Phase III Study of ASP015K

−A Randomized, Double-Blind, Placebo-Controlled Confirmatory Study of Efficacy and Safety of ASP015K in Patients with Rheumatoid Arthritis (RA) Who Had an Inadequate Response to MTX−

ISN/Protocol 015K-CL-RAJ4

ClinicalTrials.gov Identifier: NCT02305849

Date of Statistical Analysis Plan: Version 1.0, dated 18 January 2018

Sponsor: Astellas Pharma Inc. (API)

2-5-1 Nihonbashi-Honcho, Chuo-ku, Tokyo

STATISTICAL ANALYSIS PLAN

Final Version, dated 18-Jan 2018

Phase III Study of ASP015K

A Randomized, Double-Blind, Placebo-Controlled Confirmatory Study of Efficacy and Safety of ASP015K in Patients with Rheumatoid Arthritis (RA) Who Had an Inadequate Response to MTX

ISN: 015K-CL-RAJ4

Astellas Pharma Inc. (API)

Table of Contents

I. LIST OF ABBREVIATIONS AND KEY TERMS

List of Abbreviations

Abbreviations	Description of abbreviations
ACR	American College of Rheumatology
AE	Adverse Event
ALP	Alkaline Phosphatase
ALT	Alanine Transaminase
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
AST	Aspartate Transaminase
ASP015K	Astellas Pharmaceuticals compound 015K
BDRM	Blinded Data Review Meeting
BMI	Body Mass Index
BUN	Blood urea nitrogen
anti-CCP antibody	anti-cyclic citrullinated peptide antibody
CDAI	Clinical Disease Activity Index
CI	Confidence Intervals
CK/CPK	Creatine kinase /creatine phosphokinase
CK-MB	Creatine kinase MB isozyme
CRF	Case Report Form
CRO	Contract research organization
CRP	C-reactive protein
CSR	Clinical Study Report
DAS	Disease activity score
DBP	Diastolic Blood Pressure
DILI	Drug induced liver injury
DIP	Distal interpharangeal joint
DMARD	Disease-modifying antirheumatic drug
DSMB	Data and Safety Monitoring Board
ECG	Electrocardiogram
eCRF	electronic Case Report Form
EOS	End of Study
ESR	Erythrocyte sedimentation rate
ET	Early termination
EU	European Union
EULAR	European League Against Rheumatism
FAS	Full Analysis Set
FDA	Food and Drug Administration
GCP	Good Clinical Practice
GD	Global Development
GFR	Glomerular filtration rate
GMP	Good Manufacturing Practice
γ-GTP	γ-glutamyl transpeptidase (GGT)
H	High
HAQ-DI	Health Assessment Questionnaire – Disability Index
HBc antibody	Hepatitis B core antibody
HBs antigen/antibody	Hepatitis B surface antigen/antibody
hCG	Human chorionic gonadotropin
HBV	Hepatitis B virus

Abbreviations	Description of abbreviations
HCV	Hepatitis C virus
HDL	High-density lipoprotein
hERG	Ether-a-go-go related gene
HIV	Human immunodeficiency virus
HSA	Human serum albumin
ICH	International Conference on Harmonization
INR	International normalized ratio
IP	Interphalangeal joint
IRB	Institutional review board
ISN	International study number
IV	Intravenous
JAK	Janus kinase
L	Low
LLN	Lower Limit of Normal
LLOQ	Lower limit of quantification
LOCF	Last observation carried forward
MCP	Metacarpophalangeal joint
MCS	Mental component score of the SF-36v2®
MDRD	Modification of diet in renal disease
MedDRA	Medical Dictionary for Regulatory Activities
MMP-3	Matrix metalloproteinase 3
MTP	Metatarsophalangeal joint
mTSS	Modified Total Sharp Score
MTX	Methotrexate
N	Normal
NCI-CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
NDA	New drug application
NK	Natural killer
NRI	Non-responder imputation
NSAID	Non-steroidal anti-inflammatory drugs
PCS	Phisical component score of the SF-36v2®
PD	Pharmacodynamic
PD1-x	Protocol Deviation 1-x
PDAS	Pharmacodynamic Analysis Set
PGA	Physician's Global assessment of arthritis
PGx	Pharmacogenomics
PIP	Proximal interpharangeal joint
PK	Pharmacokinetic
PKAS	Pharmacokinetics Analysis Set
PPS	Per-Protocol Set
PT	Preferred Term
PY	Patient-years
QD	Once a day
RA	Rheumatoid arthritis
RCS	Role/social component score of the SF-36v2®
RF	Rheumatoid factor
SAE	Serious adverse event
SAF	Safety Analysis Set
SAP	Statistical Analysis Plan

Abbreviations	Description of abbreviations
SAS	Statistical Analysis Software
SBP	Systolic Blood Pressure
SDAI	Simplified disease activity index
SF-36v2®	Short form health survey – 36 questions, version 2
SFL	Screening failure log
SGA	Subject's global assessment of arthritis
SGAP	Subject's global assessment of arthritis pain
SI	International System of Units
SJC	Swollen joint count
SOC	System Organ Class
SOP	Standard operating procedure
TBL	Total bilirubin
TEAE	Treatment Emergent Adverse Event
TJC	Tender joint count
TLF	Tables, Listings and Figures
TNF	Tumor necrosis factor
ULN	Upper Limit of Normal
VAS	Visual analog scale
WBC	White blood cell
WHO-DD	World Health Organization Drug Dictionary
WPAI	Work Productivity and Activity Impairment Questionnaire

List of Key Terms

Terms	Definition of terms
Baseline	The last non-missing measurement prior to initial dosing of study drug(Screening or Day1)
Study period	Period of time from the time point when written informed consent is obtained from the subject and the end of the final assessment/observation specified in the protocol.
Treatment period	Period of time from the first study drug administration for the treatment period to the end of Week 52 visit or the end of the withdrawal visit in case of withdrawal.
Follow-up period	The follow-up period will be 4 weeks (on a per-protocol basis) starting after the early-termination visit for withdrawn subjects and after the Week 52 visit for subjects who completed the study but are not willing to enroll into the extension study. For subjects who completed the study and wish to enroll into the extension study, the follow-up period will last for a maximum of 4 weeks starting after the Week 52 visit and ending at the initiation of the extension study treatment.
Subject	An individual who participates in a clinical trial as a recipient of the study drug.
Week 12/ET	Week 12 or early termination before Week 12
Week 28/ET	Week 28 or early termination before Week 28
EOT	Week 52 or early termination before Week 52

Discontinuation	When a subject who was enrolled in the study discontinues his/her participation in the study before the completion of all the protocol-specified items of the study.
Variable	Any quantity that varies; any attribute, phenomenon or event that can have different qualitative or quantitative values.

1 INTRODUCTION

This Statistical Analysis Plan (SAP) contains a more technical and detailed elaboration of the principal features of the analysis described in the protocol, and includes detailed procedures for executing the statistical analysis of the primary and secondary endpoints and other data.

The SAP is finalized and signed prior to unblinding.

This statistical analysis is coordinated by the responsible biostatistician of GD, API. Any changes from the analyses planned in the SAP will be justified in the Clinical Study Report (CSR).

2 FLOW CHART AND VISIT SCHEDULE

FLOW CHART

Obtain informed consent	→ Screening period (The screening test period must be within 28 days prior to the baseline visit.) Confirmation of inclusion/exclusion criteria Preliminary registration	→ Baseline Confirmation of inclusion/exclusion criteria Enrollment & randomization	→ Treatment period: maximum of 52 weeks, ranging from the start of the study treatment to the completion of tests/assessments schedule at the end of treatment (Week 52 or early termination) See figure below.	→ Entry into the extension study Subjects not enrolling into the extension study will undergo follow-up observation 4 weeks after the end of the study treatment.

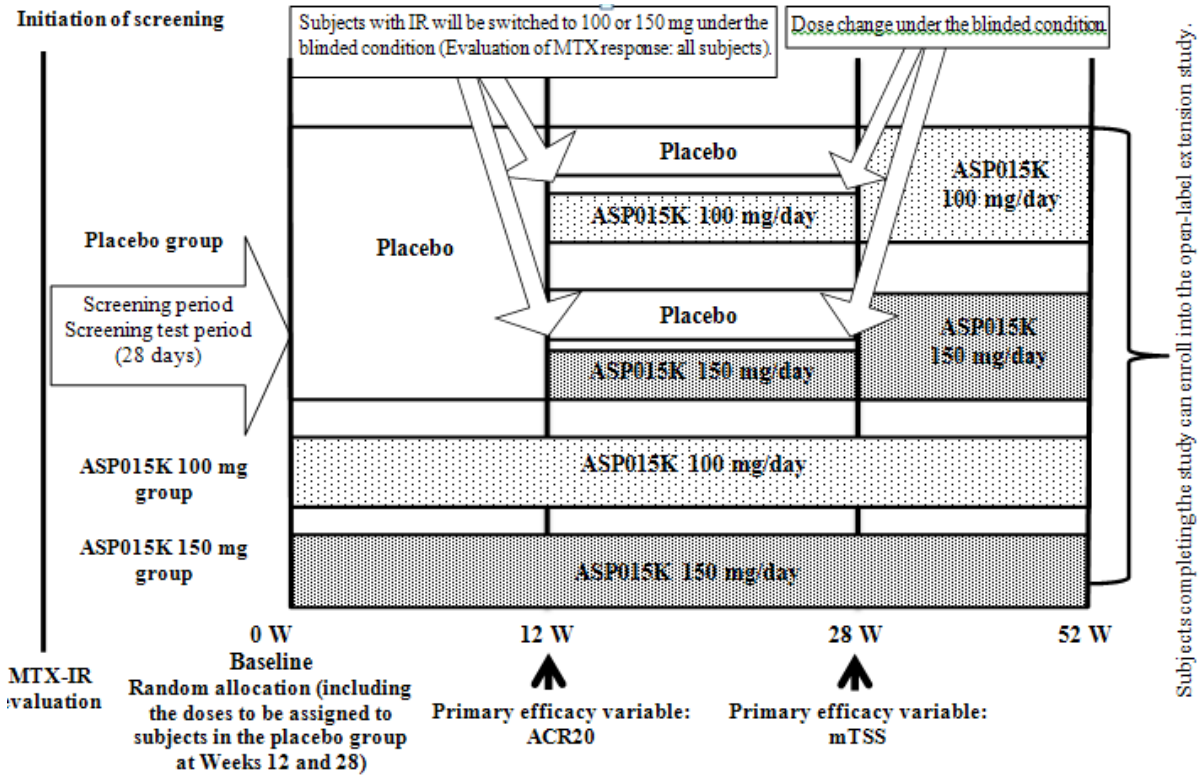

Table 1 Schedule of Assessments

Visit	Initiation of Screening[1)]	Screening Test Period	0 W	4 W	8 W	12 W	16 W	20 W	24 W	28 W	32 W	36 W	40 W	44 W	48 W	52 W	Early Termination	Follow-up[2)]	Unscheduled Visit
Visit day	**−1 at latest**	**−28 to −1**	**1**	**29**	**57**	**85**	**113**	**141**	**169**	**197**	**225**	**253**	**281**	**309**	**337**	**365**	**Within 2 days of withdrawal**	**28 days after Week 52 or early termination visit**	**-**
Allowable range from specified date				±7	±7	±7	±7	±7	±7	±7	±7	±7	±7	±7	±7	±7	+2	±7	-
Informed consent	○																		
Inclusion/exclusion criteria	○	○	○																
Subject registration	○		○			○				○						○	○	○	
Inadequate response to MTX	○		○																
Demographics	○	○	○																
Medical history	○	○	○																
Height		○																	
TB		○																	
Anti-CCP antibody, RF		○																	
Safety																			
Physical examination		○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	▲
Body weight		○				○				○			○			○	○	○	▲
Vital signs		○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	▲
Hepatitis examination		○																	▲
Hematology[3)]		○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	▲
Serum biochemistry[3)]		○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	▲
Urinalysis[3)]		○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	▲
Fasting lipid profile test			○			○		○		○			○			○	○	○	▲
Chest radiography		○								○						○	▲		▲
12-lead ECG		○								○						○	○		▲
Central ECG[6)]			○	○															

Visit	Initiation of Screening[1)]	Screening Test Period	0 W	4 W	8 W	12 W	16 W	20 W	24 W	28 W	32 W	36 W	40 W	44 W	48 W	52 W	Early Termination	Follow-up[2)]	Unscheduled Visit
Visit day	**−1 at latest**	**−28 to −1**	**1**	**29**	**57**	**85**	**113**	**141**	**169**	**197**	**225**	**253**	**281**	**309**	**337**	**365**	**Within 2 days of withdrawal**	**28 days after Week 52 or early termination visit**	**-**
Allowable range from specified date				±7	±7	±7	±7	±7	±7	±7	±7	±7	±7	±7	±7	±7	+2	±7	-
Pregnancy test[5)]		○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	▲
Adverse event			○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	▲
Efficacy																			
Hand and foot radiography		○	○			○[4)]				○						○	○		
CRP and ESR		○ (CRP only)	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	▲
TJC/SJC (68/66 joints)		○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	▲
PGA (VAS), SGA (VAS), and subject's assessment of pain (VAS)			○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	▲
SF-36 v2®			○	○	○	○				○						○	○	○	▲
WPAI			○	○	○	○				○						○	○	○	▲
HAQ-DI			○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	▲
PK/PD																			▲
Blood sampling for trough concentration			○	○	○	○		○		○			○			○	○		
Blood sampling for post-dose drug concentration				○															
Blood sampling for PD (biomarkers)			○	○	○	○				○						○	○	○	▲
Blood sampling for PD (lymphocyte subset)			○	○	○	○				○						○	○	○	▲
Informed consent for PGx research (relevant sites only)			○																

Visit	**Initiation of Screening[1)]**	**Screening Test Period**	**0 W**	**4 W**	**8 W**	**12 W**	**16 W**	**20 W**	**24 W**	**28 W**	**32 W**	**36 W**	**40 W**	**44 W**	**48 W**	**52 W**	**Early Termination**	**Follow-up[2)]**	**Unscheduled Visit**
Visit day	**−1 at latest**	**−28 to −1**	**1**	**29**	**57**	**85**	**113**	**141**	**169**	**197**	**225**	**253**	**281**	**309**	**337**	**365**	**Within 2 days of withdrawal**	**28 days after Week 52 or early termination visit**	**-**
Allowable range from specified date				±7	±7	±7	±7	±7	±7	±7	±7	±7	±7	±7	±7	±7	+2	±7	-
Blood sampling for PGx (relevant sites only)			○																
Study drug																			
Confirmation of remaining unused drug				○	○	○	○	○	○	○	○	○	○	○	○	○	○		
Prescription			○	○	○	○	○	○	○	○	○	○	○	○	○				

The symbol "○" designates mandatory items, whereas the symbol "▲" denotes the optional items to be investigated on the basis of the clinical judgment of the investigator/sub-investigator.

For details, see the respective sections of the protocol.

1) Subjects who have signed the informed consent will undergo evaluation of MTX response at the screening visit.

2) Follow-up will not be performed for subjects starting the extension study immediately after the Week 52 visit.

3) For details of the test parameters, refer to Section 5.4.4.

4) At Week 12, subjects with < 20% improvement from baseline in TJC and SJC will be required to undergo radiographic examination of their hands and feet.

5) For pregnancy tests, serum samples will be used at the screening visit, and urine samples will be used at and after the baseline visit.

6) Central ECG should be performed on the same day as blood sampling for post-dose drug concentration at Week 4 or Week 8 before and after study drug administration.

3 STUDY OBJECTIVE(S) AND DESIGN

3.1 Study Objective(s)

The main objective of this study is to evaluate the efficacy of ASP015K (100 and 150 mg/day) versus placebo administered in combination with MTX in RA patients who had an inadequate response to MTX. The primary efficacy variables are the ACR20 response rate at Week 12 and suppression of joint destruction (change from baseline in mTSS) at Week 28. This study will also investigate pharmacokinetic, pharmacodynamic, and safety profiles. Furthermore, this study will examine the long-term efficacy and safety (52 weeks).

