assessing whether randomization was conducted as planned.

The list of re-screened subjects and corresponding subject identifiers will be provided. These subjects' identifiers will be listed. Only the second subject identifier will be used in statistical descriptions and analyses, while the first identifier will not be considered/excluded from the disposition counts.

12.2 Protocol Deviations

The Table 7 summarizes how protocol deviations will be handled for both primary and final analysis:

Table 7: Protocol deviations

Analysis	Period covered for protocol deviations reporting	Treatment Group	Observation Period	Analysis Set
Primary analysis	12-Week treatment period	Placebo or M2951 50 mg BID	First dose of 12-Week treatment period up to first dose of OLE, or last observed dose during the 12-Week treatment period if subject either prematurely discontinues from treatment or declines to participate in OLE	Randomized
Final analysis	OLE	Placebo/M2951 50 mg BID	First dose of OLE to last observed dose	Randomized and entered into OLE
	12-Week treatment period + OLE	M2951 50 mg BID/M2951 50 mg BID	First dose to last observed dose	Randomized and entered into OLE

12.2.1 Important Protocol Deviations

The following summary tables and listings of important protocol deviations will be provided (separately for pre-/post inclusion deviations):

- Table providing frequency for each type of important protocol deviation
- Listing of important protocol deviations

12.2.2 Clinically Important Protocol Deviations

Clinically important protocol deviations, as defined in Section 10.1 and Section 10.2, will be summarized and listed.

Deviations, or changes to the procedures or events that may have a relevant influence on the PK/PD analysis or impact the quality of the PK/PD data may lead to exclusion from the PK/PD analysis sets. Examples include, but may not be limited to, the following:

- Adverse events, vomiting, diarrhea etc.
- Vomiting following oral dosing (these instances will be discussed on a case-by-case basis)
- Sample processing errors that may lead to inaccurate bioanalytical results
- Inaccurate dosing or dosing errors
- Concomitant medication violations (see Section 6.5.2. of the protocol)

In the case of a deviation or event that may affect PK, PK data collected during the affected study day will be excluded from the study results after consultation and agreement with Sponsor or sponsor representative. Other changes to the procedures or events which do not impact the quality of the PK data will not be considered clinically important protocol deviations. A common example of a protocol deviation that is not clinically important is a missed PK blood sample or deviations from PK blood collection times, with the requirement that the actual time of collection is recorded. An exception to a PK blood sample deviation defined as not clinically important is the predose sample which will be considered a deviation if it is not collected predose, and the 6 hour sample if missing or collected 10 minutes outside the scheduled window.

If subjects or select M2951 concentrations are excluded from the PK analysis set, the excluded PK data will be listed in the CSR and flagged.

13 Demographics and Other Baseline Characteristics

Demographics and baseline characteristics will only be summarized for the primary analysis.

Demographics and baseline characteristics will be summarized on SAF and presented by treatment group and overall. All data will be listed on SAF.

13.1 Demographics

- Demographic characteristics
  - Sex: male, female
  - Race: White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Other
  - Ethnicity: Hispanic or Latino, Not Hispanic or Latino
  - Age (years) at informed consent: summary statistics
  - Age categories :
    - < 65 years, ≥ 65 years
- Geographic Region (eg Europe, North America)
- European Economic Area (Yes/No)

Specifications for computation:

- Age (years):

- (date of given informed consent - date of birth + 1) / 365.25
- In case of missing day for at least one date, but month and year available for both dates:

  For the derivation of age, the day of informed consent and the day of birth will be set to 1 and the formula above will be used
- In case of missing month for at least one date, but year available for both dates:

  For the derivation of age, the day and the month of informed consent and the day and month of birth will be set to 1 and the formula above will be used

- Site codes will be used for the determination of the subject's geographic region and European Economic Area membership.

13.2 Medical History

The medical history will only be summarized for the primary analysis.

The medical history will be summarized using Medical Dictionary for Regulatory Activities (MedDRA), current version, PT as event category and MedDRA SOC body term as Body System category. The MedDRA version used will be indicated in footnote. Medical history will be tabulated by SOC and PT. SOC and PT will be alphabetically sorted. Medical history will be also listed.

13.3 Other Baseline Characteristics

13.3.1 Disease history

Information on RA baseline disease characteristics collected during screening will only be summarized for the primary analysis.

Information on RA baseline disease characteristics collected during screening will be summarized in total and by treatment arm. Descriptive statistics will be presented for:

- Time (months) since confirmed diagnosis of RA according to 2010 ACR/The EULAR RA classification criteria. Time will be computed as (Date of Informed Consent Signature – Date of confirmed diagnosis of RA) / 30.4375. If the date of confirmed diagnosis of RA is entirely missing, then time since confirmed diagnosis of RA will not be computed. In case of partial missing date of confirmed diagnosis of RA, this date will be imputed as follows:
  - If only the day is missing, then it will be replaced by the first day of the month
  - If both day and month are missing, then it will be replaced by the first of January
- The count of Tender Joint Count 68 (TJC68)
- The count of Swollen Joint Count 66 (SJC66)

- Physician's Global Assessment of Disease Activity using horizontal Visual Analog Scale (VAS) in mm
- Subject's Global Assessment of Disease Activity using horizontal VAS in mm
- Subject's Assessment of Pain using horizontal VAS in mm
- Subject's Assessment of Physical Function as measured by HAQ-DI will be derived (Appendix 1: HAQ-DI scoring) and described overall and by dimensions (dressing and grooming, getting up, eating, walking, hygiene, reach, grip, activities)
- hsCRP (mg/L) as a continuous variable and by category (normal/high based on central normal laboratory ranges)
- ESR (mm/Hr) as a continuous variable and by category (low/normal/high based on local normal ranges collected in eCRF)
- Acute-phase reactant (at least one test result is needed for the classification): normal hsCRP and normal ESR, abnormal hsCRP or abnormal ESR. Abnormal refers to either low or high normal range classification.
- RF (expressed in kU/L) as a continuous variable and classified as negative (≤ Upper Limit of Normal [ULN]) and positive (> ULN)
- Anti-CCP antibodies classified as a continuous variable and as negative (≤ ULN) and positive (> ULN)
- Serology (at least one test result is needed for the classification): negative RF and negative anti-CCP (≤ ULN), low-positive RF or low-positive anti-CCP (> ULN, ≤ 3*ULN), high-positive RF or high positive anti-CCP (> 3*ULN)
- DAS28-hsCRP, derived using formula presented in Appendix 2, summarized as a continuous variable and by disease activity category (< 2.6 [remission], ≥ 2.6 to ≤ 3.2 [low], > 3.2 to ≤ 5.1 [moderate] and > 5.1 [high]).

13.3.2 Other

Physical examination will only be summarized for the primary analysis.

Descriptive statistics will be presented for physical examination as continuous variables: height (cm), weight (kg) and BMI (kg/m²).

14 Previous or Concomitant Medications/Procedures

Medications/procedures will be presented according to Table 8:

Table 8: Data handling for medications/procedures

Analysis	Period covered	Treatment groups	Data summarized	Analysis sets
Primary analysis	12-Week treatment period	Placebo or M2951 50 mg BID	Previous medications + concomitant medications	SAF
Final analysis	OLE	Placebo/M2951 50 mg BID	Concomitant medications only	OLE SAF
	12-Week treatment period + OLE	M2951 50 mg BID/M2951 50 mg BID	Concomitant medications only	OLE SAF

The definition of previous or concomitant medication is presented in Table 9:

Table 9: Definition of previous/concomitant medication

Analysis	Period covered	Treatment groups	Definition
Primary analysis	12-Week treatment period	Placebo or M2951 50 mg BID	**Previous medications** are medications, other than trial medications, which either: 1. started and stopped before first administration of any IMP (Placebo or M2951). 2. started prior to the first administration of IMP (Placebo or M2951) and are taken by subjects on or after the first administration of IMP.
			Concomitant medications are medications, other than IMPs, which either: 1. started on or after the first administration of any IMP (Placebo or M2951). 2. started prior to the first administration of IMP (Placebo or M2951) and are taken by subjects on or after the first administration of IMP.
Final analysis	OLE	Placebo/M2951 50 mg BID	**Concomitant medications** are medications, other than IMPs, which either: 1. started on or after the first administration of M2951 during OLE. 2. started prior the first administration of M2951 during OLE and are taken by subjects on or after the first administration of M2951 during OLE.
	12-Week treatment period + OLE	M2951 50 mg BID/M2951 50 mg BID	**Concomitant medications** are medications, other than IMPs, which either: 1. started on or after the first administration of M2951 during 12-Week treatment period. 2. started prior the first administration of M2951 during 12-Week treatment period and are taken by subjects on or after the first administration of M2951 during 12-Week treatment period.

Partial dates will be handled as follows:

- For previous medications, in case the date values will not allow a medication to be unequivocally allocated to previous medication, the medication will be considered as previous medication.

- For concomitant medications, in case the date values will not allow a medication to be unequivocally allocated to concomitant medication, the medication will be considered as concomitant medication.

The ATC-2nd level and PT will be tabulated as given from the World Health Organization Drug Dictionary (WHO-DD) current version. In case multiple ATCs are assigned to a drug, all ATC-2nd level will be used for reporting.

The number and proportion of subjects with previous or concomitant medications (previous medications will not be summarized for final analysis) will be separately summarized by treatment group and will be presented by descending frequency of ATC 2nd level term and then by descending frequency of PT in total column. If multiple ATCs/PTs have the same frequency, they will be sorted alphabetically. The WHO-DD version used will be indicated in footnote.

Previous or concomitant medications will be also listed.

All concurrent procedures, which were undertaken any time on trial, will be summarized. Concurrent procedures will be classified by medical review. Number of subjects with concurrent procedures (prior, on or after the first day of IMP) overall and by type of procedure (as classified by medical review) will be summarized by treatment group.

Concurrent procedures will be also listed.

15 Treatment Compliance and Exposure

Exposure

Exposure will be separately analyzed for the 12-week treatment period, for the OLE period and overall (ie 12-Week treatment period + OLE).

Exposure for 12-week treatment period will only be presented for primary analysis, whereas exposure for OLE period and overall exposure will only be presented for final analysis.

The Table 10 summarizes how exposure time will be computed:

Table 10: Exposure time

Analysis	Period covered	Treatment groups	Exposure time	Analysis sets
Primary analysis	12-Week treatment period	Placebo or M2951 50 mg BID	From first dose of any IMP during 12-Week treatment period to last observed dose of 12-Week treatment period	SAF
Final analysis	OLE	Placebo/M2951 50 mg BID	From first dose of M2951 to last dose of M2951 during OLE	OLE SAF
	12-Week treatment period + OLE	M2951 50 mg BID/M2951 50 mg BID	From the first dose of M2951 during 12-Week treatment period to the last dose of M2951 during OLE	OLE SAF

12-week treatment period

Exposure for the 12-week treatment period will be presented on SAF.

For the 12-week treatment period, subjects will receive 50 mg bid M2951 or placebo administered as 2 matching oral capsules twice daily for 84 days. Each treatment kit contains enough medication for administration for 1 week + 2 days (ie 36 capsules). On trial visit days, subjects will be given treatment kits containing the number of capsules needed up to next planned trial visit. Subjects will self-administer IMP in a blinded manner at a set time each day (± 2 hours) as two matching capsules bid (morning and evening) for 12 weeks:

- Subjects randomized to placebo arm will take four placebo capsules filled with mannitol
- Subjects randomized to M2951 arm will take four capsules filled with 25 mg of M2951.

On trial visit days, IMP will be administered during the trial visit after the scheduled trial visit procedures (other than post treatment PK/PD sampling) are completed. Exposure time in weeks will be calculated according to the following formula:

$$\text{exposure (weeks)} = (\text{date of last dose - date of first dose} + 1)/7$$

First dose refers to the first administration of any IMP in 12-week treatment period.

Exposure time will be presented by summary statistics and according to the following categories:

- ≤ 1 week
- > 1 to 2 weeks
- > 2 to 4 weeks
- > 4 to 8 weeks
- > 8 weeks

The cumulative actual dose (mg) per subject for the 12-week treatment period will also be summarized for the active treatment group. The cumulative actual dose is defined as the sum of the total dose levels that the subject received between first and last treatment dose: number of ingested capsules * 25mg. The number of ingested capsules will be computed as the difference between the dispensed amount of capsules and the returned amount of capsules. Cumulative actual dose cannot be calculated for placebo treatment group.

Study drug administrations will also be listed by treatment group, and subject, with start/end dates of administration, and reason for dose change or no dose (if applicable).

OLE period

Exposure for the OLE period will only be presented on patients from OLE SAF who switch from Placebo to M2951 50 mg BID.

After completing the treatment period, subjects will be offered the opportunity to participate in an optional 6-month M2951 extension period on Week 12 (Visit 8). Subjects who received M2951 during the treatment period will continue with the same dose of M2951 (in tablet formulation). All subjects in the placebo group during the treatment period will discontinue placebo and start treatment with 50 mg bid M2951 tablets on Day 1 of OLE. On trial visit days, the IMP will be administered during the trial visit after the scheduled trial visit procedures are completed.
Subjects will take 2 matching oral tablets twice daily for 6 months. Subjects will self-administer the IMP at a set time each day (± 2 hours).

Exposure time in weeks will be calculated according to the following formula:

$$\text{exposure (weeks)} = (\text{date of last dose - date of first dose} + 1)/7$$

First dose refers to the first administration of M2951 in OLE period.

Exposure time will be presented by summary statistics and according to the following categories:

- ≤ 4 weeks
- > 4 to 8 weeks
- > 8 to 12 weeks
- > 12 to 16 weeks
- > 16 weeks

The cumulative actual dose (mg) per subject for the OLE period will also be summarized for both treatment groups. It is defined as for the 12-week treatment period.

Overall

Overall exposure will be presented on OLE SAF only for subjects who received M2951 during the 12-week treatment period and continue with the same dose of M2951 during OLE period.

Exposure time in weeks will be calculated according to the following formula:

$$\text{exposure (weeks)} = (\text{date of last dose - date of first dose} + 1)/7$$

First dose refers to the first administration of M2951 in the study.

Exposure time will be presented by summary statistics and according to the following categories:

- ≤ 6 weeks
- > 6 to 12 weeks
- > 12 to 18 weeks
- > 18 to 24 weeks
- > 24 weeks

The cumulative actual dose (mg) per subject will be computed as well. It is defined as for the 12-week treatment period or OLE period.

Study drug administrations will also be listed by treatment group, and subject, with start/end dates of administration and reason for dose change (if applicable).

Compliance

Compliance will be separately analyzed for the 12-week treatment period, the OLE period and overall, in the same manner as exposure was analyzed (see Table 10).

For each analysis, a listing of kit number with date of dispense, number of ingested and returned capsules/tablets will also be provided.

12-week treatment period

Compliance for the 12-week treatment period will be presented on SAF.

For the 12-week treatment period, compliance with treatment is defined as the number of capsules taken during a period divided by the number of capsules that should have been taken during that period, multiplied by 100 to yield a percentage, ie:

compliance for X weeks of treatment =

$$100*\left(\frac{\left((36 * X) - \text{number of returned capsules for X\_wk period}\right)}{4 * \text{number of days between first intake and last intake of X\_wk period}}\right)$$

Compliance with treatment and the number of ingested capsules will be tabulated overall (ie from first intake to last intake of 12-week treatment period).

OLE period

Compliance for the OLE period will be presented on OLE SAF.

For the OLE period, compliance with treatment is defined using the same formula as for the 12-week treatment period, replacing capsules with tablets.

Compliance with treatment and the number of ingested tablets will be tabulated overall (ie from first intake to last intake during OLE period).

Overall

Overall compliance will be presented on OLE SAF only for subjects who received M2951 during the 12-week treatment period and continue with the same dose of M2951 during OLE period.

Overall compliance with treatment is defined using the same formula as for the 12-week treatment period, replacing capsules with capsules/tablets.

Compliance with treatment and the number of ingested capsules/tablets will be tabulated overall (ie from first intake to last intake of the study).