3.2 Study Design

This study is a multi-center, randomized, placebo-controlled, double-blind, parallel-group, confirmatory study to evaluate the efficacy and safety of ASP015K (100 and 150 mg/day) administered in combination with MTX in patients with RA who had an inadequate response to MTX. The geographical scope of this study will be confined to Japan. MTX response is evaluated at the initiation of screening and then after the screening period, subjects will be randomly assigned at a 1:1:1 ratio to receive ASP015K 100 mg, ASP015K 150 mg, or placebo QD in combination with MTX. Subjects will orally receive the study drug after breakfast for 52 weeks. The target number of subjects is 170 patients per group.

Re-screening is allowed only in situations in which a subject underwent the screening procedures (i.e., scans and laboratory work) and owing to logistical circumstances, the allocated time window for these tests has expired. Re-screening is not permitted in cases in which the initial test results do not support eligibility based on inclusion and exclusion criteria. The subject must be classified as a screening failure at that time point.

Based on the consideration of the ethical issues related to long-term placebo use, inadequate responders in the placebo group, as determined by < 20% improvement from baseline in tender or painful joint count (TJC) and swollen joint count (SJC), will be switched at Week 12 to either ASP015K 100 mg or ASP015K 150 mg, and the dosage will be maintained until the end of treatment. In addition, subjects receiving placebo at Week 28 will be switched to either ASP015K 100 mg or ASP015K 150 mg, and the dosage will be maintained until the end of treatment. The ASP015K dose that will be started for placebo group subjects at Week 12 or Week 28 (100 or 150 mg) will be randomly chosen at baseline. The dose will be switched under the blinded condition.

Subjects who complete this study will be eligible for participation in the open-label extension study (015K-CL-RAJ2). Subjects will make a follow-up visit after the Week 52 visit if they do not enroll into the extension study on the day of the Week 52 visit.

Safety data obtained from this study will be reviewed by the Independent Data and Safety Monitoring Board (DSMB), which will make recommendations on study continuation, termination, or protocol revision from the safety perspective.

3.3 Randomization

The person responsible for assigning study drugs will randomly assign the study drugs and retain the treatment code under secrecy until the breaking of the treatment code.

The Case Registration Center will assign subjects who have been found to be eligible for enrollment in accordance with the case registration procedures to 1 of the following groups: ASP015K 100 mg, ASP015K 150 mg, or placebo groups at baseline in 1:1:1 ratio. At Week 12, subjects in the placebo group whose tender and swollen joint counts have not improved by 20% from baseline will be assessed as having an inadequate response, and they will be switched to receive either ASP015K at a dose of 100 mg or ASP015K at a dose of 150 mg. Subjects who are continuing to receive placebo at Week 28 will be switched to receive either ASP015K at a dose of 100 mg or ASP015K at a dose of 150 mg.

The dose of ASP015K (100 mg or 150 mg) to be administered to the placebo group from Week 12 or Week 28 is decided randomly at baseline, and study drugs must always be switched under double-blind conditions.

All drug numbers are assigned randomly by the Case Registration Center under blind conditions.

4 SAMPLE SIZE

A total of 510 treated patients, with 170 in each treatment group

[Rationale for the sample size]

The sample size calculation was based on ASP015K Japanese Phase IIb study (015K-CL-RAJ1) result and other Japanese studies for patients with RA [XELJANZ®, tablet, 5 mg, CTD (2013)] [CIMZIA, sc, 200 mg, Syringe, CTD (2012)]. In the 015K-CL-RAJ1 study, the primary efficacy endpoint, which was Percentage of Subjects achieving ACR20 response at Week 12/ET, was 54.5% in the 100 mg, and 65.5% in the 150 mg. Therefore, in the 015K-CL-RAJ4 study, it can be assumed to achieve same magnitude of ACR20 response at Week 12/ET for both 100 mg and 150 mg. For ACR20 response at Week 12/ET in the placebo group, it was assumed 25% based on other Japanese drugs for RA.

Meanwhile, the change in mTSS at Week 28 was estimated to be 0.5 in ASP015K 150 mg group, considering the results of the Japanese subgroup analysis of the Oral Scan Study of Tofacitinib, a drug of same class and indication (5 mg group: −0.0, 10 mg group: 0.5), Study CDP870-041 of Cimzia (100 mg group: 1.05, 200 mg group: 0.21, 400 mg group: 0.65), and Study RA0006 (200 mg group: 0.44). Similarly, the change in mTSS at Week 28 was estimated to be 2.0 in the placebo group, considering the results of the Japanese subgroup analysis of the Oral Scan Study (placebo group: 1.4), Study CDP870-041 (placebo group: 2.78), and Study RA0006 (placebo group: 2.49). The standard deviation of the change was also estimated and set to 4.0 based on the results reported from the same studies [Japanese subgroup analysis of the Oral Scan Study (placebo group: 3.15, 5 mg group: 1.09, 10 mg

group: 1.54), Study CDP870-041 (placebo group: 5.15, 100 mg group: 2.85, 200 mg group: 2.01, 400 mg group: 2.95), and Study RA0006 (placebo group: 5.52, 200 mg group: 1.83)].

On this assumption, the necessary sample size was calculated to enable to analyze ACR20 using the Fisher's exact test and the change in mTSS at Week 28 using the t-test. When a closed testing procedure (in Section 7.4.1.1 Primary Analysis) considering multiplicity is used in the primary analysis, the sample size needed to verify the hypotheses Step 1 through Step 3 at a 2-sided 5% significance level with 90% power is 151 subjects per group. By considering about 10% drop-out, this study will enroll 170 patients in each treatment group. Moreover, from the long term safety point of view, this sample size is also needed to keep 300 subjects for 6 month period, 100 subjects for 1 year period in both ASP015K 100 mg and 150 mg among RAJ2/RAJ3/RAJ4 studies.

5 ANALYSIS SETS

In accordance with International Conference on Harmonization (ICH) recommendations in guidelines E3 and E9, the following analysis sets will be used for the analyses. FAS will be the primary data set for efficacy analyses.

5.1 Full Analysis Set (FAS)

The FAS will consist of all subjects who are randomized and receive at least one dose of the study drug. This will be the primary data set for efficacy analyses.

5.2 Per Protocol Set (PPS)

5.2.1 PPS for Assessing ACR and DAS at Week 12 (PPS-ACR)

The PPS-ACR includes all subjects of the FAS who meet the following criteria. PPS-ACR is defined based on the data up to Week 12.

- Have no violation of inclusion criteria;
- Do not meet any exclusion criteria possibly interfering with the efficacy evaluation;
- Period of study treatment between the initiation of administration and Week 12 is 8 weeks (56 days) or longer (those who have discontinued study treatment due to the lack of efficacy will be included in PPS);
- Have treatment compliance of 75% or higher during the 12 weeks of study administration (Week 12 or at early termination before Week 12);
- Have evaluable ACR20 at the end of Week 12 (Week 12 or at early termination before Week 12);
- Have no major protocol violations after registration.

Subjects who are considered to have no impact on the efficacy evaluation at the Case Review Meeting will also be included in the PPS-ACR, if they are not satisfying these criteria.

5.2.2 PPS for the Assessment of the Inhibition of Joint Destruction at Week 28 (PPS-MTSS)

The PPS-MTSS includes all subjects of the FAS who meet the following criteria. PPS-MTSS is defined based on the data up to Week 28.

- Have no violation of inclusion criteria;
- Do not meet any exclusion criteria possibly interfering with the efficacy evaluation;
- Period of study treatment between the initiation of administration and Week 28 is 8 weeks (56 days) or longer (those who have discontinued study treatment due to the lack of efficacy will be included in PPS);
- Have treatment compliance of 75% or higher during the 28 weeks of study administration (Week 28 or at early termination before Week 28)
- Have evaluable change in mTSS from baseline at the end of Week 28 (Week 28 or at early termination before Week 28)
- Have at least 1 mTSS data at baseline and from Week 12
- Have no major protocol violations after registration

Subjects who are considered to have no impact on the efficacy evaluation at the Case Review Meeting will also be included in the PPS-MTSS, if they are not satisfying these criteria.

5.3 Safety Analysis Set (SAF)

The SAF is defined as all subjects who received at least one dose of the study drug.

5.4 Pharmacokinetic Analysis Set (PKAS)

The PKAS will consist of all subjects who receive at least one dose of the study drug and who provide samples for drug concentrations for at least one time point.

5.5 Pharmacodynamic Analysis Set (PDAS)

All subjects who receive at least one dose of the study drug and who provide samples for the determination of pharmacodynamic parameters at least one time point will be included in the PDAS.

6 ANALYSIS VARIABLES

6.1 Efficacy Endpoints

6.1.1 Primary Efficacy Endpoints

There are two primary efficacy endpoints.

- ACR20 response rate at Week 12 or early termination (Week 12/ET)

- Change from baseline in mTSS at Week 28 or early termination (Week 28/ET) (suppression of joint destruction)

Definitions

ACR20 response rate at Week 12 or early termination (Week 12/ET)

It is defined as percentage of subjects achieving ACR 20 response at Week 12/ET, based on C-reactive protein (CRP). ACR20 is a binary variable, with levels responder and non-responder.

Responder :

A subject will be defined as an ACR20-CRP responder at any time point (e.g. Week 12/ET) if the subject meet ALL the following American College of Rheumatology (ACR) response criteria:

- At least 20% reduction from baseline at the time point (e.g. Week 12/ET) in the number of 68 tender joint count (TJC-68)

AND

- At least 20% reduction from baseline at the time point (e.g. Week 12/ET) in the number of 66 swollen joint count (SJC-66)

AND

- At least 20% reduction from baseline at the time point (e.g. Week 12/ET) in ANY 3 or more of the 5 following ACR components
  - Subject's global assessment of arthritis pain (SGAP) (assessed using a 100 mm VAS ; score of 0 indicates no pain, and score of 100 indicates very severe pain)
  - Subject's global assessment of arthritis (SGA) (assessed using a 100 mm VAS; score of 0 indicates no disease activity, and score of 100 mm indicates very severe disease activity)
  - Physician's global assessment of arthritis (PGA) ((assessed using a 100 mm VAS; score of 0 indicates no disease activity, and score of 100 mm indicates very severe disease activity)
  - Health Assessment Questionnaire – Disease Index (HAQ-DI; score ranges from 0 to 3 with higher scores indicating greater disability)
  - CRP (Higher values indicate greater inflammation)

Baseline is defined in Section 7.11.3. A negative percent change indicates a reduction from baseline (i.e., a favorable outcome).

If the baseline value is 0 in some of ACR components, then that component is regarded as non-responder at that component and used for ACR calculation. The handling of missing data in ACR components is described at Section 7.11.1.

Non-responder: A subject will be defined as an ACR20-CRP non-responder at the time point (e.g. Week 12/ET), if the subject does not meet the ACR20-CRP responder criteria.

Change from baseline in mTSS at Week 28 or early termination (Week 28/ET) (suppression of joint destruction)

The severity of joint destruction in hands and feet (bone erosion and joint fissure narrowing) will be assessed by the third party through the mTSS method by scoring the severity of bone erosion and joint fissure narrowing of the target joints and calculate the total score. The target joints and the scoring method are described in the SOP prepared by the third party institution.

The results of the assessment will be sent to the sponsor by the third party institution.

The higher scores in mTSS represent greater damage and the positive change from baseline in mTSS represents unfavorable values.

6.1.2 Secondary Efficacy Variables

6.1.2.1 Categorical Variables

- Percentage of subjects achieving ACR 20/50/70-CRP response at each analysis visit (except ACR20 Week 12/ET)
- Percentage of subjects achieving ACR 20/50/70-ESR response at Week 12/ET

  Note: ACR 20-ESR response is similar to definition in 6.1.1 by replacing CRP with ESR.

- Percentage of subjects achieving ACR20-CRP response at Week 4 and sustaining the response at Weeks 8 and 12
- Percentage of subjects achieving ACR20-CRP response at Week 8 and sustaining the response at Week 12
- Percentage of subjects achieving mTSS change from baseline <= 0.5 at each analysis visit
- Percentage of subjects achieving mTSS change from baseline <= 0 at each analysis visit
- Percentage of subjects showing Clinically Relevant Radiographic Progression: CRRP (mTSS change from baseline >= 3) at each analysis visit
- Percentage of subjects showing Rapid Radiographic Progression: RRP (Yearly progression of mTSS >= 5)

Calculation DAS-CRP/ESR response:

DAS-CRP/ESR response is consist of following parameters, and calculated according to below description.

  - TJC (28 joints)
  - SJC (28 joints)

- CRP or ESR
- SGA

[When CRP is used]
DAS28 = $0.56\sqrt{(TJC)} + 0.28\sqrt{(SJC)} + 0.36 \ln (CRP + 1) + 0.014 \times SGA + 0.96$

[When ESR is used]
DAS28 = $0.56\sqrt{(TJC)} + 0.28\sqrt{(SJC)} + 0.70 \ln ESR + 0.014 \times SGA$

Note: CRP values measured in mg/dL will be converted to mg/L for analysis purposes as: value in mg/L = value in mg/dL × 10; SGA is measured on 100 mm VAS. If any component is a missing value, then DAS28-CRP/ESR will be a missing value. If ESR is 0 then DAS28-ESR is missing.

- Percentage of subjects achieving DAS28-CRP score <2.6 at each analysis visit
- Percentage of subjects achieving DAS28-ESR score <2.6 at each analysis visit
- Percentage of subjects achieving DAS28-CRP score <=3.2 at each analysis visit
- Percentage of subjects achieving DAS28-ESR score <=3.2 at each analysis visit

Definition EULAR:

DAS28 after treatment	DAS28 improvement (DAS28 before treatment - DAS28 after treatment)		
	> 1.2	> 0.6 and ≤ 1.2	≤ 0.6
≤ 3.2	Good response	Moderate response	No response
> 3.2 and ≤ 5.1	Moderate response	Moderate response	No response
> 5.1	Moderate response	No response	No response

- Percentage of subjects in DAS28-CRP EULAR response criterion of "Good Response" at each analysis visit
- Percentage of subjects in DAS28-ESR EULAR response criterion of "Good Response" at each analysis visit
- Percentage of subjects in DAS28-CRP EULAR response criterion of "Good Response" or "Moderate Response" at each analysis visit
- Percentage of subjects in DAS28-ESR EULAR response criterion of "Good Response" or "Moderate Response" at each analysis visit
- Percentage of subjects achieving ACR/EULAR score for remission at each analysis visit
  If all of the following 4 parameters are fulfilled, it is defined as remission:
  - TJC (68 joints) <= 1
  - SJC (66 joints) <= 1
  - CRP <= 1 mg/dL
  - SGA <= 1 cm (on a VAS of 0-100 mm).

  Note: All 4 conditions must be fulfilled to be in remission. If any component is a missing value, then ACR/EULAR score for remission will be a missing value.

- Percentage of subjects with a SDAI score of <= 3.3 at each analysis visit
  SDAI score is consist of following parameters, and calculated according to below description.
  - TJC (28 joints)
  - SJC (28 joints)
  - SGA (0-10 cm VAS)
  - PGA (0-10 cm VAS)
  - CRP (mg/dL)

  SDAI = TJC +SJC +SGA +PGA + CRP

  Note: If any component is a missing value, then SDAI score will be a missing value.
- Percentage of subjects with a CDAI score of <= 2.8 at each analysis visit
  CDAI score is consist of following parameters, and calculated according to below description.
  - TJC (28 joints)
  - SJC (28 joints)
  - SGA (0-10 cm VAS)
  - PGA (0-10 cm VAS)

  CDAI = TJC +SJC +SGA +PGA

  Note: If any component is a missing value, then CDAI score will be a missing value.
- Percentage of subjects achieving HAQ-DI (<=0.5) at each analysis visit
- Percentage of subjects achieving decrease of baseline HAQ-DI of at least 0.22 at each analysis visit
- Percentage of subjects achieving SF-36v2 of difference >= 5 at each analysis visit

6.1.2.2 Continuous Variables

Continuous efficacy variables are:

- Raw value and change from baseline in the following assessments:
  - TJC (68 joints)
  - SJC (66 joints)
  - CRP
  - ESR
  - SGAP (VAS) (100 mm VAS)
  - SGA (VAS) (100 mm VAS)
  - PGA (VAS) (100 mm VAS)
  - HAQ-DI (See Section 10.5 Appendix 5: Computation of HAQ-DI Score)
  - DAS28-CRP score
  - DAS28-ESR score
  - SF-36v2®

- SDAI score
- CDAI score
- WPAI score
- mTSS
- Joint Space Narrowing Score
- Erosion Score
- Yearly progression of mTSS

Note: yearly progression will be calculated as follows.

$$\frac{[\text{the final value of mTSS up to Week 52}] - [\text{the baseline value of mTSS}]}{[\text{date of the final value of mTSS up to Week 52}] - [\text{date of baseline value of mTSS}]} \times 365.25$$

Note: if subject switched from placebo then the final value of that subject is the final value before switching placebo.

6.1.2.3 Time to Event Variables

- Time to the first ACR20-CRP response up to Week 12
- Time to the first ACR50-CRP response up to Week 12
- Time to the first ACR70-CRP response up to Week 12
- Time to the first occurrence of DAS28-CRP <2.6 up to Week 12
- Time to the first occurrence of SDAI score <= 3.3 up to Week 12
- Time to the first occurrence of CDAI score <= 2.8 up to Week 12
- Time to the first occurrence of ACR/EULAR score for remission up to Week 12

6.2 Safety Variables

- AEs
- Vital signs (body temperature, pulse rate and blood pressure in sitting position)
- Body weight
- Central ECG
- 12-lead ECG
- Central electrocardiogram
- Chest radiography
- Laboratory assessments

Treatment emergent adverse events (TEAEs) will be defined as any adverse event that started or worsened in severity after initial dose of study drug through Week52 or Withdrawal.

AE for the subjects who switched ASP015K 100 mg or ASP150 mg from Placebo at Week 12 or Week 28 will be handled as follows.

- AE onset before the first dose of week12 study drug will be considered to be occurred at Placebo.
- AE onset after the first dose of week12 study drug before the first dose of week28 study drug will be considered to be occurred at Placebo if the subject did not switch Placebo at Week 12. Else, if the subject switched ASP015K 100 mg or 150 mg at Week12, then the AE onset after the first dose of week12 study drug before the first dose of week28 study drug will be considered to be occurred at ASP015K 100 mg or 150 mg.
- AE onset after the first dose of week28 study drug will be considered to be occurred at ASP015K 100 mg or 150 mg, which is switched treatment from Placebo.

A drug-related TEAE is defined as any TEAE with possible, probable or missing relationship to study drug as assessed by the investigator. Moreover, any TEAE with missing relationship to study drug or reference drug was counted as drug related (probable).

6.3 Pharmacokinetic Variables

Refer to PK analysis plan.

6.4 Pharmacodynamic Variables

- Change from baseline in MMP-3 level
- Change from baseline in vascular endothelial growth factor (VEGF) level
- Change from baseline for following variables in lymphocyte subset assays
- CD3+/Lymphocytes (%), CD3 (cells/uL)
- CD8+/Lymphocytes (%), CD8 (cells/uL)
- CD4+/Lymphocytes (%), CD4 (cells/uL)
- CD19+/Lymphocytes (%), CD19 (cells/uL)
- CD4/CD8 Ratio
- (CD16 or CD56)+/Lymphocytes (%), CD16 or CD56 (cells/uL)
- (CD16+ and CD56+)/CD3- (%), CD16+ and CD56+ (cells/uL)
- (CD16- and CD56+)/CD3- (%), CD16- and CD56+ (cells/uL)
- (CD16+ and CD56-)/CD3- (%), CD16+ and CD56- (cells/uL)
- (CD16- and CD56-)/CD3- (%), CD16- and CD56- (cells/uL)
- CD56 bright/CD3- (%), CD56 bright (cells/uL)
- CD56 dim/CD3- (%), CD56 dim (cells/uL)

6.5 Other Variables

- Duration of RA (years), calculated as: (Date of Baseline mTSS taken – Onset Date of RA +1) / 365.25, and then rounded to one decimal place. The onset date of RA doesn't have the day (e.g., 2013-03), therefore the first day of the month will be used (2013-03-01).
  *: Missing onset date of RA will be imputed. For example, if month is missing (e.g., 2013), then the first day of January will be used (2013-01-01).

- Age at onset of RA (years), calculated as: (date of onset of RA – date of birth +1) / 365.25. Onset Date of RA will be handled in the same way as described above.
- Duration of exposure (days)
  Duration of exposure will be calculated in days, using the following formula:
  ('Date of Last Dose of Study Drug'* - 'Date of First Dose of Study Drug'**) + 1
  * = Max(EX1ENDT, EX2ENDT, EX3ENDT) [Study Drug and Reference Drug Dosing-page of the CRF]
  ** = EX1STDT [Study Drug Dosing -page of the CRF]
- Overall treatment compliance (during overall treatment period)
  In the study, the subjects will be instructed to take 2 tablets once per day, totaling a daily dose of study dose. Therefore, overall treatment compliance will be based on number of tablets and calculated as follows:

  Treatment compliance (%) =

  [Total number of tablets actually received in the treatment period]

  ----------x 100
  [Total number of tablets planned to receive in the treatment period]

  where total number of tablets planned to receive in the treatment period = 2 × number of days the subject was in the treatment period (the previous date of Week 52 visit [for completed subjects] / date of withdrawal date [for discontinued subjects] – date of initial dose of study drug +1), and total number of tablets actually received in the treatment period will be calculated as:
  (total number of tablets dispensed) – (total number of tablets returned) – (total number of tablets lost).
- Treatment compliance during the 12 weeks of study (up to Weeks 12)
  Treatment compliance (%) =

  [Total number of tablets actually received up to Weeks 12]

  ----------x 100
  [Total number of tablets planned to receive up to Weeks 12]

  where total number of tablets planned to receive up to Weeks 12 = 2 × number of days the subject was during 12 weeks (the previous date of Weeks 12 visit date [for completed Week 12 subjects] / date of withdrawal date [for discontinued before Week 12 subjects] – date of initial dose of study drug +1), and total number of tablets actually received up to Weeks 12 will be calculated as:
  (total number of tablets dispensed up to Weeks 12) – (total number of tablets returned up to Weeks 12) – (total number of tablets lost up to Weeks12).
- Treatment compliance during the 28 weeks of study (up to Weeks 28)

  Treatment compliance (%) =

[Total number of tablets actually received up to Weeks 28]

----------x 100
[Total number of tablets planned to receive up to Weeks 28]

where total number of tablets planned to receive up to Weeks 28 = 2 × number of days the subject was during 28 weeks (the previous date of Weeks 28 visit date [for completed Week 28 subjects] / date of withdrawal date [for discontinued before Week 28 subjects] – date of initial dose of study drug +1), and total number of tablets actually received up to Weeks 28 will be calculated as:
(total number of tablets dispensed up to Weeks 28) – (total number of tablets returned up to Weeks 28) – (total number of tablets lost up to Weeks28).

*If Weeks 28 visit date is missing but not discontinue before Day 197, then the total number of tablets planned to receive up to Weeks 28 = 2 × 196 (i.e. days from the previous date of scheduled visit at Week 28 – the date of initial dose of study drug + 1)

- Number of prior DMARD biologics, calculated as number of unique kinds of DMARD biologics prior to initial dose from following CRF page.
  [Concomitant Medication 2B -page of the CRF]
  [Biologic DMARD -page of the CRF]
- Number of prior DMARDs (including biologics), calculated as number of unique kinds of DMARD prior to initial dose from following CRF page.
  [Concomitant Medication 2B -page of the CRF]
  [Concomitant Medication(Non-Biologic DMARD) -page of the CRF]
  [Non Biologic DMARD_MTX -page of the CRF]
  [Non Biologic DMARD except for MTX -page of the CRF]
  [Biologic DMARD -page of the CRF]
- MTX Dose at Baseline (mg/week) is defined as the dose which is taken during 6 days before “date of initial dose of study drug” up to “the initial dose of study drug”.

7 STATISTICAL METHODOLOGY

7.1 General Considerations

- For continuous variables, descriptive statistics will include the number of subjects (n), mean, standard deviation (SD), median, minimum, and maximum.
- For categorical variables: number and percentages of subjects will be described.
- For time-to-event variables: number and percentage of subjects with the event using Kaplan-Meier method, the cumulative event rate will be estimated and a plot will be constructed.
- All data processing, summarization, and analyses will be performed using SAS Drug Development (ver. 4.5), and PC-SAS (ver. 9.4) or higher versions.
- Specifications for table, figures, and data listing formats can be found in the TLF specifications for this study.

- For the definition of subgroups of interest can be referred to section 7.8.
- Summaries based on FAS and PPS (e.g. disposition, baseline and efficacy data) will be presented by actual treatment groups, unless specifically stated otherwise. Safety analysis and other summaries based on SAF will also be presented by actual treatment groups.
- Initial Randomization Set of efficacy analyses except for mTSS, erosion score, and joint space narrowing score are those of the efficacy data up to Week 12 and include the primary analysis (ACR20 at Week 12/ET) as described 7.4.1.1. This set will be analyzed by the initially treated groups (ASP015K 100 mg, ASP015K 150 mg, and Placebo) and statistical testing will be performed for ASP015K 100 mg, and 150 mg, compared with Placebo. In addition, statistical testing will also be conducted at Week 28/ET.

Table 3 Initial Randomization Set

Initial Randomization Arm Code	**Initial Randomization Label**	**Analysis Scope**
PLACEBO	Placebo	Screening, Week 0 to Week 12, Week 12/ET, Week 28/ET
100MG	100 mg	
150MG	150 mg	

- Treatment Sequence Set of efficacy analyses except mTSS, erosion score, and joint space narrowing score are those of the efficacy data up to Week 52. Taking into consideration a switch from placebo to active treatment at Week 12 or Week 28, this set will be analyzed by following groups defined below. The four types of "Placebo to ASP015K xx mg at Week xx" groups are not for primary objective and are conducted optionally for efficacy and safety.

Table 4 Treatment Sequence Set

Treatment Sequence Code	**Treatment Sequence Label**	**Week0- Week12**	**Week 12- Week 28**	**Week 28- Week 52**	**Analysis Scope**
SEQ1	100mg,	ASP015K 100 mg	ASP015K 100 mg	ASP015K 100 mg	Screening, Week 0 to Week 52, EOT
SEQ2	150mg	ASP015K 150 mg	ASP015K 150 mg	ASP015K 150 mg	
SEQ3	Placebo to 100mg at Week 12	Placebo	ASP015K 100 mg	ASP015K 100 mg	
SEQ4	Placebo to 150mg at Week 12	Placebo	ASP015K 150 mg	ASP015K 150 mg	
SEQ5	Placebo to 100mg at Week 28	Placebo	Placebo	ASP015K 100 mg	
SEQ6	Placebo to 150mg at Week 28	Placebo	Placebo	ASP015K 150 mg	

  - 100 mg: subjects who initially treated ASP015K 100 mg
  - 150 mg: subjects who initially treated ASP015K 150 mg

  - Placebo to 100 mg at Week 12: subjects who initially treated Placebo group and switched 100 mg at Week 12 and at least one dose of ASP015K 100 mg drug after switched
  - Placebo to 150 mg at Week 12: subjects who initially treated Placebo group and switched ASP015K 150 mg at Week 12 and at least one dose of ASP015K 150 mg drug after switched
  - Placebo to 100 mg at Week 28: subjects who initially treated Placebo group and switched ASP015K 100 mg at Week 28 and at least one dose of ASP015K 100 mg drug after switched
  - Placebo to 150 mg at Week 28: subjects who initially treated Placebo group and switched ASP015K 150 mg at Week 28 and at least one dose of ASP015K 150 mg drug after switched

- The analyses for mTSS (Appendix1, Appendix 2), erosion score, and joint space narrowing score will be performed at Week 28/ET and Week 52/ET, and statistical testing will also be conducted at Week 28/ET, Week 52/ET , compared with Placebo. These analyses include primary analysis (mTSS change from baseline at Week 28/ET).
- MedDRA11.1 will be used as the coding dictionary for adverse event and medical history.
- Statistical hypothesis testing will be performed only if specified. All statistical comparisons of treatment groups will be versus placebo (i.e., each ASP015K treatment versus placebo) and will be made using two sided tests at the 0.05 significance level, unless stated otherwise. Multiplicity adjustment will be done in the primary analysis and other analysis if specified.
- Statistical analysis including hypothesis testing for efficacy variables to compare 100 mg or 150 mg versus Placebo, only the data for the pairwise comparison will be used.
- All statistical results will be presented, as appropriate, by treatment group, by treatment group and scheduled visit; the schedule of assessments is provided in 7.11.4.
- In addition, demographics and other baseline characteristics, and safety results will be provided for all ASP015K treatment groups combined ("ASP015K Total") and for all treatment groups, including placebo, combined ("Total"). Efficacy presentations will not contain combined treatment groups.
- Chi-square test is continuity corrected in all analysis.
- Confidence interval for binary outcome is continuity corrected in all analysis.
- The values below the lower limit of quantitation (BQL) for β-D-glucan, hCG, Troponin, CK-MB will be treated as it is and these variables are not used for descriptive statistics and displayed in the listings. As for β-D-glucan, it is used for shift-from baseline analysis.
- If anti-CCP antibody (U/mL) is beyond measurement (i.e. >-500 U/ml) then Anti-CCP antibody is regarded as 500 U/ml, and categorized in >-500 U/ml. Also, if anti-CCP antibody (U/mL) is below measurement (i.e. <0.6 U/ml) then Anti-CCP antibody is regarded as 0.6 U/ml, and categorized in < 4.5 U/mL.

- CRP (mg/dL) is below measurement (i.e. <0.01 mg/dL) then CRP is regarded as 0.01 U/ml, and used for the calculation of DAS28-CRP, etc. and categorized in <1.0 mg/dL.
- If Rheumatoid Factor (IU/mL) is below measurement (i.e. <3 IU/mL), then Rheumatoid Factor is regarded as 3 IU/mL and categorized in <- 15 IU/ml.
- Baseline for safety is defined as last non-missing value before the first dose of study drug for all subjects including placebo assigned subjects.
- Baseline for efficacy is defined as value at Day1 before the first dose of study drug for all subjects including placebo assigned subjects.
- Change from baseline to post-baseline will be calculated as: post-baseline value - baseline value. If the baseline value is missing, then that subject is not included in the calculation at any visit.
- Percent change from baseline to post-baseline will be calculated as: 100 × (change /baseline). If the baseline value is 0 or missing, or post-baseline value is missing, then Percent change from baseline to post-baseline is missing.
- If onset date of AE is Unknown, then missing onset date of AE will be imputed as following steps.
    1. If only the day of the month is missing (e.g., 2013-03), then the first day of the month will be used (2013-03-01),
    2. If both day and month are missing (e.g., 2013), then the first day of January will be used (2013-01-01).
    3. If imputed onset date of AE is earlier than the first date of study drug, then the onset date of AE is onset at the first date of study drug, because AE data in CRF is collected only after the first date of study drug.

7.2 Study Population

7.2.1 Disposition of Subjects

The following subject data will be summarized and presented with regard to All Subjects with Screening Failure Log and Case Report Form.

- Screened
- Screen failed
- Randomized

The following subject data will be summarized and presented by Initial Randomization Set and Treatment Sequence Set:

- Subjects randomized and dosed (same as the FAS)
- Subjects who prematurely discontinued from the study period during overall period
- Subjects who prematurely discontinued from the study period by 4 weeks
- Subjects who included/excluded of FAS, PPS-ACR, PPS-MTSS, SAF, PKAS, and PDAS

- Subjects who were excluded from FAS, PPS, SAF, PKAS, and PDAS will be summarized by reason for exclusion.

The following analysis will be conducted for Initial Randomization Set.
- For time-to-withdrawal from initial dose up to Day 92: Kaplan-Meier plot will be constructed.

Withdrawal Due to Lack of Efficacy:

In addition to the disposition summaries, the percentage of subject withdrawal due to lack of efficacy as the primary reason for withdrawal will be compared (each ASP015K dose group versus placebo) using Fisher's exact test, and the analysis will be based on SAF by Initial Randomization Set. Fisher's exact p-values (No multiplicity adjustment) will be provided.

The following subject data will be summarized and presented by Initial Randomization Set and Treatment Sequence Set
- Subjects who discontinued after starting the study period will be summarized for primary reason for withdrawal during overall period.
- Subjects who discontinued after starting the study period will be summarized for primary reason for withdrawal by 4 weeks.