16 Endpoint Evaluation

For all efficacy endpoints, data selection for primary and final analyses will be handled as indicated in the Table 11:

Table 11 : Data handling for efficacy analysis

Analysis	Population	Period covered	Treatment groups	Data to be analyzed
Primary analysis	mITT Analysis Set	12-Week treatment period until start of OLE or discontinuation of treatment prior to OLE, whichever is applicable	Placebo or M2951 50 mg BID	All data collected during the 12-Week treatment period up to start of OLE if applicable
Final analysis	OLE mITT Analysis set	OLE period only	Placebo/M2951 50 mg BID only	Data from OLE period only, or posterior to OLE (Safety follow-up / End of Study)
		Both 12-Week treatment period and OLE period	M2951 50 mg BID/M2951 50 mg BID only	Data from both periods (12-Week treatment and OLE period), ie all data collected throughout the study

The Table 12 summarizes the analyses of efficacy endpoints:

Table 12 : Summary of efficacy endpoints for the primary analysis

Analysis	Endpoint	Method	Missing Data Handling
Primary endpoint	ACR20 at Week 12/Day 84	(1) Primary: Computation of response rate, point estimate and 2-sided 80% CI for Δ based on the "corrected Miettinen-Nurminen" method (2) Supportive**: logistic regression with treatment and baseline hsCRP as covariates	Primary: LOCF-NR (Table 3) Sensitivity: See Table 13
Key secondary endpoint	hsCRP CFB at Week 4/Day 28 and Week 12/Day 84	(1) Primary: ANCOVA model (2) Supportive: Hodges-Lehmann estimation for 80% CI	Primary: LOCF

Analysis	Endpoint	Method	Missing Data Handling
Other secondary endpoints	ACR20 response at each visit	(1) Primary: Computation of response rate, point estimate and 2-sided 80% CI for Δ based on the "corrected Miettinen-Nurminen" method (2) Supportive: logistic regression with treatment and baseline hsCRP as covariates	Primary: LOCF-NR (Table 3) Sensitivity: See Table 13
	ACR50/ACR70 response at each visit	(1) Primary: Computation of response rate, point estimate and 2-sided 80% CI for Δ based on the "corrected Miettinen-Nurminen" method (2) Supportive: logistic regression with treatment and baseline hsCRP as covariate	Primary: LOCF-NR (Table 3) Sensitivity: See Table 13
	CFB and percent CFB at each visit for each ACR component, ESR, anti-CCP, RF	Summary Statistics	As observed
	CFB in DAS28-hsCRP score at Week 4/Day 28, and Week 12/Day 84	(1) Primary : ANCOVA model (2) Supportive: Hodges-Lehmann estimation for 80% CI and ANCOVA model with baseline hsCRP as covariate	Primary: LOCF of each component prior to DAS28 calculation
	Percentage of subjects with DAS28-hsCRP < 3.2 (well-controlled disease) at Week 12/Day 84	(1) Primary: Computation of response rate, point estimate and 2-sided 80% CI for Δ based on the "corrected Miettinen-Nurminen" method (2) Supportive: logistic regression with treatment and baseline hsCRP as covariate	Primary: LOCF-NR (Table 3)
	Percentage of subjects with DAS28-hsCRP < 2.6 at Week 12 / Day 84	(1) Primary: Computation of response rate, point estimate and 2-sided 80% CI for Δ based on the "corrected Miettinen-Nurminen" method (2) Supportive: logistic regression with treatment and baseline hsCRP as covariate	Primary: LOCF-NR (Table 3)
** All supportive analysis will be based on the primary missing data imputation method. No sensitivity analysis of missing data imputation will be performed for all supportive analyses.			

The Table 13 summarizes the sensitivity analyses to be performed on missing data imputations:

Table 13 : Sensitivity analysis for the primary analysis

Endpoint	Method
ACR Response	**Non-responder imputation (NR)**: ACR20 response rate will be evaluated on mITT analysis set. Subjects with missing ACR score will be considered as non-responders.
	Completely Missing at Random pattern (MCAR): ACR20 response rate will be evaluated as observed without any LOCF imputation on mITT analysis set.
	Missing At Random pattern (MAR): ACR20 will be imputed using multiple imputations process using MI SAS procedure as follows: 1. Monotone missing data structure will be created as follows: intermediate (non-monotone) missing data will be multiply imputed using the Markov chain Monte Carlo (MCMC) method and assuming MAR and multivariate normality. Transformation of data will only be used if there is a clear deviation from normality. The SAS procedure PROC MI with the MCMC option will be used with seed number=20170524. The number of burn-in iterations will be set to 200, which is the default value. Nevertheless, if diagnosis plots show that the convergence has not yet occurred, this will be adjusted. 2. Then, each component will be imputed with treatment group, hsCRP at baseline and data at previous visit from all ACR20 components as covariates**.** 3. Imputation will be repeated 1000 times. The ACR20 will then be calculated for each of the multiply imputed data sets from the imputed components. Results will be combined using MIANALYZE SAS procedure with the Rubin's rules.
	Missing Not At Random pattern (MNAR): ACR20 will be imputed using multiple imputations process with MI SAS procedure for the MNAR pattern, using controlled-arm data with the MNAR statement from SAS MI procedure.

Subgroup analyses will be performed on the randomized subjects with either:

- hsCRP at screening ≥ 3.6 mg/L
- hsCRP at baseline ≥ 3.6 mg/L
- hsCRP at screening ≥ 5 mg/L
- hsCRP at baseline ≥ 5 mg/L

The Table 14 summarizes the analyses to be performed for each subgroup:

Table 14 : Subgroup analysis for the primary analysis

Endpoint	Method	Missing Data Handling
ACR20/ACR50/ACR70 response at each visit	Primary: Computation of response rate, point estimate and 2-sided 80% CI for Δ based on the "corrected Miettinen-Nurminen" method	Primary: LOCF-NR (Table 3)
CFB in hsCRP and DAS28-hsCRP score at Week 4/Day 28 and Week 12/Day 84	Primary : ANCOVA model Supportive: Hodges-Lehmann estimation for 80% CI, only if the assumption of normality is violated	Primary and supportive: LOCF of each component prior to DAS28 calculation
Percentage of subjects with DAS28-hsCRP < 3.2 (well-controlled disease) at Week 12/Day 84	Primary: Computation of response rate, point estimate and 2-sided 80% CI for Δ based on the "corrected Miettinen-Nurminen" method	Primary: LOCF-NR (Table 3)
Percentage of subjects with DAS28-hsCRP < 2.6 at Week 12 / Day 84	Primary: Computation of response rate, point estimate and 2-sided 80% CI for Δ based on the "corrected Miettinen-Nurminen" method	Primary: LOCF-NR (Table 3)

16.1 Primary Endpoint Analyses

The primary endpoint, ACR20 response rate at Week 12, will be presented on the mITT Analysis Set.

The ACR20 is a primary efficacy measure for which, at Week 12, a subject must have at least 20% improvement in the following ACR Core Set values.

- TJC (being the number of joints assessed as "Pain/Tender only" or "Pain/Tender and Swollen" among the 68 joint count) and
- SJC (being the sum of joints assessed as "Swollen only" or "Pain/Tender and Swollen" among the 66 joint count) and
- An improvement of at least 20% in at least 3 of the following 5 assessments:
  - Subject's Global Assessment of Disease Activity
  - Subject's Assessment of Pain

- Subject's Assessment of Physical Function as measured by the HAQ-DI
- Physician's Global Assessment of Disease Activity
- Acute phase reactant as measured by hsCRP.

In this trial, ACR20 response calculations will use the HAQ-DI for the subject's assessment of physical function and hsCRP as the measure of acute phase reactant.

Improvement of at least 20% of these assessments is defined as follows:

- TJC (68 joint count): a decrease of at least 20% of the number of tender joint compared to baseline
- SJC (66 joint count): a decrease of at least 20% of the number of swollen joint compared to baseline
- Subject's Global Assessment of Disease Activity: a decrease of at least 20% on the VAS compared to baseline.
- Subject's Assessment of Pain: a decrease of at least 20% on the VAS compared to baseline.
- Subject's Assessment of Physical Function: a decrease of at least 20% of the HAQ-DI score compared to baseline.
- Physician's Global Assessment of Disease Activity: a decrease of at least 20% on the VAS compared to baseline
- Acute phase reactant: a decrease of at least 20% of hsCRP measurement compared to baseline

The hypotheses tested are as follows:

- H0: Difference in ACR20 response rates at Week 12 (M2951 versus placebo) ≤ 0.
- H1: Difference in ACR20 response rates at Week 12 (M2951 versus placebo) > 0.

Tests will be based on the 2-sided 80% CI for Δ.

The Week 12 ACR20 proportion and a 2-sided 80% CI will be provided for each treatment group. The difference Δ in ACR20 proportions between M2951 + MTX and placebo + MTX, and a 2-sided 80% CI for Δ, based on the "corrected Miettinen-Nurminen" method (3) - method of Mee interval corrected with Miettinen and Nurminen factor n/(n-1), n being the total number of evaluable randomized subjects - will be provided.

The 2-sided 80% CI for Δ based on the "corrected Miettinen-Nurminen" method will be computed with SAS® FREQ procedure, using RISKDIFF and CL=MN statements.

The primary analysis will be based on the mITT Analysis Set using post baseline LOCF for imputation of missing values for the ACR components. If a subject discontinued prior to Week 12, then the subject will be considered as ACR20 non-responder. The null hypothesis H0: $\Delta \leq 0$ will be rejected at the 1-sided 0.10 level if the lower limit of the 2-sided 80% CI exceeds zero.

The primary analysis will be repeated for subjects with baseline hsCRP ≥3.6mg/L, as detailed in Table 14.

Additional sensitivity analyses will be performed on ACR20 response as described in Table 13.

16.2 Secondary Endpoint Analyses

16.2.1 Key secondary endpoint analysis

The key secondary endpoint is the change in hsCRP from Baseline to Week 4 and will be presented on the mITT Analysis Set.

Descriptive statistics will be presented by treatment group for CFB as well as percent CFB of hsCRP. A point estimate and 2-sided 80% CI for the difference Δ in Week 4 hsCRP mean CFB between subjects treated with M2951 + MTX and subjects treated with placebo + MTX (reference) will be provided.

Point estimate and its CI limits will be estimated with SAS® GLM procedure using LSMEANS statement. The ANCOVA will include baseline as a covariate and treatment group as an independent variable. A sensitivity analysis will be performed using a non-parametric method. The Hodges-Lehmann difference median estimate and 80% CI will be calculated with SAS® NPAR1WAY procedure using HL statement. Asymptotic and exact CIs will be computed.

Assumptions for the ANCOVA models:

Validity of the model will be verified as follows:

- The homogeneity of variances between treatment groups will be checked using the Levene's test. A scatter plot of residuals versus fitted values will be also performed.
- In order to assess the independence of error terms, a residual lag plot will be performed, ie constructed by plotting residual (i) against residual (i-1).
- A scatter plot of the response variable against the covariate, using separate symbols for each treatment group will be performed to verify the presence of linear relationship between the covariate and the response variable, and that all treatment regression lines have the same slope.
- Homogeneity of regression slopes will be verified by first running the model with the interaction between the covariate and the treatment group. This assumption will be validated if this interaction is not significant. If this assumption is verified, then the interaction will be removed from the model.
- Normality of residuals will be checked by computing Skewness and Kurtosis, which should fall within the interval [-2;2]. Normal QQ-plot of residuals will be performed as well.

16.2.2 Other secondary endpoint analyses

The following efficacy endpoint analyses will be performed for the primary analysis and presented on the mITT Analysis Set:

- ACR20 response at each visit
- ACR50/ACR70 response at each visit
- Change and percent change in hsCRP from Baseline to Week 12
- CFB in DAS28-hsCRP score at Week 4 and Week 12
- Percentage of subjects with DAS28-hsCRP < 3.2 (well-controlled disease) at Week 12
- Percentage of subjects with DAS28-hsCRP < 2.6 at Week 12
- CFB in ESR, anti-CCP, RF (at Week 4 and Week 12)
- CFB to Week 12 in subject's self-assessments including:
  - Global assessment of disease activity (VAS)
  - Self-assessment of pain (VAS)
  - Self-assessment of disability (HAQ-DI)
- CFB to Week 12 in Physician's Global Assessment of Disease Activity (VAS).

The Week 12 hsCRP CFB will be analyzed in the same way as the Week 4 hsCRP CFB, ie descriptive statistics, ANCOVA model and sensitivity analysis using Hodges-Lehmann difference median estimate and 80% CI. This analysis will also be applied to DAS28-hsCRP score CFB at Week 4 and Week 12. The DAS28-hsCRP will be derived using the formula presented in Appendix 2.

ACR50 and ACR70 are defined in the same way as the ACR20 using at least 50% and 70% improvement, respectively. The ACR50 and ACR70 proportions will be analyzed in the same way as the Week 12 ACR20 proportions.

This analysis will also be applied to Week 12 "DAS28-hsCRP < 3.2" proportion and Week 12 "DAS28-hsCRP < 2.6" proportion, but without the sensitivity analysis.

Proportions of ACR20/50/70, "DAS28-hsCRP < 3.2" and "DAS28-hsCRP < 2.6" will also be presented by treatment group and time point as observed, ie without any imputation.

Descriptive statistics

Descriptive statistics will be tabulated by treatment group and time point for the following continuous variables, both absolute value and CFB. The summary statistics will be provided as observed without imputation of missing data.

- DAS28-hsCRP score
- hsCRP (CFB and percent CFB)
- ESR
- anti-CCP
- RF

- TJC68 (CFB and percent CFB)
- SJC66 (CFB and percent CFB)
- subject's global assessment of disease activity (CFB and percent CFB)
- subject's assessment of pain (CFB and percent CFB)
- subject's self-assessment of disability (HAQ-DI) (CFB and percent CFB)
- Physician's Global Assessment of Disease Activity (CFB and percent CFB)

Each of the following parameters will be also displayed by categories. The number and percentage of subjects at Week 4 and Week 12 will be computed and a shift table from baseline to Week 4 and Week 12 respectively will be provided. For all shift tables, data will be summarized as observed, ie no imputation on missing data. The following parameters will be presented into categories:

- hsCRP (mg/L) by category (normal/high based on central normal laboratory ranges)
- DAS28-hsCRP, derived using the formula presented in Appendix 2, summarized by disease activity category (< 2.6 [remission], ≥ 2.6 to ≤ 3.2 [low], > 3.2 to ≤ 5.1 [moderate] and > 5.1 [high]).
- ESR (mm/Hr) by category (low/normal/high based on local normal ranges collected in eCRF)
- Acute-phase reactant (at least one test result is needed for the classification): normal hsCRP and normal ESR, abnormal hsCRP or abnormal ESR. Abnormal refers to either below or above the normal range classification.
- RF (expressed in kU/L) classified as negative (≤ ULN) and positive (> ULN)
- anti-CCP classified as negative (≤ ULN) and positive (> ULN)
- Serology (at least one test result is needed for the classification): negative RF and negative anti-CCP (≤ ULN), low-positive RF or low-positive anti-CCP (> ULN, ≤ 3*ULN), high-positive RF or high positive anti-CCP (> 3*ULN)

Box plots will be performed for each ACR component and DAS28 at each timepoint. Both treatment groups will be presented in the same graph. One plot will present the CFB and another plot for the percent CFB). Data will be displayed as observed, ie without any imputation.

Line plots will be generated for ACR response at each visit over time for both primary and final analysis as follows:

- For the primary analysis, both treatment groups will be displayed.
- For the final analysis, only the M2951 50 mg BID/ M2951 50 mg BID subjects will be presented.
- Data will be presented using the LOCF-NR method, as for the primary endpoint.

16.3 Other Endpoint Analyses

16.3.1 Open-label Extension Endpoint Analyses

Descriptive statistics will be provided for data observed in the OLE period, and no imputation of missing data will be performed:

- Continuous endpoints: summary statistics will be provided for each group.
- Categorical endpoints: number and frequency of subjects in each category will be reported for each group.

The following efficacy endpoint analyses will be performed for the final analysis and presented on the OLE mITT Analysis Set:

- ACR20/ACR50/ACR70 response at Month 6
- Absolute value, CFB and percent CFB in hsCRP at Month 6
- Absolute value, CFB in DAS28-hsCRP score at Month 6
- Percentage of subjects with DAS28-hsCRP < 3.2 (well-controlled disease) at Month 6
- Percentage of subjects with DAS28-hsCRP < 2.6 at Month 6
- Absolute value, CFB in ESR, anti-CCP, RF at Month 6
- Absolute value, CFB and percent CFB in TJC68, SJC66 at Month 6
- Absolute value, CFB and percent CFB to Month 6 in subject's self-assessments including:
  - Global assessment of disease activity (VAS)
  - Self-assessment of pain (VAS)
  - Self-assessment of disability (HAQ-DI)
- Absolute value, CFB and percent CFB to Month 6 in Physician's Global Assessment of Disease Activity (VAS).