[Analysis population: FAS]
- Subjects who applied to the improvement rate from baseline in tender/swollen joint count(68/66) less than 20% at Week12

7.2.2 Demographic and Other Baseline Characteristics

All demographic and other baseline characteristics (defined below) will be summarized by treatment group. To explore the imbalance among treatment groups, comparisons with respect to categorical variables (e.g., sex) will be based on chi-squared test (continuity corrected), and comparisons with respect to continuous variables (e.g., BMI) will be based on a one-way ANOVA model with fixed effect for treatment group. All analyses/summaries except for (7.2.2.4) will be based on the SAF, FAS, and PPS.

If the imbalance of the factors between treatment groups is found (at a two-sided 0.05 significance level) and considered to clinically affect the primary variable, an analysis adjusted for the factor will be performed to assess the effect of the primary variable on the primary analysis.

7.2.2.1 Demographics

The following demographic variables will be summarized and presented for each treatment group.

Table 1. Demographic Variables and Analysis Methods

Item	Classification	Testing Method
Age (at the time of Informed Consent)	Measurement value	one-way ANOVA
	<65, >= 65	-
Sex	Male, Female	chi-squared test
Height (cm)	Measurement value	one-way ANOVA
Body Weight[Screening] (kg)	Measurement value	one-way ANOVA
	<- 40 kg, 40 kg <- 60 kg, 60 kg <- 80 kg, > 80 kg	-
BMI [Screening] (kg/m2)	Measurement value	one-way ANOVA
Complications	No, Yes	
Previous Medications	No, Yes	
Concomitant Medications	No, Yes	
Concomitant DMARD except for MTX at Baseline	No, Yes	chi-squared test
Concomitant Steroid at Baseline*	No, Yes	chi-squared test
Prednisone dose at Baseline (mg/day) *	Measurement value	one-way ANOVA
	None, 0 <- 5 mg/day, > 5 mg/day	

*: Concomitant Steroid and Prednisone dose is applied for “Oral Use”.

7.2.2.2 Baseline Disease Activity

The following baseline disease activity variables will be summarized by descriptive statistics, and imbalance between treatment groups will be considered based on a one-way ANOVA model with fixed effect for treatment group.

- TJC-68
- TJC-28
- SJC-66
- SJC-28
- SGAP (100 mm VAS)
- SGA (100 mm VAS)
- PGA (100 mm VAS)
- HAQ-DI (scale 0 - 3)
- CRP (mg/dL)
- ESR (mm/hr)
- DAS28-CRP
- DAS28–ESR
- SDAI score
- WPAI
- SF-36V2®

- mTSS
- Erosion Score
- Joint Space Narrowing Score
- Estimated mTSS yearly progression
  Note: yearly progression from onset to baseline; baseline mTSS/ Duration of RA (years)

In addition, following baseline efficacy categorical variables will be summarized as well.

- Baseline mTSS (<=Median, > Median)
- Baseline CRP (< 1.0, >= 1.0)
- Baseline DAS28-CRP (<= 3.2, 3.2 <- 5.1, > 5.1)
- Baseline DAS28-ESR (<= 3.2, 3.2 <- 5.1, > 5.1)

All summaries will be provided separately for SAF, FAS and PPS.

7.2.2.3 RA History

The following RA history variables will be summarized and presented for each treatment group. All summaries will be provided separately for SAF, FAS and PPS.

Table 2. RA History and Analysis Methods

Item	Classification	Testing Method
ACR 1991 Revised Criteria for Global Functional Status in RA	Class I, Class II, Class III	chi-squared test
Steinbrocker Classification	Stage I, Stage II, Stage III, Stage IV	chi-squared test
Prior Surgical Procedure to Treat RA	No, Yes	chi-squared test
Duration of RA (years)	Measurement value	one-way ANOVA
	<5 years, >= 5 years	-
	< 1 year, 1 year -<5 years, 5 years -< 10 years, >- 10 years	-
Onset Age of RA (years), calculated as: (onset date of RA – date of birth+1) / 365.25	-	one-way ANOVA
Anti-CCP Antibody (U/mL)	Measurement value	one-way ANOVA
	Low-positive or negative (<- 3 x ULN), High-positive (> 3 x	-

		ULN)	
		< 4.5 U/mL, 4.5 -< 13.5 U/mL, 13.5 -< 90 U/mL, 90 -< 180 U/mL, 180 -< 500 U/mL, >- 500 U/ml	-
Rheumatoid Factor (IU/ml)		Measurement value	one-way ANOVA
		Low-positive or negative (<- 3 x ULN), High-positive (> 3 x ULN)	-
		<- 15 IU/ml, 15 <- 45 IU/ml, 45 <- 150 IU/ml, > 150 IU/ml	-
Prior MTX	Maximum Dose	Fixed Unit: mg/week	-
	Reactivity	Response, Inadequate Response, Unknown	-
	Tolerance	Intolerance, Tolerance, Unknown	-
MTX Dose at Baseline (mg/week)		Measurement value	one-way ANOVA
		0 <- 8 mg/week, 8 <- 12 mg/week, > 12 mg/week	
		0 <- 6 mg/week, 6 <- 8 mg/week, 8 <- 10 mg/week, 10 <- 12 mg/week, 12 <- 16 mg/week, 16 <- mg/week	
Prior Non-Biologic DMARD Except for MTX	Use	Non-User, User	chi-squared test
Prior Anti-TNF DMARD Use	Use	Non-User, User	chi-squared test
Prior Biologic DMARD	Use	Non-User, User	chi-squared test
Number of Prior DMARD Biologics		0, 1, 2, >= 3	chi-squared test
Number of Prior DMARDs (including biologics)		0, 1, 2, >= 3	chi-squared test
Prior Adalimumab	Use	Non-User, User	-
	Reactivity	Response , Inadequate Response, Unknown	-

	Tolerance	Intolerance, Tolerance, Unknown	-
Prior Etanercept	Use	Non-User, User	-
	Reactivity	Response, Inadequate Response, Unknown	-
	Tolerance	Intolerance, Tolerance, Unknown	-
Prior Golimumab	Use	Non-User, User	-
	Reactivity	Response, Inadequate Response, Unknown	-
	Tolerance	Intolerance, Tolerance, Unknown	-
Prior Infliximab	Use	Non-User, User	-
	Reactivity	Response, Inadequate Response, Unknown	-
	Tolerance	Intolerance, Tolerance, Unknown	-
Prior Certolizumab Pegol	Use	Non-User, User	-
	Reactivity	Response, Inadequate Response, Unknown	-
	Tolerance	Intolerance, Tolerance, Unknown	-
Prior Abatacept	Use	Non-User, User	-
	Reactivity	Response, Inadequate Response, Unknown	-
	Tolerance	Intolerance, Tolerance, Unknown	-
Prior Tocilizumab	Use	Non-User, User	-
	Reactivity	Response, Inadequate Response, Unknown	-
	Tolerance	Intolerance, Tolerance, Unknown	-
Prior Rituximab	Use	Non-User, User	-
	Reactivity	Response, Inadequate Response, Unknown	-
	Tolerance	Intolerance, Tolerance, Unknown	-
Prior Denosumab	Use	Non-User, User	-
	Reactivity	Response, Inadequate Response, Unknown	-
	Tolerance	Intolerance, Tolerance, Unknown	-

7.2.2.4 Medical History

Medical history will be coded using MedDRA/J (Version 11.1) and summarized. All summaries will be provided for SAF.

7.2.3 Previous and Concomitant Medications

Previous DMARD medications will be summarized as described in 7.2.2.3.

In addition to this, previous and concomitant medications will be coded using WHODDE(B2) (V2011SEP) and summarized by "Non-Biologic DMARD" and "previous and concomitant medications except for Non-Biologic DMARD" with preferred WHO name, respectively from following CRF page. Concomitant medications are defined as any drug medications after the first dose of study drug up to the last dose date of study drug and before the last efficacy evaluation at Week52 or Withdrawal.

[Concomitant Medication 2B -page of the CRF]
[Concomitant Medication (Non-Biologic DMARD)]
Subjects taking the same medication multiple times will be counted once per medication.

All summaries will be provided for SAF.

7.3 Study Drugs

7.3.1 Exposure

The following information on drug exposure will be presented for each treatment group for the SAF and presented by overall period and treatment period up to Week 12, separetely.

Duration of exposure will be summarized in two ways.

- Descriptive statistics will be presented by treatment group.
- Exposure time will be categorized according to the following categories by treatment group:
  <Category for Treatment period up to Week 12>
  - 1 - < 29 days
  - 29 - < 57 days
  - 57 - < 85 days
  - 85 days -

  < Category for Overall Period>
  - 1 - < 85 days
  - 85 - < 197 days

- 197 - < 281 days
- 281 - < 365 days
- 365 days -

Counts and percentages of subjects in each of these categories will be summarized for each treatment group for the SAF.

7.3.2 Treatment Compliance

Overall treatment compliance, treatment compliance during the 12 weeks of study (up to Weeks 12), treatment compliance during the 28 weeks of study (up to Weeks 28) with the dosing schedule will be examined for subjects in the SAF whose total study drug count and first and last days of treatment are known. These analyses will be summarized in two ways for the SAF:

- Descriptive statistics will be presented by Initial treatment set and 4 types of Placebo treated subjects.
- Descriptive statistics up to Week 12 will be presented by Initial treatment set. Similarly, descriptive statistics up to Week 28 will be presented by Initial treatment set.
- Percent compliance will be categorized according to the following categories by treatment group:
  - <50%
  - 50% - <75%
  - 75% - <90%
  - 90% - 100% or greater
- Percent compliance up to Week 12 will be categorized according to the following categories by treatment group Similarly, Percent compliance up to Week 28 will be presented by Initial treatment set:
  - <50%
  - 50% - <75%
  - 75% - <90%
  - 90% - 100% or greater

For Drug Suspension and Interruption, following will be summarized.

- Experience for Drug Suspension
- Experience for Drug Interruption
- The Number of Experience for Drug Suspension per Subject
- The Number of Experience for Drug Suspension per Subject, Categorized
  - 0
  - 1
  - 2
  - >= 3

- The Number of Experience Drug Interruption Per Subject, Categorized
  - 0
  - 1
  - 2

7.4 Analysis of Efficacy

Efficacy analysis will be conducted on FAS and PPS. The interpretation of results from statistical tests will be based on the FAS. In the primary analysis, the purpose of using PPS is to assess the robustness of the results from the statistical tests based on FAS. Efficacy analysis up to Week 12 including primary analysis will be performed by Initial Randomization Set. Efficacy analysis up to Week 52 will be performed by Treatment Sequence Set for secondary purpose. Details for analysis groups will be described in 10.1 Appendix1 and 10.2 Appendix2..

7.4.1 Analysis of Primary Endpoint(s)

Summary for analysis of primary endpoints are displayed at Section 10.1 Appendix 1.

7.4.1.1 Primary Analysis

Primary analysis will be conducted on FAS. However, the assessment of joint destruction will be performed in subjects with at least 1 data at baseline and after Week 12. Therefore, post-baseline values for mTSS, Erosion Score, Joint Space Narrowing Score before lower allowance limit of Week 12 (Day48) will not be used for the analysis (it will be displayed in the listing).

For ACR20-CRP response at Week 12/ET, pairwise comparisons to placebo will be performed at each ASP015K dose level using Fisher's exact test in the primary analysis. For mTSS change from baseline at Week 28/ET, pairwise comparisons to placebo will be performed at each ASP015K dose level using rank analysis of covariance (RANCOVA) with treatment group (Placebo, ASP015K 100 mg, ASP015K 150 mg) as factor and baseline rank mTSS as covariate in the primary analysis.

Closed testing procedure shown below will be used for multiplicity adjustment in the primary analysis.

Four null hypotheses will be constructed:

- H_{01}: ACR20-CRP response at Week 12/ET in ASP015K 150 mg is equal to that in placebo
- H_{02}: ACR20-CRP response at Week 12/ET in ASP015K 100 mg is equal to that in placebo
- H_{03}: mTSS change from baseline at Week 28/ET in ASP015K 150 mg is equal to that in placebo

- H_{04}: mTSS change from baseline at Week 28/ET in ASP015K 100 mg is equal to that in placebo

The accompanying alternative hypotheses are:

- H_{11}: ACR20-CRP response at Week 12/ET in ASP015K 150 mg is not equal to that in placebo
- H_{12}: ACR20-CRP response at Week 12/ET in ASP015K 100 mg is not equal to that in placebo
- H_{13}: mTSS change from baseline at Week 28/ET in ASP015K 150 mg is not equal to that in placebo
- H_{14}: mTSS change from baseline at Week 28/ET in ASP015K 100 mg is not equal to that in placebo

Step 1. ACR20-CRP response at Week 12/ET: ASP015K 150 mg vs. Placebo

Step 2. ACR20-CRP response at Week 12/ET: ASP015K 100 mg vs. Placebo

Step 3. mTSS change from baseline at Week 28/ET: ASP015K 150 mg vs. Placebo

Step 4. mTSS change from baseline at Week 28/ET: ASP015K 100 mg vs. Placebo

The null hypothesis at Step 1 will be tested at a two-sided significance level of 0.05.

If it is statistically significant, the next step will be initiated and implemented in the same manner. Otherwise, it is completion of the hypothesis test. These hypothesis tests continue up to Step 4 unless it is rejected.

For each comparison, only the data for two treatment groups to be compared will be used. For the missing imputation of ACR20-CRP at Week 12/ET, Last Observation Carried Forward (LOCF) methodology will be used.

For subjects who discontinued at or before Week 28 or switched to receive ASP015K instead of placebo at Week 12 due to the lack of efficacy, mTSS at Week 28 will be extrapolated using linear extrapolation method based on the mTSS at baseline and early termination or Week 12 (before switching).

Cumulative probability plot for the change from baseline in mTSS score at Week 28/ET will be created.

7.4.1.2 Sensitivity Analysis

Multiplicity adjustment is not executed for these analyses.

ACR20-CRP response at Week 12/ET

In order to assess the robustness of findings from the primary efficacy analysis, the following sensitivity analyses will be performed:

- LOCF for components and NRI for response (see Section 7.11.1. for details)

- Analysis using the PPS-ACR as the analysis set with closed testing procedure
  Note: The FAS will be replaced with the PPS as analysis set, to assess robustness results from FAS.
- Observed data (data as collected, no imputation)
- Multiple Imputation, assuming Missing at Random Mechanism (see Section 7.11.1. for details)
- Placebo Multiple Imputation (pMI) (see Section 7.11.1. for details)

mTSS change from baseline at Week 28/ET

In order to assess the robustness of findings from the primary efficacy analysis, the following sensitivity analyses will be performed:

- Rank Analysis of Covariance with LOCF (see Section 7.11.1. for details)
- Rank Analysis of Covariance with Observed data (data as collected, no imputation)
- Rank Analysis of Covariance (LEP) using the PPS-mTSS as the analysis set with closed testing procedure
- Analysis of Covariance (ANCOVA) with treatment group (Placebo, ASP015K 100 mg, ASP015K 150 mg) and baseline mTSS as covariate. Imputation method is linear extrapolation. 95% confidence interval will be calculated.
- Analysis of Covariance (ANCOVA) with treatment group (Placebo, ASP015K 100 mg, ASP015K 150 mg) and baseline mTSS as covariate. Imputation method is Multiple Imputation (see Section 7.11.1. for details).

7.4.1.3 Subgroup Analysis

In order to evaluate homogeneity of treatment effects across subjects with different demographic and baseline characteristics, subgroup analysis of the primary efficacy variables will be performed, respectively. In addition, logistic regression model will be performed for each subgroup category and odds ratios, and 95% confidence intervals will be provided for each subgroup. The subgroups in Section 7.8 will be analyzed. As for subgroup analysis, it is not executed any statistical testing. These subgroup analyses will be displayed graphically as Dot-and-Forest plot.

7.4.1.4 Adjustment for imbalances in demographics and other baseline characteristics

As for the variables in which imbalance is present in 7.2.2, adjusted analysis will be performed to evaluate the influence.

ACR20-CRP response at Week 12/ET

Logistic model with each treatment group as a factor and the variables in which imbalance is present as a covariate.

mTSS change from baseline at Week 28/ET

Rank Analysis of Covariance (RANCOVA) with treatment group (Placebo, ASP015K 100 mg, ASP015K 150 mg) and adjusted variable (if imbalance will be detected) as factor and baseline rank mTSS as covariate.