Each of the following parameters will be also displayed into categories. The number and percentage of subjects at Month 6 will be computed and a shift table from baseline to Month 6 will be provided. For all shift tables, data will be summarized as observed, ie no imputation on missing data. The following parameters will be presented into categories:

- hsCRP (mg/L) by category (normal/high based on central normal laboratory ranges)
- DAS28-hsCRP, derived using formula presented in Appendix 2, summarized by disease activity category (< 2.6 [remission], ≥ 2.6 to ≤ 3.2 [low], > 3.2 to ≤ 5.1 [moderate] and > 5.1 [high]).
- ESR (mm/Hr) by category (low/normal/high based on local normal ranges collected in eCRF)

- Acute-phase reactant (at least one test result is needed for the classification): normal hsCRP and normal ESR, abnormal hsCRP or abnormal ESR. Abnormal refers to either low or high normal range classification.

- RF (expressed in kU/L) classified as negative (≤ ULN) and positive (> ULN)

- anti-CCP classified as negative (≤ ULN) and positive (> ULN)

- Serology (at least one test result is needed for the classification): negative RF and negative anti-CCP (≤ULN), low-positive RF or low-positive anti-CCP (> ULN, ≤ 3*ULN), high-positive RF or high positive anti-CCP (> 3*ULN)

16.3.2 Pharmacokinetic Endpoint Analyses

Analysis of the PK of M2951 and the associated figures will be the responsibility of the clinical pharmacokineticist at PPD . The PK summaries and data listings as well as the statistical analysis of the PK parameters will be the responsibility of the study biostatistician at PPD .

For the PK analysis, predose sample M2951 concentrations that are BLQ will be assigned a numerical value of zero. On Day 1, any missing predose value will also be assigned a numerical value of zero for the calculation of AUC. Any anomalous concentration values observed at predose will be identified in the CSR. All postdose BLQ concentrations will be set to 'zero' in all cases. Graphical displays of mean (± SD for linear profiles only) plasma concentration-time profiles will be presented on linear and semi-logarithmic scales by dose level with Day 1 and Day 29 data on the same plot.

Non-compartmental computation of M2951 PK parameters will be performed. Pharmacokinetic parameters will be evaluated and listed for all patients who provide sufficient concentration-time data.

Individual PK parameters will be calculated using actual elapsed time from dose with a maximum of 14 significant digits in the time data (or using scheduled time if actual time is not available). The predose sample will be considered as if it had been taken simultaneously with the administration of IMP.

For each subject in the PK analysis set the following PK parameters will be calculated for M2951 on Day 1 and Day 29 following the morning dose on Days 1 and 29:

$AUC_{0\text{-}6h}$ Area under the plasma concentration-time curve from time zero to 6 hours after morning dosing, calculated according to the mixed log linear trapezoidal rule (ie, linear up/log down). Units: h*ng/mL.

C_{max} Maximum plasma concentration observed after morning dose, obtained directly from the observed concentration versus time data. Units: ng/mL.

t_{max} Time to reach maximum plasma concentration, obtained directly from the observed concentration versus time data. Units: h.

C_{trough} the concentration observed immediately before next dosing (Day 29 only). Units: ng/mL.

$R_{acc(AUC0-6h)}$ Accumulation ratio for AUC, calculated as ratio of AUC's: $AUC_{0-6h, Day29}/AUC_{0-6h, Day1}$.

$R_{acc(Cmax)}$ Accumulation ratio for C_{max}, calculated as ratio of C_{max}'s: $C_{max, Day29}/C_{max, Day1}$.

16.3.3 Pharmacodynamic Endpoint Analyses

Absolute B cell numbers and subtypes (mature-naïve, non-switched, immature/transitional, memory), are assessed at screening and Days 1, 29, and 85 (end of treatment), prior to the morning dose in the 12-week treatment period. In addition absolute B cell numbers and subtypes will be assessed in the optional OLE period at Month 3, Month 6, and at the End of Trial Visit. Immunoglobulin levels (IgA, IgM, and IgG and subclasses) will be assessed at screening, Day 57 and Day 85 (end of treatment) in the treatment period, and at Month 3 and Month 6 visits in the OLE period.

Absolute B cell, B cell subtypes, and immunoglobulin levels (observed, change from baseline, and percent change from baseline values) will be summarized using the relevant PD analysis set. Descriptive statistics for these variables (mean, SD, CV%, median, minimum/maximum, Q1 and Q3 percentile, and interquartile range) will be presented.

The mean (SD) absolute value for each treatment group will be plotted versus time point for B cell, B cell subsets (mature-naïve, non-switched, immature/transitional, memory), and IgM, IgA, IgG immunoglobulin levels and IgG subclasses IgG1-IgG4. Similar plots of the mean (SD) percent change from baseline for each treatment group versus time point will be generated for B cell and B cell subtypes (mature-naïve, non-switched, immature/transitional, memory), NK cell, T cell, T cell subtypes (CD4 and CD8) and IgM, IgA, and IgG (and IgG subclasses) immunoglobulin levels.

16.3.4 Exploratory Endpoints

CCI [redacted]

CCI

CCI

17 Safety Evaluation

The subsections in this section include specifications for summarizing safety endpoints that are common across clinical trials such as AEs, laboratory tests and vital signs.

For all safety endpoints, data selection for primary and final analyses will be handled as indicated in the Table 15:

Table 15 : Data handling for safety analysis

Analysis	Population	Period covered	Treatment groups	Data to be analyzed
Primary analysis	Safety Analysis Set	12-Week treatment period until start of OLE or discontinuation of treatment prior to OLE, whichever is applicable	Placebo or M2951 50 mg BID	All Data collected during the 12-Week treatment period up to start of OLE if applicable
Final analysis	OLE Safety Analysis set	OLE period only	Placebo/M2951 50 mg BID only	Data from OLE period only, or posterior to OLE (Safety follow-up / End of Study)
		Both 12-Week treatment period and OLE period	M2951 50 mg BID/M2951 50 mg BID only	Data from both periods (12-Week treatment and OLE period), ie all data collected throughout the study

17.1 Adverse Events

All analyses described in Section 17.1 will be based on treatment-emergent adverse events (TEAEs) if not otherwise specified.

TEAEs will be defined according to Table 16.

Table 16: Definition of TEAE

Primary analysis	TEAEs will be defined as AEs starting on or after first treatment administration of any IMP (placebo or M2951) until first dose of M2951 during OLE (or until the end of study if the patient declined to participate in OLE period). Only AEs from 12-Week treatment period will be considered.
Final analysis	• For subjects who switch from placebo to M2951 50 mg BID, TEAEs will be defined as AEs starting on or after first treatment administration of M2951 50 mg BID during OLE until the end of study. Only AEs from OLE period will be considered. • For patients who continue with M2951 50 mg BID during OLE, TEAEs will be defined as AEs starting on or after first treatment administration of M2951 during 12-Week double blind treatment period. AEs from both periods (12-Week treatment and OLE) will be considered.

Incomplete AE-related dates will be handled as follows:

- In case the onset date is missing completely or missing partially but the onset month and year, or the onset year are equal to the start of IMP then the onset date will be replaced by the minimum of start of IMP and AE resolution date.
- In all other cases the missing onset day or missing onset month will be replaced by 1.
- Incomplete stop dates will be replaced by the last day of the month (if day is missing only), if not resulting in a date later than the date of subject's death. In the latter case, the date of death will be used to impute the incomplete stop date.
- Further information collected after the cut-off for an analysis (such as a fatal outcome) may be extracted from the Safety data base and presented separately in the CSR.

17.1.1 All Adverse Events

AEs will be coded according to the latest MedDRA version available at the time of analysis. The severity of AEs will be graded using National Cancer Institute - Common Terminology Criteria for Adverse Events (NCI-CTCAE version 4.03) toxicity grades. AEs with missing classification concerning IMP relationship will be considered related to the IMP.

All AEs recorded during the course of the trial (ie assessed from signature of informed consent until the end of the Follow-up/End of Trial visit) will be coded according to the MedDRA and assigned to a SOC and PT.