7.4.1.5 Time to the first response analysis

Time to the first ACR20-CRP response analysis will be analyzed as following procedure. This analysis will be applied during first 12 weeks of treatment period, because placebo treated subjects can be switched to ASP015K after Week 12.

- Time to the first response analysis will be analyzed using KM profile plots for each treatment group, and log-rank test will be used for comparing each ASP015K dose group time-to-event profile versus placebo. This analysis will be applied up to Week 12.
- For each subject, time to the first event (days) will be defined as the number of days from the date of initial dose of study drug to the date the first occurrence of the event, and calculated as: date of the first occurrence of the event (most recent assessment among each components if the assessment date is different among each components) – date of initial dose of study drug +1. In addition to this, Cox proportional hazards model with a main effect for treatment group will be used comparing each ASP015K dose group versus placebo.
- Subjects who complete the treatment period through Week 12 without having the event will be censored at one day before the dosing at Week 12 day up to Day 92, and the censored time will be calculated as: one day before the dosing at Week 12 up to Day 92 – date of initial dose of study drug +1.
- In the time to response analysis, each subject will be classified as either a responder (event = 1 in time-to-event model) or a non-responder (event = 0 in time-to-event model), based on the definition in Section 6.1.1. The definition of event is the same as the definition of the binary response.
- Subjects who prematurely discontinued the treatment period through Week 12 (i.e. before the dosing at Week 12) without having the event up to Day 92 will be censored at the date of early discontinuation, and the censored time will be calculated as: date of early discontinuation from the treatment period up to Day 92 – date of initial dose of study drug +1.

7.4.2 Analysis of Secondary Endpoints

Refer to Section 10.2 Appendix 2. They include categorical, continuous, and time-to-event variables, with "X" indicating that the variable will be analyzed. As for secondary variables, no multiplicity adjustment will be done. Basically, analysis of secondary endpoints except for mTSS, erosion score, and joint space narrowing score will be performed by Initial Randomization Set and Treatment Sequence Set defined at Secrion7.1.

The secondary variables and analyses are as follows:

<Initial Randomization Set>

- Categorical variables at each visit (including Baseline, Weeks 4, 8, 12, 12/ET, Week 28/ET) will be analyzed using Fisher's exact test, as described for the primary efficacy variable, unless otherwise specified in Section10.2 Appendix 2.
- Raw value and change from baseline variables at each visit (Baseline, Weeks 4, 8, 12, 12/ET, Week 28/ET) will be analyzed using the ANCOVA model with fixed effects for treatment, and baseline value as a covariate.
- Time to the first response analysis will be analyzed similarly as described in 7.4.1.5.
  - ACR50-CRP
  - ACR70-CRP
  - DAS28-CRP <2.6
  - SDAI score <= 3.3
  - CDAI score <= 2.8
  - ACR/EULAR score for remission

  As for DAS28-CRP < 2.6, CDAI score <= 2.8, SDAI score <= 3.3, ACR/EULAR score for remission, if the event is occurred at baseline, then that subject is excluded from this analysis.
- Graphical analysis (including Baseline, Weeks 4, 8, 12, 12/ET, Week 28/ET).
  - ACR20/50/70-CRP will be plotted for over time.
  - The mean-standard deviation plot of actual values and changes from baseline will be presented for DAS28-CRP and DAS28-ESR.
  - To visually explore possible relationships between ACR20-CRP and DAS28-CRP, 2-panel line graphs of DAS28-CRP over time will be plotted for ACR20-CRP responders (in 1 panel) and non-responders (in another panel). Similar graphs will be provided for DAS28-ESR.

<Treatment Sequence Set>

- Categorical variables at each visit (including Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT) will be summarized as described in Section10.2 Appendix 2.
- Raw value and change from baseline variables at each visit (including Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT) will be summarized by descriptive statistics.
- Graphical analysis (including Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT).
  - ACR20/50/70-CRP will be plotted for over time.
  - The mean-standard deviation plot of actual values and changes from baseline will be presented for DAS28-CRP and DAS28-ESR.

<mTSS, Erosion Score, and Joint Space Narrowing Score>

In this analysis, mTSS, Erosion Score, and Joint Space Narrowing Score will be imputed using linear extrapolations as described in the primary analysis.

- Categorical variables at each visit (including Week 28/ET, Week 52/ET) will be analyzed using Fisher's exact test.
  - mTSS change from baseline >= 3
  - mTSS change from baseline <= 0.5
  - mTSS change from baseline <= 0
- Yearly progression of mTSS will be analyzed

- Percentage of subjects showing Rapid Radiographic Progression: RRP (Yearly progression of mTSS >= 5)

- Raw value and change from baseline variables at each visit (including Week 28/ET, EOT) will be analyzed using the RANCOVA model with treatment group (Placebo, ASP015K 100 mg, ASP015K 150 mg) as factor and baseline rank mTSS as covariate.
  - mTSS (Week 52/ET)
  - Erosion Score (Week 28/ET, Week 52/ET)
  - Joint Space Narrowing Score (Week 28/ET, Week 52/ET)
- Graphical analysis (including Week 28/ET, Week 52/ET).
  - Cumulative probability plot for the change from baseline in mTSS score, Erosion Score, Joint Space Narrowing Score will be presented at Week 28/ET, Week 52/ET.
  - Scatter plot for mTSS score at baseline to mTSS at Week 28/ET, Week 52/ET will be presented.
  - Scatter plot for mTSS score at baseline to mTSS change from baseline at Week 28/ET, Week 52/ET will be presented.

7.5 Analysis of Safety

In general, For AE Analysis up to Week 12 will be performed by Initial Randomization Set in SAF. AE Analysis from "Week 12 to the Last Evaluation" will be performed by Treatment Sequence Set in SAF. AE 100 Patient-Years analysis will be performed by Groups for Patient-Years. Details for analysis groups will be described in 10.3 Appendix3.

For other safety analysis except for AE up to Week 12 will be performed by Initial Randomization Set in SAF. For other safety analysis except for AE up to Week 52 will be performed by Treatment Sequence Set in SAF. Details for analysis groups will be described in 10.3 and 10.4.

7.5.1 Adverse Events

If an AE occurs on the same date as the initial dose date, the subject will be asked to select one of the following (see the eCRF page for AE): onset before initial dose of study drug,

onset after initial dose of study drug. Any AEs occurring after initial dose of study drug through the follow-up period will be considered treatment emergent.

Adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA).

All adverse events reported during the study period will be presented in subject listings.

Only TEAEs in the study period will be analyzed, and 2 types of analyses will be performed: TEAE occurrence analysis and TEAE per 100 patient-years analysis, as described below. TEAE occurrence analysis will be displayed by whole treatment period and by the following period. Primarily, TEAE occurrence analysis for Week 0 to Week 12 will be conducted in the view of comparison with placebo.

TEAE occurrence analysis for overall treatment period will be conducted in focus for ASP015K dose groups.

1. Week 0 to Week 12
2. Week 12 to Week 28
3. Week 28 to Week 52 or Later

TEAE per 100 patient-years analysis will be displayed above category, and by Week 0 to Week 12, and overall period.

7.5.1.1 Adverse events

The coding dictionary for this study will be MedDRA/J 11.1. It will be used to summarize AEs by SOC and PT. Subjects reporting more than one AE for a given MedDRA PT will be counted only once for that term. Subjects reporting more than one AE within a SOC will be counted only once for the SOC total. Subjects reporting more than one AE will be counted only once in the overall AE total.

SOCs will be presented by descending frequency for ASP015K Total column, and PTs within SOC will be presented by decreasing frequency in the ASP015K Total.

Number and percentage of subjects with TEAE in the following AE categories will be summarized by SOC and PT for each treatment group, ASP015K Total and Total:

An overview table will include the following details:

- Number and percentage of subjects with TEAE
- Number and percentage of subjects with drug related TEAE
- Number and percentage of subjects with death
- Number and percentage of subjects with serious TEAE
- Number and percentage of subjects with drug related serious TEAE
- Number and percentage of subjects with Grade 3 or Higher in Severity TEAE
- Number and percentage of subjects with TEAE leading to study drug discontinuation

- Number and percentage of subjects with drug related TEAE leading to study drug discontinuation
- Number and percentage of subjects with serious TEAE leading to permanent discontinuation of study drug or reference drug
- Number and percentage of subjects with drug related serious TEAE leading to permanent discontinuation of study drug or reference drug

The number and percentage of subjects with TEAEs, as classified by SOC and PT will be summarized for each treatment group. Summaries will be provided for:

- TEAEs
- Drug related TEAEs, defined as any TEAE with possible, probable or missing relationship to study drug as assessed by the investigator.
  Subjects reporting more than one TEAE for a given PT will be counted only once for that term in the most related category reported.
- Serious TEAEs
- Drug related serious TEAEs
- TEAEs by severity based on National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE) Grade (Grade 1 = mild, Grade 2 = moderate, Grade 3 = severe or medically significant, Grade 4 = life threatening, Grade 5 = death related to AE)
  Subjects reporting more than one TEAE for a given PT will be counted only once for that term in the most severe category reported.
  If a subject has an AE which has a missing severity, then the subject will be counted in the severity category of "Missing" (i.e., missing severity will not be imputed).
- Drug related TEAEs by severity
- Grade 3 or higher TEAEs by severity
- TEAEs by relationship to study drug
  Subjects reporting more than one TEAE for a given PT will be counted only once for that term in the most related category reported.
  TEAE with missing relationship to study drug missing will be counted as drug related. Moreover, any TEAE with missing relationship to study drug or reference drug was counted as drug related (probable).
- TEAEs leading to death
- TEAEs leading to permanent discontinuation of study drug
- Drug related TEAEs leading to permanent discontinuation of study drug
- Serious TEAEs leading to permanent discontinuation of study drug
- Drug related serious TEAEs leading to permanent discontinuation of study drug
- TEAEs excluding serious adverse events that equal to or Exceed a threshold of 5% in Any Treatment Group
- Most common TEAEs (> 5% in any treatment group)
- Most common serious TEAEs (> 5% in any treatment group)
- The number and percentage of subjects with TEAEs by subgroups defined in section 7.8.

- The number and percentage of subjects with TEAEs of special interest.

  Following adverse events of special interest are defined in the Appendix 10.8:

  - Serious Infections
  - Malignancies
  - Herpes Zoster Related Disease (Herpes Zoster and Varicella)
  - Herpes Zoster
  - Varicella
  - Infections That Require Intravenous Anti-infectious Therapy

In addition, the subgroup analysis of Sex, Age Group, Concomitant Steroid at Baseline, and Predonisone Dose at Baseline will be conducted.

7.5.1.2 Adverse Events Per 100 Patient-Years

In order to adjust for differences in subjects' durations in the study and the potential differential dropout rates between the treatment groups, TEAEs per 100 patient-years (PYs) will be calculated for each treatment group as follows:

Definition 1:

TEAEs per 100 PYs = 100 × (Number of subjects who had at least 1 incidence / Total PYs

Total PYs will be calculated by summing individual subjects' durations.

If subjects had at least 1 predefined AE, then the duration of these subjects are summed as from initial dose up to first incidence of predefined AEs.

If subjects had no predefined AE, then the duration of these subjects are summed as from initial dose through follow-up. Summed duration will be divided by 365.25 to represent per year.

Definition 2:

TEAEs per 100 PYs = 100 × (Number of TEAEs for all subjects in the treatment group / Total PYs in the treatment group).

For number of TEAEs for all subjects in the treatment group, multiple occurrences of the same TEAE in the same subject will be counted multiple times.

Total PYs will be calculated by summing subjects' durations in study through the follow-up period. Each subject's duration will be calculated as: (last date of follow-up in the study – date of initial dose of study drug + 1) / 365.25.

Patient-Year analysis will be calculated by Placebo, 100 mg, 150 mg, 100mg+150mg, ASP015K Total, Total.

The number of TEAEs and AEs per 100 PYs will be provided by Definition 2, for all TEAEs (if analyzable), by SOC and PT for each treatment group, ASP015K Total and Total.

Patient-Year analysis for following specified AE will also be conducted by Definition 1 with 95% confidence interval.

- Serious Infections
- Malignancies
- Herpes Zoster Related Disease (Herpes Zoster and Varicella)
- Herpes Zoster
- Varicella
- Infections That Require Intravenous Anti-infectious Therapy

Note: Definition will be provided in the Appendix 8.

Confidence Interval for TEAEs per 100 patient-years will be calculated for AE of special Interest based on following formula (assuming Poisson distribution).

- Upper limit of CI (TEAEs per 100 PYs) : $100 * \exp\left(\log\left(\frac{A}{T}\right) + Z_{alpha} * \sqrt{\frac{1}{A}}\right)$
- Lower limit of CI (TEAEs per 100 PYs) : $100 * \exp\left(\log\left(\frac{A}{T}\right) - Z_{alpha} * \sqrt{\frac{1}{A}}\right)$

A: Number of subjects who had at least 1 incidence

T: Total PYs by 100 PYs as one unit

$Z_{(alpha)}$: alpha % point of standard normal distribution

alpha = 0.025

In addition, the subgroup analysis of Sex, Age Group, Concomitant Steroid at Baseline, and Predonisone Dose at Baseline will be conducted.

7.5.1.3 Subgroup Analyses

In order to assess whether or not AEs vary across specific patient groups, AE occurrence will be summarized by SOC and PT for the subgroups specified in section 7.8.

7.5.2 Clinical Laboratory Evaluation

All Laboratory Tests

For clinical laboratory parameters (hematology, chemistry, urinalysis and fast lipid profile), raw values at each scheduled visit, change from baseline at each post-baseline visit, shift from baseline and shift from reference range will be summarized using descriptive statistics or frequency tabulation. No statistical hypothesis testing will be performed. For Fasting Lipids Profile, all analysis will be conducted using the measured values which are confirmed under fasting condition. In addition, all Lipid Profile values including non-fasting condition will also be summarized for descriptive statistics.

- For continuous variables, descriptive statistics of Raw values at each scheduled visit will be summarized at each time point using SI unit. For the ratio T-Chol to HDL and the ratio LDL to HDL, descriptive statistics will be summarized at each post-baseline visit.
- For continuous variables, descriptive statistics of change from baseline value will be summarized at each post-baseline visit using SI unit.
- For categorical variables, frequencies and percentage will be displayed at each scheduled visit.
- Shift-from-baseline table: Shift from baseline to highest value up to Week 12 or during entire period, and baseline to lowest value up to Week 12 or during entire period, for all laboratory variables. If the laboratory value of same subject is measured lower from normal range at one visit, and higher from normal range at the other visit, that subject will be counted as both low and high from baseline up to Week 12 or during entire period.

Laboratory Tests for NCI-CTC Toxicity Grading

Laboratory test results for the select laboratory tests will be graded programmatically using the standardized NCI-CTC toxicity grading criteria (Grade 0, 1, 2, 3, 4) specified in the 10.8 Appendix8 for this study. Shift-from-baseline to worst value up to Week 12 or during entire period will be provided.

Note: AEs in this study are graded the investigator using NCI-CTAE Grade (1, 2, 3, 4, 5).

Laboratory Tests for Reference Range

Test results for the select laboratory tests in Table 3 will be summarized against the reference range indicated. Number and percentage for the Highest/Lowest values of each subjects included in pre-specified reference range for select laboratory up to Week 12 or during entire period, will be provided.