The number and percentage of subjects experiencing at least one TEAE will be summarized according to MedDRA SOCs and PTs by treatment group, relationship to IMP, and severity. If a subject experiences more than one occurrence of the same TEAE during the trial, the subject will only be counted once for that treatment (the worst severity and the worst relationship to trial treatment will be tabulated).

In case a subject had events with missing and non-missing grades, the maximum of the non-missing grades will be displayed.

A TEAE summary table will include a row for the overall frequency of TEAEs of the following types:

- Any TEAE
- IMP related TEAE
- Serious TEAE
- IMP related serious TEAE
- TEAE with severe intensity (NCI-CTCAE grade ≥ 3, NCI-CTCAE grade ≥ 4)
- IMP related TEAE with severe intensity (NCI-CTCAE grade ≥ 3, NCI-CTCAE grade ≥ 4)
- TEAE leading to death
- IMP related TEAE leading to death

A separate TEAE summary table will present the following:

- TEAE leading to interruption of IMP
- IMP related TEAE leading to interruption of IMP
- TEAE leading to withdrawal of IMP
- IMP related TEAE leading to withdrawal of IMP
- TEAE leading to dose reduction
- IMP related TEAE leading to dose reduction
- TEAE leading to study termination
- IMP related TEAE leading to study termination

The TEAE tables to be prepared are listed below.

	Overall frequency	By primary SOC and PT	By PT only
TEAE overview summary	X	*NA*	*NA*
TEAE leading to discontinuation of IMP/study/dose reduction of IMP overview summary	X	*NA*	*NA*
TEAE by SOC and PT	X	X	
TEAE by PT	X		X
IMP related TEAE by SOC and PT	X	X	
Serious TEAE by SOC and PT	X	X	
IMP related serious TEAE by SOC and PT	X	X	
Non-serious TEAE by SOC and PT*		X	
TEAE by worst grade, SOC and PT		X	
IMP related TEAE by worst grade, SOC and PT		X	
TEAE leading to death by SOC and PT	X	X	
IMP related TEAE leading to death by SOC and PT	X	X	
TEAE leading to withdrawal of IMP by SOC and PT	X	X	
TEAE leading to study termination	X	X	

*A table with all AEs will be first provided and then only AEs exceeding a frequency of 5% in at least one of the treatment groups (> 5%), by SOC and PT will be provided.

Group/SOC terms will be sorted alphabetically. PTs within each group/SOC will be sorted by active dose descending frequency, and alphabetically if multiple PTs have the same frequency.

Pre-treatment and TEAEs will be listed separately by treatment group and subject. Specific data listings of pre-treatment AEs and TEAEs will be performed for subjects with a gap > 4 weeks between 12-Week treatment period and OLE period.

A listing of TEAEs leading to withdrawal of IMP, a listing of TEAEs leading to study termination, if any, will be provided.

Exposure adjusted incidence rates are calculated as number of subjects with AE divided by the sum of the individual times of all subjects in the safety population from start of treatment to first onset of AE. The incidence rate multiplied with 1000 would give the number of AEs expected in 1000 subjects within 1 time unit (for example 1 year). The exposure adjusted incidence rates of TEAEs will be presented by SOC and PT.

17.2 Deaths, Other Serious Adverse Events, and Other Significant Adverse Events

17.2.1 Deaths

A summary of deaths will be provided including (clinicaltrials.gov requirement):

- Number and percentage of (all) deaths
- Number and percentage of the primary cause of death (categories: disease progression, adverse event, unknown, other)

The tabulation of TEAEs leading to death is described in Section 17.1.1. A listing of deaths, if any, will be provided.

In case there is no death in the trial, only the summary of death required by clinicaltrials.gov will be performed, neither tabulation of TEAE leading to death will be edited, nor the listing of death.

17.2.2 Serious Adverse Events

The tabulation of serious TEAEs is described in Section 17.1.1. A subject listing of serious TEAEs will be provided.

17.3 Clinical Laboratory Evaluation

The following laboratory parameters will be measured during the trial as part of the safety evaluation:

- Hematology
- Biochemistry
- Urinalysis
- Coagulation

The clinical laboratory safety tests to be measured in this trial are provided in the protocol (refer to Section 7.4.3 Table 4 of the Clinical Trial Protocol [CTP]).

Continuous protocol-specified clinical laboratory findings (hematology, biochemistry, urinalysis, coagulation) will be summarized by treatment using descriptive statistics (see Section 11). The statistics will be presented for baseline, each time point during the study, and CFB to each time point, when applicable.

Laboratory results will be classified according to NCI-CTCAE Version 4.03 as provided by the central laboratory. In case a laboratory parameter has bi-directional toxicities (eg Potassium) both directions will be presented for the given parameter (ie Potassium Low and Potassium High). On-treatment values are results of assessments done from the first IMP administration on Day 1. Laboratory results containing a modifier such as “<” or “>=” will be handled case by case for summary statistics and will be reported as collected in the database in subject data listings.

A shift table of baseline versus post-baseline based on the worst NCI-CTCAE grade will be presented by treatment group for hematology and biochemistry.

Additional laboratory results that are neither classified in protocol nor part of NCI-CTCAE will be presented according to the categories based on normal ranges: below normal limits (Low), within normal limits (Normal), and above normal limits (High). Classification into categories will be presented in subject data listings only.

Boxplots of the laboratory values (by treatment arm) by time point will be provided for:

- Hemoglobin
- Absolute reticulocyte count
- White blood cell count
- Absolute neutrophil count
- Absolute lymphocyte count
- Platelet count
- Alanine aminotransferase (ALT)
- Aspartate aminotransferase (AST)
- Gamma-glutamyl transferase (GGT)
- Total bilirubin
- Amylase
- Lipase
- Creatinine
- Blood urea nitrogen (BUN)

Boxplots for laboratory parameters where toxicity grades are defined based on the ratio of the parameter values and the ULN will not be displayed using the unit of measurement but instead using the ratio of the measured value over ULN. This comprises the following parameters:

- ALT
- AST
- GGT
- Total bilirubin
- Creatinine
- Amylase
- Lipase

Listings of individual data with a flag for abnormal values will be provided. Specific data listings of laboratory data will be performed for subjects with a gap > 4 weeks between 12-Week treatment period and OLE period.

Laboratory values that are outside the normal range will also be flagged in the data listings, along with corresponding normal ranges. Any out-of-range hematology, biochemistry, urinalysis and coagulation values that are identified as being clinically significant (ie, values classified as low or high based on normal range and/or value with grade ≥ 1) will be shown in a data listing.

17.4 Vital Signs

Vital signs (body temperature (°C), SBP (mmHg), DBP (mmHg), respiratory rate (breaths/min) and pulse rate (beats/min)) will be summarized by treatment group using descriptive statistics (see Section 11) for baseline, each applicable time point and CFB to each time point.

Body temperature, SBP, DBP, respiratory rate and pulse rate will be analyzed with shift tables of maximum CFB using the categories defined in Table 17:

Table 17: Vital signs categories

Parameter	Unit	Shift	Baseline categories	Post-baseline categories (absolute change)
Temperature	°C	Increase	<37 / ≥37 - <38 / ≥38 - <39 /≥39 - <40 / ≥ 40	≤0* / >0 - <1 / ≥1 - <2 /≥2 - <3 / ≥3
Pulse rate	bpm	Increase and decrease	<100 / ≥100	≤0* / >0 - ≤20 / >20 - ≤40 / >40
SBP	mmHg	Increase and decrease	<140 / ≥140	≤0* / >0 - ≤20 / >20 - ≤40 / >40
DBP	mmHg	Increase and decrease	<90 / ≥90	≤0* / >0 - ≤ 20 / >20 - ≤40 / >40
Respiratory rate	breaths/min	Increase and decrease	<20 / ≥20	≤0* / >0 - ≤5 / >5 - ≤10 / >10

* This category will include the subjects with no changes or decrease/increase in the increase/decrease part of the table respectively.

A listing of maximum CFB and a listing of all vital signs data will be provided.

17.5 12-Lead Electrocardiogram (ECG)

The 12-lead ECG data will be listed and summarized for observed values and CFB values by treatment group using descriptive statistics:

- Ventricular rate (beats/min)

- Pulse rate interval (msec)
- QRS (msec)
- QT (msec)
- Fridericia corrected QT (QTcF) (msec).

QTcF values will be categorized according to their calculated values into the categories

- ≤ 430 msec,
- > 430 – 450 msec,
- > 450 – 480 msec,
- > 480 – 500 msec,
- > 500 msec

and categorized according to their CFB into the categories

- ≤ 30 msec,
- > 30 – 60 msec,
- > 60 msec.

The ECG outlier values will also be tabulated. The categories described above will be summarized by treatment group in frequency tables using number and percentage of subjects.

A listing of ECG quantitative values, morphological and rhythm results will be produced.

A shift table of rhythm results, from baseline to end of treatment, of the number and percentage of subjects for each category (Sinus rhythm, Atrial fibrillation, Other, Missing and Total) will be provided. This table will only be performed for the primary analysis.

A shift table of morphological assessments, from baseline to end of observation, of the number and percentage of subjects for each interpretation category (Normal, Abnormal, Missing, and Total) will also be provided, the end of observation referring to the last available observation of the 12-week treatment period for the primary analysis and the last available observation of the study for the final analysis.

17.6 Physical Examination

No summary table will be provided since physical examination findings during screening will be recorded as medical history events and findings during the trial as AEs.

17.7 Pregnancy test

Results of pregnancy test (serum and urine beta human chorionic gonadotropin for women only) will be listed.

17.8 B cell (CD19+) count

B cell (CD19+) count data will be listed by treatment group and time point (where applicable).

17.9 Urine Protein/Creatinine Ratio

Urine Protein/Creatinine Ratio data will be listed by treatment group and time point (where applicable).

17.10 Urinalysis Microscopic Evaluation

Urinalysis Microscopic Evaluation data will be listed by treatment group and time point (where applicable).

17.11 Chest X-ray evaluations

CXR evaluations will be listed and tabulated by treatment group using the number and percentage of subjects for each interpretation category (Normal, Abnormal Not Clinically Significant, Clinically Significant, and Abnormal Overall).

17.12 Other safety evaluations

Other characteristics like viral serology, Quantiferon TB test, thyroid-stimulating hormone and prior surgeries will be listed only.

18 References

(1) Cohen SB, Emery P, Greenwald MW, et al. Rituximab for rheumatoid arthritis refractory to anti-tumor necrosis factor therapy. Arthritis Rheum. 2006; 54(9):2793-806.

(2) Kremer JM, Cohen S, Wilkinson BE, et al. A phase IIb dose-ranging study of the oral JAK inhibitor tofacitinib (CP-690,550) versus placebo in combination with background methotrexate in patients with active rheumatoid arthritis and an inadequate response to methotrexate alone. Arthritis Rheum. 2012; 64(4):970-81.

(3) Miettinen OS, Nurminen M. Comparative analysis of two rates. Stat Med. 1985; 4(2):213-26.

(4) US Department of Health and Human Services. Common Terminology Criteria for Adverse Events Version 4.0. 2009 (v4.03: June 14, 2010).

(5) Anderson J, Caplan L, Yazdany J, et al. Rheumatoid Arthritis Disease Activity Measures: American College of Rheumatology Recommendations for Use in Clinical Practice. Arthritis Care & Research. 2012; 64(5):640-47.

(6) Stanford Patient Education Research Center Website [internet]. 2016. [cited 2016 April 14]. Available from: http://patienteducation.stanford.edu/research/haq20.html

19 Appendices

19.1 Appendix 1: HAQ-DI scoring

The HAQ-DI is a subject-reported questionnaire that consists of 24 questions referring to 8 domains: dressing and grooming, arising, eating, walking, hygiene, reach, reach, grip, and activities.

The questionnaire scores the subject's self-perception (20 questions): 0=without any difficulty, 1=with some difficulty, 3=unable to do, and reports use of special aids or devices and/or the need for assistance from another person:

Domains	Questions
Dressing and grooming	1. Dress yourself, including tying shoelaces and doing buttons? 2. Shampoo your hair?
Arising	3. Stand up from a straight chair? 4. Get in and out of bed?
Eating	5. Cut your meat? 6. Lift a full cup or glass to your mouth? 7. Open a new milk carton?
Walking	8. Walk outdoors on flat ground? 9. Climb up five steps?
	Check any aids or devices that you usually use for any of the above activities (questions 1 to 9): Devices used for dressing (button hook, zipper pull, etc.) / Special or built up chair / Built up or special utensils / Cane / Walker / Crutches / Wheelchair / Other
	Check any categories for which you usually need help from another person (questions 1 to 9): Dressing and grooming / Arising / Eating / Walking
Hygiene	10. Wash and dry your body? 11. Take a tub bath? 12. Get on and off the toilet?
Reach	13. Reach and get down a 5 pound object (such as a bag of sugar) from just above your head? 14. Bend down to pick up clothing from the floor?
Grip	15. Open car doors? 16. Open jars which have been previously opened? 17. Turn faucets on and off?
Activities	18. Run errands and shop? 19. Get in and out of a car? 20. Do chores such as vacuuming or yard work?
	Check any aids or devices that you usually use for any of the above activities(questions 10 to 20): Raised toilet seat / bathtub seat / bathtub bar / long-handled appliances in bathroom / long-handled appliances for reach / jar opener (for jar previously opened) / Other
	Check any categories for which you usually need help from another person (questions 10 to 20): Hygiene / Reach / Gripping an opening things / Errands and Chores

The highest score for any component question of each of the eight domains determines the score for that domain. Each domain must have at least one question answered. Otherwise, the domain score is set to missing.

The non-missing domain scores are adjusted, based upon the patient's use of any aid, device or assistance (multiple aids, devices, or assistance could be checked). The relationship between aids, devices help from another person and the disability domain is shown below:

Domains	Aids or devices	Help from another person
Dressing and grooming	Devices used for dressing / Other	Dressing and grooming
Arising	Special or built up chair / Other	Arising
Eating	Built up or special utensils / Other	Eating
Walking	Cane / Walker / Crutches / Wheelchair / Other	Walking
Hygiene	Raised toilet seat / bathtub seat / bathtub bar / long-handled appliances in bathroom / Other	Hygiene
Reach	long-handled appliances for reach / Other	Reach
Grip	jar opener / Other	Gripping an opening things
Activities	Other	Errands and Chores

If either aid or device and/or help from another person are checked for a domain or 'other' aid/device/assistance was needed, then the score for that domain is raised from 0 or 1 to 2 and unchanged if already scored at 2 or 3. If no aid, device and assistance were needed, the score for that domain will remain as raw score.

The disability index is the mean of the eight domain scores. If at least two of the domains are missing, the disability index cannot be obtained. If fewer than 2 of the domain scores is missing, divide the sum of the domains by the number of answered domains. The disability scores between 0 and 3, with higher score indicating greater disability.

19.2 Appendix 2: DAS28-hsCRP formula

As per CTP Section 7.3.1.3, the DAS28-hsCRP will be computed according to below formula:

DAS28-hsCRP=0.56*√(TJC28)+0.28*√(SJC28) + 0.014*GH +0.36*ln(hsCRP+1) +0.96

Where,

- TJC28=28 Tender Joint Counts (14 joints on each side of the subject's body: 2 shoulders, 2 elbows, 2 wrists, 10 metacarpophalangeal joints, 2 interphalangeal joints of the thumb, 8 proximal interphalangeal joints, and 2 knees)
- SJC28=28 Swollen Joint Counts (same joints used as for TJC28)
- ln(hsCRP)=natural logarithm of hsCRP
- GH=the General Health component of the DAS (ie Subject's Global Assessment of Disease Activity).