Table 3. Select Laboratory Variables and Corresponding Reference Ranges

Select Laboratory Variable	Reference Range	Highest/Lowest
Aspartate Aminotransferase (AST)	>= 1 × ULN to < 2 × ULN >= 2 × ULN to < 3 × ULN>= 3 × ULN	Highest
Alanine Aminotransferase (ALT)	>= 1 × ULN to < 2 × ULN >= 2 × ULN to < 3 × ULN >= 3 × ULN	Highest
Alkaline Phosphatase (ALP)	> 2 × ULN to <= 3 × ULN > 3 × ULN to <= 5 × ULN > 5 × ULN	Highest
Total Bilirubin	> 1.5× ULN	Highest
Total Bilirubin	>= 1 × ULN to < 2 × ULN >= 2 × ULN to < 3 × ULN >= 3 × ULN	Highest
Low-Density Lipoprotein (LDL)	> 160 mg/dL	Highest
Low-Density Lipoprotein (LDL)	< 100 mg/dL 100 to < 130 mg/dL 130 to < 160 mg/dL 160 to < 190 mg/dL >= 190 mg/dL	Highest
Hemoglobin (HGB)	>=8.0 to < 10.0 g/dL < 8.0 g/dL	Lowest
Hemoglobin (HGB)	>= (Baseline-2) to <(Baseline-1) g/dL <(Baseline-2) g/dL	Lowest
Hemoglobin (HGB)	• Mild to Moderate: decrease from baseline >=1 to <=2 g/dL, • Severe: decrease from baseline >2 to <3 g/dL or absolute value >7 and <8 g/dL, • Potentially Life Threatening: decrease from baseline >=3 or absolute value <=7 g/dL	Lowest
Creatine Phosphokinase (CPK)	> 500 to <= 2,000 U/L > 2,000 U/L	Highest
Creatine Phosphokinase (CPK)	> 5 × Baseline	Highest
Creatine Phosphokinase (CPK)	> 2× ULN to <= 5 × ULN > 5 × ULN to <= 10 × ULN > 10 × ULN	Highest

Select Laboratory Variable	Reference Range	Highest/Lowest
Creatinine	> 1.5 × Baseline to <= 3.0 × Baseline > 3.0 × Baseline	Highest
Absolute Neutrophil Count (ANC)	>= 1,500 to < 2,000 / mm^3 >= 500 to < 1,500 / mm^3 < 500 / mm^3	Lowest
Lymphocytes	>= 200 to < 500 /uL < 200 /uL	Lowest
Lymphocytes	>= 1,500 to < 2,000 / mm^3 >= 500 to < 1,500 / mm^3 < 500 / mm^3	Lowest
Platelets	>= 2 × 10^4 to < 5 × 10^4 /uL < 2 × 10^4 /uL	Lowest
Platelets	> 600,000 /uL	Highest
ULN = Upper Limit Normal Each categories are mutually exclusive.		

7.5.2.1 Liver Abnormalities

The following potentially clinically significant criteria in liver function tests for Alkaline Phosphatase (ALP), Alanine Transaminase (ALT), total bilirubin, Aspartate Transaminase (AST) and their combination are defined. The subject's highest value during the investigational period will be used.

Parameter	Criteria
ALT	> 3xULN
	> 5xULN
	> 10xULN
	> 20xULN
AST	> 3xULN
	> 5xULN
	> 10xULN
	> 20xULN
ALT or AST	> 3xULN
	> 8xULN
Total Bilirubin	> 2xULN
ALP	> 1.5xULN

ALT and/or AST OR Total Bilirubin (*)	(ALT and/or AST > 3xULN) or total bilirubin > 2xULN
ALT and/or AST AND Total Bilirubin (*)	(ALT and/or AST > 3xULN) and total bilirubin > 2xULN

(*) Combination of values measured within same sample

The number and percentage of subjects with potentially clinically significant values in liver function tests during the investigational period will be presented by treatment group.

Additionally, confirmed liver abnormities will be characterized as moderate and marked, as follow:

Moderate: ALT or AST > 3×ULN OR Total Bilirubin > 2×ULN.

Marked: ALT or AST > 3×ULN AND Total Bilirubin > 2×ULN.

Number and percentage of subjects in each category at each scheduled visit will be summarized. These combinations of elevated values are based on measured value within the same sample.

Liver Functions Plots

A matrix scatter plot of liver function test will be plotted showing the maximum ALT, AST, ALP and total bilirubin during study period crossed against each other. Different dots will be used for randomization arm.

7.5.3 Vital Signs

Raw values and changes from baseline at each scheduled visit will be summarized using descriptive statistics.

7.5.4 Electrocardiograms (ECGs)

<12-lead ECG>

- Raw values (categorical) and changes from baseline (shift from baseline) will be summarized, and the summaries will include number and percentage of subjects with normal, not clinically significant abnormal, and clinically significant abnormal 12 lead ECG findings.

< Cetnral electrocardiogram >

QTc intervals should be calculated by followed procedure:

- Bazett's formula: $QTc = QT/(RR/1000)^{1/2}$
- Fridericia's formula: $QTc = QT/(RR/1000)^{1/3}$

Change for ECG variables is defined in two ways.

- baseline: before the first dose of study drug
- predose : before the Week 4 / Week 8 dose of study drug

- Descriptive statistics of the central ECG variables (QTcF, QTcB, QT, PR, RR intervals,QRS and HR) and the corresponding changes from baseline by each visit and time of measurement will be calculated. In addition, the corresponding changes from pre-dose to post-dose at Week 4 / Week 8 will be calculated.
- The QT, QTc interval variables will be summarized by the frequencies of subjects with following categories up to Week 12, and each time points.
  - < 300 msec
  - < 330 msec
  - < 360 msec
  - > 450 msec
  - > 480 msec
  - > 500 msec
- The change from baseline of QT, QTc interval variables will be summarized by the frequencies of subjects with following categories up to Week 12, and each time points. In addition, the corresponding changes from pre-dose to post-dose at Week 4 / Week 8 will be calculated.
  - > 30 msec increase
  - > 60 msec increase
  - > 30 msec decrease
  - > 60 msec decrease

7.5.5 Chest radiography

Chest radiography will be provided as a subject listing only.

7.5.6 Pregnancies

Pregnancies will be provided as a subject listing only.

7.6 Analysis of PK

Refer to PK SAP.

7.7 Analysis of PD

Following analysis is executed for PDAS.

Vascular Endothelial Growth Factor (VEGF), Matrix metalloproteinase 3 (MMP-3) , following Lymphocytes subset variables.

- CD3+/Lymphocytes (%), CD3+ (cells/uL)
- CD8+/Lymphocytes (%), CD8+ (cells/uL)
- CD4+/Lymphocytes (%), CD4+ (cells/uL)
- CD19+/Lymphocytes (%), CD19+ (cells/uL)
- CD4/CD8 Ratio
- (CD16 or CD56) + /Lymphocytes (%), (CD16 or CD56)+ (cells/uL)
- (CD16+ and CD56+)/CD3- (%), CD16+ and CD56+ (cells/uL)
- (CD16- and CD56+)/CD3- (%), CD16- and CD56+ (cells/uL)
- (CD16+ and CD56-)/CD3- (%), CD16+ and CD56- (cells/uL)
- (CD16- and CD56-)/CD3- (%), CD16- and CD56- (cells/uL)
- CD56 bright/CD3- (%), CD56 bright (cells/uL)
- CD56 dim/CD3- (%), CD56 dim (cells/uL)

- Raw values and changes from baseline will be summarized using descriptive statistics.

7.8 Subgroups of Interest

Primary efficacy endpoint and selected safety variables (treatment emergent adverse events) will be summarized by the treatment group for the subgroups defined on the basis of the categorized variables listed below:

- Sex (Male, Female)
- Age Group (< 65 years, >= 65 years)
- Prior Anti-TNF DMARD Use (User, Non-user)
- Prior Biologic DMARD Use (User, Non-user)
- Duration of RA (years) (< 5 years, >= 5 years)
- Number of Prior DMARDs Used (0, 1, 2, >= 3)
- Number of Prior Biologic DMARDs Used (0, 1, 2, >= 3)
- MTX Dose (mg/week) at Baseline (0 <- 8 mg/week, 8 <- 12 mg/week, > 12 mg/week)
- Concomitant Steroid at Baseline (No, Yes)
- Prednisone Dose at Baseline (mg/day) (None, 0 <- 5 mg/day, > 5 mg/day)
- Baseline DAS28-CRP (<= 3.2, 3.2 <- 5.1, > 5.1) *
- Baseline DAS28-ESR (<= 3.2, 3.2 <- 5.1, > 5.1) *
- Baseline mTSS (<= Median, > Median) *
- Baseline CRP (< 1.0, >= 1.0) *
- Baseline Anti-CCP Antibody (Low-positive or negative (<- 3 x ULN), High-positive (> 3 x ULN)) *
- Baseline Rheumatoid Factor (Low-positive or negative (<- 3 x ULN), High-positive (> 3 x ULN)) *
- mTSS at Week 28 Imputation Status (Yes, No)**
- Body Weight (kg) at Screening (<- 40 kg, 40 kg <- 60 kg, 60 kg <- 80 kg, > 80 kg)

*: Only applied to efficacy analysis.
**: Only applied to mTSS subgroup analysis. If mTSS at Week 28 does not exists and Week 28/ET is linear extrapolated from previous data, or switched to receive ASP015K instead of placebo at Week 12 due to the lack of efficacy, those subjects are regarded as "Yes". If mTSS at Week 28 exists and Week 28/ET is not imputed by linear extrapolation, then those subjects are regarded as “No”.).

For more details refer to section 7.4.1.3, 7.5.1.3.

7.9 Other Analyses

Not Applicable.

7.10 Interim Analysis (and Early Discontinuation of the Clinical Study)

Not applicable for efficacy analysis. Review of safety data and safety evaluation will be completed by an independent DSMB during the study in accordance with separate SOP.

7.11 Handling of Missing Data, Outliers, Visit Windows, and Other Information

7.11.1 Missing Data

PD variables

Values below the LLOQ will be set to 0 for calculation of descriptive statistics. As defined in 7.11.6, Week 12/ET, Week 28/ET, EOT are analysis visit using LOCF method.

Missing Data in ACR20/50/70, DAS28

Table 4 addresses the handling of missing ACR components and ACR response in the inferential analyses of ACR20/50/70, DAS28, and DAS28-related variables at Week12/ET. Note that if subject does not have any post-baseline values, then that subject will be defined as an ACR20-CRP non-responder. Moreover, if subject does not have baseline values, then that subject will be defined as an ACR20-CRP non-responder.

For ACR20/50/70, if some of ACR components include missing, then the following handling will be done.

	TJC-68	SJC-66	Remaining 5 Components	ACR20/50/70
Pattern 1	Non-Responder	Any	Any	Non-Responder
Pattern 2	Any	Non-Responder	Any	Non-Responder
Pattern 3	Any	Any	The Number of the Response is Less than 3 Components.	Non-Responder
Pattern 4	Missing	Responder or Missing	The Number of the Response is at least 3 Components, or the Number of the Missing is at least 3 Components.	Missing
Pattern 5	Responder or Missing	Missing	The Number of the Response is at least 3 Components, or the Number of the Missing is at least 3 Components.	Missing
Pattern 6	Responder or Missing	Responder or Missing	The Number of the Missing is at least 3 Components (Baseline=0 regarded as Non-Responder).	Missing
Pattern 7	Responder	Responder	The Number of the Response is at least 3 Components.	Responder

Table4. Imputation Methods for Missing ACR and DAS28 for Components and Response at Week 12/ET

Imputation Method		Explanation/Instruction
For use in all ACR20/50/70 analyses at Week 12/ET including primary analysis.		
1	LOCF components	● First LOCF all missing ACR component value(s) at Week 12 and then calculate ACR response as Week 12/ET.
For use in sensitivity analysis of ACR20-CRP at Week 12/ET		
2	LOCF components and NRI response	● If all ACR component values are missing at Week 12 for any reason and therefore the ACR response is missing, the missing ACR response at Week 12/ET will be imputed using NRI (i.e., subject is a non-responder). ● If NOT all ACR component values are missing at Week 12 for any reason, all missing ACR component value(s) at Week 12/ET will first be imputed using LOCF and then the ACR response will be calculated.
3	Observed	● Data as observed (i.e., data as reported in the study database) will be analyzed. Missing ACR response components will remain missing if reported in the study database as missing. Missing ACR response will remain missing. *: ACR20 at Week 12 is used for this sensitivity analysis

Imputation Method		Explanation/Instruction
4	Multiple Imputation, assuming Missing at Random Mechanism	Multiple imputation (MI) will be conducted in the following procedure. <Step1>Imputation Part 1) Multiple Imputation for Non-Monotone Missing data If non-monotone missing pattern exists, then those non-monotone missing data will be imputed by adapting Markov Chain Monte Carlo method, using each non-missing ACR components data up to Week 12. Imputation will be conducted by each ACR components (i.e. TJC-68, SJC-66, SGAP, SGA, PGA, HAQ-DI, CRP), and treatment groups, respectively. 2) Multiple Imputation for Monotone Missing data Completing the imputation for non-monotone missing data, each ACR components (i.e. TJC-68, SJC-66, SGAP, SGA, PGA, HAQ-DI, CRP) at Week 12 will be imputed, respectively, based on the regression model with treatment group (Placebo, ASP015K 100 mg, ASP015K 150 mg) as factor and all values up to Week 12 as covariate. In this method, using seed as 12345, 1000 imputed datasets for each ACR components will be created based on SAS ver. 9.4 PROC MI procedure, monotone reg statement. Then, ACR20 response at Week 12 will be calculated with imputed datasets for each ACR components based on the same method described in Section 6.1.1. <Step2>Analysis Part Based on 1000 Imputed datasets created at Step1, ACR20 at Week 12 in each treatment group will be summarized by Imputed datasets. Also, the difference in ACR20 at Week 12 between ASP015K 100 mg / 150 mg and placebo and their standard error will be calculated. Moreover, the log odds ratio to placebo in ACR20 at Week 12 and its standard error will be calculated by each Imputed datasets using logistic regression model with treatment group as factor. <Step3>Combine Part In accordance with Rubin's rule (Little and Rubin, 2002), the 1000 estimated ACR20 response rate in each treatment groups are combined. In addition, the difference between ASP015K 100 mg and placebo, the difference between ASP015K 150 mg and 95% confidence interval for the difference will be combined. Moreover, the Wald's chi-squared test using logistic regression model with treatment group (Placebo, ASP015K 100 mg, ASP015K 150 mg) as factor will also be calculated.

Imputation Method		Explanation/Instruction
5	Placebo Multiple Imputation (pMI)	Placebo Multiple Imputation, assuming the statistical behavior of ASP015K 100 mg or 150 mg treated patients who discontinue becomes that of placebo-treated patients after the time of dropout will be conducted in the following procedure. <Step1>Imputation Part 1) Multiple Imputation for Non-Monotone Missing data If non-monotone missing pattern exists, then those non-monotone missing data will be imputed by adapting Markov Chain Monte Carlo method, using each non-missing ACR components data up to Week 12. Imputation will be conducted by each ACR components (i.e. TJC-68, SJC-66, SGAP, SGA, PGA, HAQ-DI, CRP), and treatment groups, respectively. 2) Multiple Imputation for Monotone Missing data Completing the imputation for non-monotone missing data, each ACR components (i.e. TJC-68, SJC-66, SGAP, SGA, PGA, HAQ-DI, CRP) at Week 12 will be imputed, respectively, based on the regression model using only data from the placebo group up to Week 12. In this method, using seed as 12345, 1000 imputed datasets for each ACR components will be created based on SAS ver. 9.4 PROC MI procedure, monotone reg statement. Then, ACR20 response at Week 12 will be calculated with imputed datasets for each ACR components based on the same method described in Section 6.1.1. <Step2>Analysis Part For 1000 Imputed datasets created at Step1, ACR20 at Week 12 in each treatment group will be calculated. Also, the difference in ACR20 at Week 12 between ASP015K 100 mg / 150 mg and placebo and their standard error will be calculated. Moreover, the log odds ratio to placebo in ACR20 at Week 12 and its standard error will be calculated by each Imputed datasets using logistic regression model with treatment group as factor. <Step3>Combine Part In accordance with Rubin's rule (Little and Rubin, 2002), the 1000 estimated ACR20 response rate in each treatment groups are combined. In addition, the difference between ASP015K 100 mg and placebo, the difference between ASP015K 150 mg and 95% confidence interval for the difference will be combined. Moreover, the Wald's chi-squared test using logistic regression model with treatment group (Placebo, ASP015K 100 mg, ASP015K 150 mg) as factor will also be calculated.

Imputation Method		Explanation/Instruction
For use in all DAS28 and DAS28-based analyses, unless otherwise specified in Section 10.2 Appendix 1		
6	LOCF DAS28 ACR components	• LOCF all missing DAS28 ACR component value(s) and then calculate the DAS28 score.

Missing Data in mTSS

For mTSS, Erosion Score, Joint Space Narrowing Score, will be imputed by linear extrapolation at Week 28 evaluation (Day197) described as Week 28/ET. Namely, for subjects who discontinued at or before Week 28 (after Day48) or for initially placebo treated subjects who switched to receive ASP015K at Week 12 due to the lack of efficacy, mTSS at Week 28 (Day 197) will be extrapolated using linear extrapolation method based on the mTSS at baseline and early termination (after Day48) or Week 12 (Day85) (before switching). For subjects who initially treated placebo, if the subjects have the measurement data at Week 28 (from Day 160 to Day 234) and not switched to ASP015K yet, then that data will be used (No imputation will be done). For Week 52 (Day365) evaluation, described as Week 52/ET, the value will be extrapolated as well (at baseline and early termination (after Day48) or Week 12/Week 28 (before switching for placebo treated subjects). If the subjects have the measurement data at Week 52 (from Day 328 to Day 402) and not initially randomized to placebo, then that data will be used (No imputation will be done).

For use in sensitivity analysis, mTSS, Erosion Score, Joint Space Narrowing Score will be imputed using LOCF. Moreover, Observed analysis will also be conducted.

In addition, Multiple Imputation for sensitivity analysis will be conducted in the following procedure.

1) mTSS at Week 28 will be imputed based on the regression model treatment group as factor and all mTSS values up to Week 28 as covariate. For subjects who discontinued at or before Week 28 or switched to receive ASP015K instead of placebo at Week 12 due to the lack of efficacy, mTSS data will be applied before switching. In this method, using seed as 12345, 1000 Imputed datasets will be created based on SAS ver. 9.4 PROC MI procedure, monotone regression statement.
2) For 1000 Imputed datasets created at Step1, Analysis of Covariance (ANCOVA) with treatment group (Placebo, ASP015K 100 mg, ASP015K 150 mg) and as factor and baseline mTSS as covariate will be conducted by each Imputed datasets and the lsmeans difference between ASP015K each group and placebo will be put.
3) In accordance with Rubin's rule (Little and Rubin, 2002), the 1000 estimated difference value and standard error are combined, and 95% confidence interval will be calculated. Moreover, statistical test will be conducted.

Missing Data in HAQ-DI

For HAQ-DI, there are 20 items in 8 categories, with each category having 2 or 3 questions. A subject must have a score in ≥ 6 of the 8 categories, otherwise the HAQ-DI cannot be computed and will be considered missing for data summarization and analysis purposes. If responses to individual questions within a category are missing, they are not imputed. Therefore, the score in each category is based on non-missing responses. A category score is missing when all responses within a category are missing.

Missing Data in SF-36v2

For SF-36v2, there are 8 scales (domains). Missing values will be imputed if at least half of the items in the domain that included the missing item score(s) are non-missing. In that case, the missing value will be imputed with the average score of the non-missing item scores in this domain.

Missing Data in WPAI

For WPAI Scale Score, each score will be computed on condition that the all of the questions for each score will not be non-missing.

Safety Data Imputation Methods

For Safety data, as defined in 7.11.5, Week 12/ET is analysis visit using LOCF method.

7.11.2 Outliers

All values will be included in the analyses.

7.11.3 Visit Windows

The acceptable time ranges of the efficacy and safety examinations, observations, etc. from the date of initial study treatment (Day 1) are defined as follows. If there are multiple data available over the same period of time, the data obtained on the date closest to the reference date will be utilized; the later date will be applied if the number of days from the reference date is equal. Subjects who have no evaluable post-dose data on ACR20, ACR50, and ACR70 response rates will be counted as non-responders and included in the analysis at the end of Week 12 (Week 12 or at early termination before Week 12).

7.11.4 Handling of Schedule of Assessments of Efficacy Variables

For data after the end of study treatment, data obtained within +2 days after the last dose is taken will be included in analysis. Baseline for efficacy is defined as value at Day 1 before the first dose of study drug for all subjects including placebo assigned subjects. For Week

12/ET, or Week 28/ET, the values at closest date to reference date will be used. As for the subjects who switched to receive ASP015K instead of placebo at Week 12, the last observation before first dose of ASP015K at Week 12 will be used for the analysis for visit Week 4, 8, 12, 12/ET. In addition, the subjects who switched to receive ASP015K instead of placebo at Week 28, the last observation before the first dose of ASP015K at Week 28 will be used for the analysis for visit Week 4, 8, 12, 12/ET, 28/ET. For EOT, the closest value to the reference date will be used regardless of the switching from placebo.

(1) TJC (68 joints)/SJC (66 joints), laboratory test (CRP)

Table5. Visit Window for TJC (68 joints)/SJC (66 joints), laboratory test (CRP)

Time points defined in analysis	Reference date*	Acceptable time range
Screening**	Day −28 to Day −1	Day −28 to Day −1
Baseline	Day 1	Day 1
Week 4 to Week 48***	Day 7×x + 1	Day (7×x − 6) to Day (7×x + 8) and before drug administration on the visit day
Week 52	Day 365	Day 358 to Day 372
Week 12/ET (Week 12 or at early termination before Week 12)	Day 85	Day 2 to Day 92
Week 28/ET (Week 28 or at early termination before Week 28)	Day 197	Day 2 to Day 204
EOT (Week 52 or at early termination)	Day 365	Day 2 to Day 372

*: Day 1 represents the first day of study treatment. The day before is expressed as Day −1.

**: PGA, SGA, subject's assessment of pain, and CRP are not applicable.

***: Represents the reference date and acceptable time range for Week x.

(2) Hand and foot radiography

Table6. Visit Window for Hand and foot radiography

Time points defined in analysis	Reference date*	Acceptable time range
Baseline	Day 1	Day 1 to Day 30
Week 28	Day 197	Day 160 to Day 234
Week 52	Day 365	Day 328 to Day 402
Week 28/ET (Week 28 or at early termination before Week 28)	Day 197	Day 48 to Day 234
EOT (Week 52 or at early termination)	Day 365	Day 48 to Day 402

*: Day 1 represents the first day of study treatment. The day before is expressed as Day −1.

(3) PGA and SGA (VAS), subject's assessment of pain (VAS), HAQ-DI, laboratory test (ESR)

Table7. Visit Window for PGA and SGA (VAS), subject's assessment of pain (VAS), HAQ-DI, laboratory test (ESR)

Time points defined in analysis	Reference date*	Acceptable time range
Baseline	Day 1	Day 1
Week 4 to Week 48**	Day 7×x + 1	Day (7×x − 6) to Day (7×x + 8) and before drug administration on the visit day
Week 52	Day 365	Day 358 to Day 372
Week 12/ET (Week 12 or at early termination before Week 12)	Day 85	Day 2 to Day 92
Week 28/ET (Week 28 or at early termination before Week 28)	Day 197	Day 2 to Day 204
EOT (Week 52 or at early termination)	Day 365	Day 2 to Day 372

*: Day 1 represents the first day of study treatment. The day before is expressed as Day −1.

**: Represents the reference date and acceptable time range for Week x.

(4) SF-36 v2®, WPAI

Table8. Visit Window for SF-36 v2®, WPAI

Time points defined in analysis	Reference date*	Acceptable time range
Baseline	Day 1	Day 1
Week 4 to Week 28**	Day 7×x + 1	Day (7×x − 6) to Day (7×x + 8) and before drug administration on the visit day
Week 52	Day 365	Day 358 to Day 372
Week 12/ET (Week 12 or at early termination before Week 12)	Day 85	Day 2 to Day 92
Week 28/ET (Week 28 or at early termination before Week 28)	Day 197	Day 2 to Day 204
EOT (Week 52 or at early termination)	Day 365	Day 2 to Day 372

*: Day 1 represents the first day of study treatment. The day before is expressed as Day −1.

**: Represents the reference date and acceptable time range for Week x.

(5) QT assessment

Table9. Visit Window for QT assessment

Time points defined in analysis	Reference date	Acceptable time range
Baseline	Day 1	Day 1 Before the initiation of treatment with the study drug
Week 4 or Week 8 (before study drug administration)	Day 29 or Day 57	Before the initiation of treatment with the study drug on Reference date ± 7
Week 4 or Week 8 (after study drug administration)	Day 29 or Day 57	2 hours post-dose (reference time, but within the range of 1 hour to 4 hours post-dose is acceptable) on Reference date ± 7
Minimum on Treatment up to Week 12		Minimum value at allowance Range (Day 2 to Day 92) before first dose of Week 12 study drug
Maximum on Treatment up to Week 12		Maximum value at allowance Range (Day 2 to Day 92) before first dose of Week 12 study drug

7.11.5 Handling of Schedule of Assessments of Safety Variables

For data obtained after the end of study treatment, the data obtained within + 2 days after the last dose is taken will be included in analysis. Baseline for safety is defined as last non-missing value before the first dose of study drug for all subjects including placebo assigned subjects. For Week 12/ET, or Week 28/ET, the values at closest date to reference date will be used. As for the subjects who switched to receive ASP015K instead of placebo at Week 12, the last observation before first dose of ASP015K at Week 12 will be used for the analysis for visit Week 4, 8, 12, 12/ET. In addition, the subjects who switched to receive ASP015K instead of placebo at Week 28, the last observation before the first dose of ASP015K at Week 28 will be used for the analysis for visit Week 4, 8, 12, 12/ET, 28/ET. For EOT, the closest value to the reference date will be used regardless of the switching from placebo.

(1) Vital signs, laboratory test (hematology, biochemistry, urinalysis)

Table10. Visit Window for Vital signs, laboratory test (hematology, biochemistry, urinalysis)

Time points defined in analysis	Reference date*	Acceptable time range
Screening	Day −28 to Day −1	Day −28 to Day −1
Baseline	Day 1	Day 1
Week 4 to Week 48**	Day 7×x + 1	Day (7×x − 6) to Day (7×x + 8) and before drug administration on the visit day
Week 52	Day 365	Day 358 to Day 372
Minimum on Treatment up to Week 12***		Minimum value at allowance Range (Day 2 to Day 92) before first dose of Week 12 study drug
Maximum on Treatment up to Week 12***		Maximum value at allowance Range (Day 2 to Day 92) before first dose of Week 12 study drug
Minimum on Treatment***		Minimum value during Entire Period (No limitation for period)
Maximum on Treatment***		Maximum value during Entire Period (No limitation for period)
Week 12/ET (Week 12 or at early termination before Week 12)	Day 85	Day 2 to Day 92
Week 28/ET (Week 28 or at early termination before Week 28)	Day 197	Day 2 to Day 204
EOT (Week 52 or at early termination)	Day 365	Day 2 to Day 372

*: Day 1 represents the first day of study treatment. The day before is expressed as Day −1.

**: Represents the reference date and acceptable time range for Week x.

***: Applied to laboratory test only.

(2) Laboratory test (fasting lipid profile test)

Table11. Visit Window for Laboratory test (fasting lipid profile test)

Time points defined in analysis	Reference date*	Acceptable time range
Baseline	Day 1	Before drug administration on Day 1 and 8 hours after the last meal If Baseline data is not measured at Day 1, then the closest day up to Day 1 after Day −28 will be used for Baseline.
Week 12	Day 85	Day 78 to Day 92
Week 20	Day 141	Day 134 to Day 148
Week 28	Day 197	Day 190 to Day 204
Week 40	Day 281	Day 274 to Day 288
Week 52	Day 365	Day 358 to Day 372
Minimum on Treatment up to Week 12		Minimum value at allowance Range (Day 2 to Day 92) before first dose of Week 12 study drug
Maximum on Treatment up to Week 12		Maximum value at allowance Range (Day 2 to Day 92) before first dose of Week 12 study drug
Minimum on Treatment		Minimum value during Entire Period (No limitation for period)
Maximum on Treatment		Maximum value during Entire Period (No limitation for period)
Week 12/ET (Week 12 or at early termination before Week 12)	Day 85	Day 2 to Day 92
Week 28/ET (Week 28 or at early termination before Week 28)	Day 197	Day 2 to Day 204
EOT (Week 52 or at early termination)	Day 365	Day 2 to Day 372

*: Day 1 represents the first day of study treatment.

**: Represents the reference date and acceptable time range for Week x.

(3) 12-lead ECG, Chest radiography

Table12. Visit Window for 12-lead ECG

Time points defined in analysis	Reference date*	Acceptable time range
Screening	Day −28 to Day −1	Day −28 to Day −1
Week 28	Day 197	Day 190 to Day 204
Week 52	Day 365	Day 358 to Day 372
Week 28/ET (Week 28 or at early termination before Week 28)	Day 197	Day 2 to Day 204
EOT (Week 52 or at early termination)	Day 365	Day 2 to Day 372

*: Day 1 represents the first day of study treatment. The day before is expressed as Day −1.

(4) Body weight

Table13. Visit Window for Body weight

Time points defined in analysis	Reference date*	Acceptable time range
Screening	Day −28 to Day −1	Day −28 to Day −1
Week 12	Day 85	Day 78 to Day 92
Week 28	Day 197	Day 190 to Day 204
Week 40	Day 281	Day 274 to Day 288
Week 52	Day 365	Day 358 to Day 372
Week 12/ET (Week 12 or at early termination before Week 12)	Day 85	Day 2 to Day 92
Week 28/ET (Week 28 or at early termination before Week 28)	Day 197	Day 2 to Day 204
EOT (Week 52 or at early termination)	Day 365	Day 2 to Day 372

*: Day 1 represents the first day of study treatment. The day before is expressed as Day −1.

7.11.6 Handling of Schedule of Assessments of Pharmacodynamic Variables

For data obtained after the end of study treatment, the data obtained within + 2 days after the last dose is taken will be included in analysis. Baseline for safety is defined as last non-missing value before the first dose of study drug for all subjects including placebo assigned subjects. For Week 12/ET, or Week 28/ET, the values at closest date to reference date will be used. As for the subjects who switched to receive ASP015K instead of placebo at Week 12, the last observation before first dose of ASP015K at Week 12 will be used for the analysis for visit Week 4, 8, 12, 12/ET. In addition, the subjects who switched to receive ASP015K instead of placebo at Week 28, the last observation before the first dose of ASP015K at Week 28 will be used for the analysis for visit Week 4, 8, 12, 12/ET, 28/ET. For EOT, the closest value to the reference date will be used regardless of the switching from placebo.

Table14. Visit Window for Pharmacodynamic Variables

Time points defined in analysis	Reference date*	Acceptable time range
Baseline	Day 1	Before drug administration on Day 1
Week 4 to Week 28**	Day 7×x + 1	Within ± 7 days of the reference date and before drug administration on the visit day
Week 52	Day 365	Day 358 to Day 372
Week 12/ET (Week 12 or at early termination before Week 12)	Day 85	Day 2 to Day 92
Week 28/ET (Week 28 or at early termination before Week 28)	Day 197	Day 2 to Day 204
EOT (Week 52 or at early termination)	Day 365	Day 2 to Day 372

*: Day 1 represents the first day of study treatment. The day before is expressed as Day −1.

**: Represents the reference date and acceptable time range for Week x.

8 DOCUMENT REVISION HISTORY

Version	Date	Changes	Comment/rationale for change
1.00	18-Jan-2018	Final Version	Document finalized

9 REFERENCES

ICH Harmonized Tripartite Guideline E 3. Structure and Content of Clinical Study Reports, November 1995. (www.ich.org; Guidelines; "Efficacy" Topics)

ICH Harmonized Tripartite Guideline E 9. Statistical Principles for Clinical Trials, February 1998. (www.ich.org; Guidelines; "Efficacy" Topics)

Xeljanz® Tablets, 5 mg CTD, 2.7 Clinical Summary, 2.7.3 Summary of Clinical Efficacy, 2.7.3.3 Comparison of Results in Subpopulations, 2.7.3.3.1 Comparison of Japanese Population and Whole Population in Global Studies (Study No. A3921044) 123-139 (2013)

Cimzia Subcutaneous Injectin 200 mg Syringe CTD, 2.7 Clinical Summary, 2.7.3 Summary of Clinical Efficacy, 2.7.3.2 Summary of Results of Individual Studies, 2.7.3.2.2 Comparison of Efficacy Results of All Studies, 2.7.3.2.2.1 Japanese Studies (1) Study CDP870-041, (2) Study RA0006 22-35 (2012)

Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. PharmacoEconomics 1993; 4(5):353-65.