19.3 Appendix 3: Important Protocol Deviations identified by programming and Clinically Important Protocol Deviations

Protocol Deviation number	Protocol Deviation	Important	Clinically Important	Type criteria	Inclusion/Exclusion criteria number	Check
PDEV01	Age < 18 or > 75 at trial entry	Yes	Yes	Eligibility and Entry Criteria	Inclusion criteria #1	If calculated age at informed consent is <18 or >75 years, or if corresponding inclusion criterion is ticked No.
PDEV02	No diagnosis of RA > 6 months' duration at Screening	Yes	Yes	Eligibility and Entry Criteria	Inclusion criteria #2	If no diagnosis of RA or time between diagnosis of RA and Screening Visit is less than 6 months or corresponding inclusion criterion is ticked No.
PDEV03	Negative RF and negative anti-CCP	Yes	Yes	Eligibility and Entry Criteria	Inclusion criteria #3	If RF (IgG-RF, IgM-RF, IgA-RF) is classified as negative (≤ULN) and Anti-CCP classified as negative (≤ULN) at the Screening visit or corresponding inclusion criterion is ticked No.
PDEV04	Absence of persistently active disease	Yes	Yes	Eligibility and Entry Criteria	Inclusion criteria #4	If subject has less than 6 swollen joints (of 66 counted) or less than 6 tender joints (of 68 counted) at Screening Visit or corresponding inclusion criterion is ticked No.

Protocol Deviation number	Protocol Deviation	Important	Clinically Important	Type criteria	Inclusion/Exclusion criteria number	Check
PDEV05	hsCRP < 3.6 mg/L	Yes	Yes	Eligibility and Entry Criteria	Inclusion criteria #5	If hsCRP < 3.6 mg/L at Screening Visit or corresponding inclusion criterion is ticked No.
PDEV06	Treatment for less than 12 weeks with 10 to 25 mg/week MTX	Yes	Yes	Eligibility and Entry Criteria	Inclusion criteria #6	If no relevant previous medication is reported in the eCRF as MTX, or a relevant medication reported with a duration less than 12 weeks or a dose not included between 10 and 25 mg, or corresponding inclusion criterion is ticked No.
PDEV07	Women of childbearing potential do not agree to use highly effective methods of contraception	Yes	Yes	Eligibility and Entry Criteria	Inclusion criteria #7	If corresponding inclusion criterion is ticked No.
PDEV08	Males do not agree to use or to have their female partners use highly effective contraception	Yes	Yes	Eligibility and Entry Criteria	Inclusion criteria #8	If corresponding inclusion criterion is ticked No.

Protocol Deviation number	Protocol Deviation	Important	Clinically Important	Type criteria	Inclusion/Exclusion criteria number	Check
PDEV09	Women of childbearing potential do not have negative serum pregnancy test at screening or negative urine pregnancy test at Day 1	Yes	Yes	Eligibility and Entry Criteria	Inclusion criteria #9	If women have a positive serum pregnancy test at Screening visit or a positive urine pregnancy test at Day 1 before randomization and dosing, or corresponding inclusion criterion is ticked No.
PDEV10	History of vaccination not compliant	Yes	Yes	Eligibility and Entry Criteria	Inclusion criteria #10	If identified by medical review or corresponding inclusion criterion is ticked No.
PDEV11	Informed consent not signed	Yes	Yes	Eligibility and Entry Criteria	Inclusion criteria #11	If missing date of informed consent, or corresponding inclusion criterion is ticked No.
PDEV12	Use or change in dose of corticosteroids	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #1	If identified by medical review or corresponding exclusion criterion is ticked Yes.
PDEV13	Initiation or change in dose for nonsteroidal anti-inflammatory drugs	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #2	If identified by medical review or corresponding exclusion criterion is ticked Yes.
PDEV14	Treatment with tofacitinib, other BTK inhibitors, or a biologic DMARD, or other immunosuppressive drugs other than MTX	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #3	If identified by medical review or corresponding exclusion criterion is ticked Yes.

Protocol Deviation number	Protocol Deviation	Important	Clinically Important	Type criteria	Inclusion/Exclusion criteria number	Check
PDEV15	Treatment with anti-CD20 therapy (eg, rituximab)	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #4	If identified by medical review or corresponding exclusion criterion is ticked Yes.
PDEV16	Immunologic disorder other than RA	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #5	If identified by medical review or corresponding exclusion criterion is ticked Yes.
PDEV17	Vaccination with live or live attenuated virus vaccine	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #6	If identified by medical review or corresponding exclusion criterion is ticked Yes.
PDEV18	Severe drug allergy or history of anaphylaxis	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #7	If identified by medical review or corresponding exclusion criterion is ticked Yes.
PDEV19	Significant viral, bacterial or fungal infection	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #8	If identified by medical review or corresponding exclusion criterion is ticked Yes.
PDEV20	History of or positive HIV, hepatitis C, hepatitis B and/or IgM antibody	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #9	If identified by medical review or corresponding exclusion criterion is ticked Yes.
PDEV21	History of or current diagnosis of active TB	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #10	If identified by medical review or corresponding exclusion criterion is ticked Yes.
PDEV22	History of splenectomy or any major surgery	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #11	If identified by medical review or corresponding exclusion criterion is ticked Yes.

Protocol Deviation number	Protocol Deviation	Important	Clinically Important	Type criteria	Inclusion/Exclusion criteria number	Check
PDEV23	History of any myocardial infarction or cerebrovascular event	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #12	If identified by medical review or corresponding exclusion criterion is ticked Yes.
PDEV24	Anticoagulation or antiplatelet therapy other than daily aspirin for cardioprotection	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #13	If identified by medical review or corresponding exclusion criterion is ticked Yes.
PDEV25	Fish oil	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #14	If identified by medical review or corresponding exclusion criterion is ticked Yes.
PDEV26	History of cancer unless considered cured > 5 years	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #15	If identified by medical review or corresponding exclusion criterion is ticked Yes.
PDEV27	Breastfeeding/lactating or pregnant women	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #16	If identified by medical review or corresponding exclusion criterion is ticked Yes.
PDEV28	Clinically significant abnormality on ECG or an active infective process or any other clinically significant abnormality on CXR, per Investigator opinion	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #17	If identified by medical review or corresponding exclusion criterion is ticked Yes.

Protocol Deviation number	Protocol Deviation	Important	Clinically Important	Type criteria	Inclusion/Exclusion criteria number	Check
PDEV29	Estimated glomerular filtration rate by the 4-variable Modification of Diet in Renal Disease of < 45 mL/min/1.73 m^2	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #18	If identified by medical review or corresponding exclusion criterion is ticked Yes.
PDEV30	TSH < 0.01 or ≥ 7.1 mIU/L at Screening	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #19	If identified by medical review or corresponding exclusion criterion is ticked Yes.
PDEV31	Clinically significant laboratory abnormality	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #20	If identified by medical review or corresponding exclusion criterion is ticked Yes.
PDEV32	B cell (CD19) count < 50% of the lower limit of normal	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #21	If identified by medical review or corresponding exclusion criterion is ticked Yes.
PDEV33	Significant cytopenia	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #22	If identified by medical review or corresponding exclusion criterion is ticked Yes.
PDEV34	Participation in any investigational drug trial	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #23	If identified by medical review or corresponding exclusion criterion is ticked Yes.

Protocol Deviation number	Protocol Deviation	Important	Clinically Important	Type criteria	Inclusion/Exclusion criteria number	Check
PDEV35	Subjects receiving medications, herbal supplements, or food known to be moderate or strong inhibitors of CYP3A or drugs mainly metabolized by CYP3A (see Section 6.5.2 of CTP)	Yes	Yes	Eligibility and Entry Criteria	Exclusion criteria #24	If identified by medical review or corresponding exclusion criterion is ticked Yes.
PDEV36	Concurrent treatment with a non-permitted drug or procedure (refer to section 6.5.2 of CTP).	Yes	case by case basis			If identified by medical review.
PDEV37	Incorrect treatment group IP dispensed to subject	Yes	case by case basis	IP Compliance		If randomized treatment is different from actual treatment
PDEV38	Non-compliant with IP usage	Yes	case by case basis	IP Compliance		Check if <80% of >120%
PDEV39	IP not taken in compliance with protocol eg, at the appropriate times, during/timed with meals/snacks	case by case basis	case by case basis	IP Compliance		From monitoring log/Medical Review
PDEV40	Subject did not sign ICF and was enrolled	Yes	Yes	Informed Consent Criteria		ICF date is greater than randomization date

Protocol Deviation number	Protocol Deviation	Important	Clinically Important	Type criteria	Inclusion/Exclusion criteria number	Check
PDEV41	Subject used prohibited medication(s)	case by case basis	case by case basis	Other Criteria		Medical review
PDEV42	Lab safety and/or PK samples not drawn according to protocol requirements	case by case basis	case by case basis	Laboratory Assessment Criteria		Monitoring deviation log, Medical review