Reilly Associates Health Outcomes Research, http://www.reillyassociates.net.

10 APPENDICES

10.1 Appendix 1: Overview of ACR20 response at Week 12/ET, mTSS change from baseline at Week 28/ET analysis

Table 14 summarizes the analyses of the **ACR20 response at Week 12/ET** and **mTSS change from baseline at Week 28/ET**, with "X" indicating that the variable will be analyzed. All analyses will be based on the FAS, unless otherwise specified.

Table 15. Overview of ACR20 response at Week 12/ET, mTSS change from baseline at Week 28/ET analysis

Analysis Type	Analysis Variable		Analysis Method
Primary / Sensitivity / Subgroup	**Primary: ACR20-CRP response at Week 12/ET**	**Pimary mTSS change from baseline at Week 28/ET**	
Primary	X[1]	X[2]	ACR20: Fisher's exact test. mTSS: RANCOVA, Cumulative probability plot
Sensitivity[3] (see Sections 7.4.1.2 and 7.11.1 for further explanation)			
LOCF		X[2]	mTSS: RANCOVA
LOCF and NRI	X		ACR20: Fisher's exact test.
Observed data	X	X[2]	ACR20: Fisher's exact test. mTSS: RANCOVA
PPS as analysis set	X	X[2]	ACR20: Fisher's exact test (LOCF) mTSS: RANCOVA
Analysis of Covariance		X	mTSS: ANCOVA (LEP)
Multiple Imputation Assuming Missing at Random	X	X	ACR20: Fisher's exact test. mTSS: ANCOVA
Placebo Multiple Imputation (pMI)	X		ACR20: Fisher's exact test.
Subgroup [4]			
Sex	X	X	ACR20: LOCF mTSS : LEP
Age Group	X	X	
Prior Anti-TNF DMARD Use	X	X	
Prior Biologic DMARD Use	X	X	
Baseline mTSS (<= Median , > Median)	X	X	
Baseline DAS28-CRP (<= 3.2, 3.2 <- 5.1, > 5.1)	X	X	

Baseline DAS28-ESR (<= 3.2, 3.2 <- 5.1, > 5.1)	X	X	
Duration of RA (years)	X	X	
Number of Prior DMARDs Used	X	X	
Number of Prior Biologic DMARDs Used	X	X	
MTX Dose (mg/week) at Baseline (0 <- 8 mg/week, 8 <- 12 mg/week, > 12 mg/week)	X	X	
Concomitant Steroid at Baseline (No, Yes)	X	X	
Prednisone Dose at Baseline (mg/day) (None, 0 <- 5 mg/day, > 5 mg/day)	X	X	
Baseline CRP (< 1.0, >= 1.0)	X	X	
Baseline Anti-CCP Antibody (Low-positive or negative (<- 3 x ULN),High-positive (> 3 x ULN))	X	X	
Baseline Rheumatoid Factor (Low-positive or negative (<- 3 x ULN),High-positive (> 3 x ULN))	X	X	
Imputation Status at Week 28 (Yes, No)		X	
Body Weight (kg) at Screening (<- 40 kg, 40 kg <- 60 kg, 60 kg <- 80 kg, > 80 kg)	X	X	
Adjusted analysis	X[5]	X[6]	Only in the case of imbalance presented. Logistic regression[5] RANCOVA[6]

Abbreviations are defined in Section I, and subgroup categories in Section 7.8

[1] Based on FAS as analysis set, Fisher's exact test as statistical test, 7.11.1 Table4. 1. LOCF components will be used.

[2] Based on FAS as analysis set, Rank Analysis of Covariance (RANCOVA) with treatment group (Placebo, ASP015K 100 mg, ASP015K 150 mg) as factor and baseline rank mTSS as covariate as statistical test. The value will be imputed using linear extrapolation method described at 7.11.1

[3] For the sensitivity analysis, Multiplicity adjustment is not executed.

[4] Subgroup categories are provided in Section 7.8

[5] Logistic regression model with effect for treatment group and adjusted variable (if imbalance will be detected).

[6] Rank Analysis of Covariance (RANCOVA) with treatment group (Placebo, ASP015K 100 mg, ASP015K 150 mg) and adjusted variable (if imbalance will be detected) as factor and baseline rank mTSS as covariate

10.2 Appendix 2: Overview of Secondary Endpoints

Table 15 summarizes the analyses of the primary and secondary efficacy variables, with "X" indicating that the variable will be analyzed. All analyses will be based on the FAS, unless otherwise specified.

Table15. Overview of Secondary Endpoints analysis

Analysis Type/Analysis Variable [1]		CRP	ESR	None	Treatment Group	Analysis Method [2]
Categorical Variables						
1.1.1	Percentage of subjects achieving ACR20/50/70 at each visit (Weeks 4, 8, 12, 12/ET,28/ET)	X	X		Placebo, 100mg, 150mg	• Fisher's exact test. • ACR20-CRP at Week 12/ET is not displayed in this analysis.
1.1.2	Percentage of subjects achieving ACR20/50/70 at each visit (Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT)	X			100mg, 150mg, Placebo to 100mg at Week 12, Placebo to 150mg at Week 12, Placebo to 100mg at Week 28, Placebo to 150mg at Week 28	No statistical test
1.2.1	Line graph of 1.1.1	X	X		Placebo, 100mg, 150mg	No statistical test
1.2.2	Line graph of 1.1.2	X			100mg, 150mg, Placebo to 100mg at Week 12, Placebo to 150mg at Week 12, Placebo to 100mg at Week 28, Placebo to 150mg at Week 28	No statistical test.
2.1	Percentage of subjects achieving ACR20 response at Week 4 and sustaining response at each visit (Week 8, 12)	X			Placebo, 100mg, 150mg	Fisher's exact test.
2.2	Percentage of subjects achieving ACR20 response at Week 8 and sustaining response at Week 12	X			Placebo, 100mg, 150mg	Fisher's exact test.
3.1	Percentage of subjects achieving DAS28 score < 2.6 at each visit (Baseline, Week 4, 8, 12, 12/ET, 28/ET)	X	X		Placebo, 100mg, 150mg	Fisher's exact test.
3.2	Percentage of subjects achieving DAS28 score < 2.6 at each visit (Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT)	X	X		100mg, 150mg, Placebo to 100mg at Week 12, Placebo to 150mg at Week 12, Placebo to 100mg at Week 28, Placebo to 150mg at Week 28	No statistical test.
3.3	Percentage of subjects achieving DAS28 score	X	X		Placebo, 100mg, 150mg	Fisher's exact test.

	<=3.2 at each analysis visit					
3.4	Percentage of subjects achieving DAS28 score <=3.2 at each analysis visit	X	X		100mg, 150mg	No statistical test.
4.1	Percentage of subjects achieving "Good Response" using DAS28 EULAR response criterion at each visit (Baseline, Week 4, 8, 12, 12/ET,28/ET)	X	X		Placebo, 100mg, 150mg	Fisher's exact test.
4.2	Percentage of subjects achieving "Good Response" using DAS28 EULAR response criterion at each visit (Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT)	X	X		100mg, 150mg	No statistical test.
5.1	Percentage of subjects achieving "Good Response" or "Moderate Response" using DAS28 EULAR response criterion at each visit (Baseline, Week 4, 8, 12, 12/ET,28/ET)	X	X		Placebo, 100mg, 150mg	Fisher's exact test.
5.2	Percentage of subjects achieving "Good Response" or "Moderate Response" using DAS28 EULAR response criterion at each visit (Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT)	X	X		100mg, 150mg	No statistical test.
6.1	Percentage of subjects in ACR/EULAR remission at each visit (Baseline, Week 4, 8, 12, 12/ET,28/ET)	X			Placebo, 100mg, 150mg	Fisher's exact test.
6.2	Percentage of subjects in ACR/EULAR remission at each visit (Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT)	X			100mg, 150mg	No statistical test.
7.1	Percentage of subjects in SDAI <= 3.3 at each visit (Baseline, Week 4, 8, 12, 12/ET,28/ET)	X			Placebo, 100mg, 150mg	Fisher's exact test.
7.2	Percentage of subjects in SDAI <= 3.3 at each visit (Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT)	X			100mg, 150mg	No statistical test.

8.1	Percentage of subjects in CDAI <= 2.8 at each visit (Baseline, Week 4, 8, 12, 12/ET,28/ET)	X			Placebo, 100mg, 150mg	Fisher's exact test.
8.2	Percentage of subjects in CDAI <= 2.8 at each visit (Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT)	X			100mg, 150mg	No statistical test.
9.1	Percentage of subjects achieving HAQ-DI (<= 0.5) at each visit (Baseline, Week 4, 8, 12, 12/ET,28/ET)			X	Placebo, 100mg, 150mg	Fisher's exact test.
9.2	Percentage of subjects achieving HAQ-DI (<= 0.5) at each visit (Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT)			X	100mg, 150mg	No statistical test.
9.3	Percentage of subjects achieving decrease of baseline HAQ-DI of at least 0.22 at each visit (Baseline, Week 4, 8, 12, 12/ET,28/ET)			X	Placebo, 100mg, 150mg	Fisher's exact test.
9.4	Percentage of subjects achieving decrease of baseline HAQ-DI of at least 0.22 at each visit (Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT)			X	100mg, 150mg	No statistical test.
10.1	Percentage of subjects showing no progression of joint damage (mTSS change from baseline <= 0.5) at each visit (Baseline, Week 28/ET, EOT)			X	Placebo, 100mg, 150mg	Fisher's exact test.
10.2	Percentage of subjects showing no progression of joint damage (mTSS change from baseline <= 0) at each visit (Baseline, Week 28/ET, EOT)			X	Placebo, 100mg, 150mg	Fisher's exact test.
10.3	Percentage of subjects showing Clinically Relevant Radiographic Progression: CRRP (mTSS change from baseline >= 3) at each analysis visit			X	Placebo, 100mg, 150mg	Fisher's exact test.
10.4	Percentage of subjects showing Rapid Radiographic Progression: RRP (Yearly progression of mTSS >= 5)			X	Placebo, 100mg, 150mg	Fisher's exact test.

11.1	Percentage of subjects achieving SF-36v2 of difference >= 5 at each visit (Weeks 4, 8, 12, 12/ET, 28/ET)	X			Placebo, 100mg, 150mg	Fisher's exact test.
11.2	Percentage of subjects achieving SF-36v2 of difference >= 5 at each visit (Weeks 4, 8, 12, 28, 52, EOT)	X			100mg, 150mg	No statistical test.
Continuous Variables						
12.1.1	Raw value and change from baseline in mTSS score at each visit (Baseline, Week 52/ET)			X	Placebo, 100mg, 150mg	Descriptive Statistics RANCOVA[3] analysis with Week 52/ET. No multiplicity adjustment.
12.1.2	Raw value and change from baseline in Erosion Score at each visit (Baseline, Week 28/ET, Week 52/ET)			X	Placebo, 100mg, 150mg	Descriptive Statistics; RANCOVA[3] analysis with each visit. No multiplicity adjustment.
12.1.3	Raw value and change from baseline in Joint Space Narrowing Score at each visit (Baseline, Week 28/ET, Week 52/ET)			X	Placebo, 100mg, 150mg	Descriptive Statistics; RANCOVA[3] analysis with each visit. No multiplicity adjustment.
12.2.1	Cumulative probability plot for the change from baseline in mTSS score at Week 52/ET			X	Placebo, 100mg, 150mg	
12.2.2	Cumulative probability plot for the change from baseline in Erosion Score at Week 28/ET, Week 52/ET			X	Placebo, 100mg, 150mg	
12.2.3	Cumulative probability plot for the change from baseline in Joint Space Narrowing Score at Week 28/ET, Week 52/ET			X	Placebo, 100mg, 150mg	
12.3.1	Scatter plot for mTSS score at baseline to Week 28/ET, Week 52/ET			X	Placebo, 100mg, 150mg	
12.3.2	Scatter plot for mTSS score at baseline to mTSS change from baseline at Week 28/ET, Week 52/ET			X	Placebo, 100mg, 150mg	
12.4.1	Yearly Progression of mTSS			X	Placebo, 100mg, 150mg	Descriptive Statistics;
13.1.1	Raw value and change from baseline in DAS28	X	X		Placebo, 100mg, 150mg	Descriptive Statistics;

	score at each visit (Baseline, Week 4, 8, 12, 12/ET,28/ET)					ANCOVA[3] analysis with each visit. No multiplicity adjustment.
13.1.2	Raw value and change from baseline in DAS28 score at each visit (Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT)	X	X		100mg, 150mg, Placebo to 100mg at Week 12, Placebo to 150mg at Week 12, Placebo to 100mg at Week 28, Placebo to 150mg at Week 28	Descriptive Statistics
13.2.1	Mean-standard deviation plot of 13.1.1	X	X		Placebo, 100mg, 150mg	Descriptive Statistics ANCOVA[3] analysis with each visit. No multiplicity adjustment.
13.2.2	Mean-standard deviation plot of 13.1.2	X	X		100mg, 150mg, Placebo to 100mg at Week 12, Placebo to 150mg at Week 12, Placebo to 100mg at Week 28, Placebo to 150mg at Week 28	Descriptive Statistics
14.1	Raw values and change from baseline in SDAI score at each visit (Baseline, Week 4, 8, 12, 12/ET,28/ET)	X			Placebo, 100mg, 150mg	ANCOVA[3] analysis with each visit.
14.2	Raw values and change from baseline in SDAI score at each visit (Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT)	X			100 mg, 150 mg	Descriptive Statistics
15.1	Raw values and change from baseline in CDAI score at each visit (Baseline, Week 4, 8, 12, 12/ET,28/ET)	X			Placebo, 100mg, 150mg	ANCOVA[3] analysis with each visit.
15.2	Raw values and change from baseline in CDAI score at each visit (Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT)	X			100 mg, 150 mg	Descriptive Statistics
16.1	Raw values and change from baseline in SF-36v2 at each visit (Baseline, Week 4, 8, 12, 12/ET,28/ET)			X	Placebo, 100mg, 150mg	ANCOVA[3] analysis with each visit.
16.2	Raw values and change from baseline in SF-36v2 at each visit (Baseline, Weeks 4, 8, 12, 28, 52, EOT)			X	100 mg, 150 mg	Descriptive Statistics
17.1	Raw values and change from baseline in WPAI at each visit (Baseline, Week 4, 8, 12,			X	Placebo, 100mg, 150mg	ANCOVA[3] analysis with each visit.

	12/ET,28/ET)					
17.2	Raw values and change from baseline in WPAI at each visit (Baseline, Weeks 4, 8, 12, 28, 52, EOT)			X	100 mg, 150 mg	Descriptive Statistics
18.1	Raw value and change from baseline in TJC68 at each visit (Baseline, Week 4, 8, 12, 12/ET,28/ET)			X	Placebo, 100mg, 150mg	ANCOVA[3] analysis with each visit.
18.2	Raw value and change from baseline in TJC68 at each visit (Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT)			X	100 mg, 150 mg	Descriptive Statistics
19.1	Raw value and change from baseline in TJC28 at each visit (Baseline, Week 4, 8, 12, 12/ET, 28/ET)			X	Placebo, 100mg, 150mg	ANCOVA[3] analysis with each visit.
19.2	Raw value and change from baseline in TJC28 at each visit (Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT)			X	100 mg, 150 mg	Descriptive Statistics
20.1	Raw value and change from baseline in SJC66 at each visit (Baseline, Week 4, 8, 12, 12/ET, 28/ET)			X	Placebo, 100mg, 150mg	ANCOVA[3] analysis with each visit.
20.2	Raw value and change from baseline in SJC66 at each visit (Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT)			X	100 mg, 150 mg	Descriptive Statistics
21.1	Raw value and change from baseline in SJC28 at each visit (Baseline, Week 4, 8, 12, 12/ET, 28/ET)			X	Placebo, 100mg, 150mg	ANCOVA[3] analysis with each visit.
21.2	Raw value and change from baseline in SJC28 at each visit (Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT)			X	100 mg, 150 mg	Descriptive Statistics
22.1	Raw value and change from baseline in CRP at each visit (Baseline, Week 4, 8, 12, 12/ET, 28/ET)	X			Placebo, 100mg, 150mg	ANCOVA[3] analysis with each visit.

22.2	Raw value and change from baseline in CRP at each visit (Screening, Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT)	X			100 mg, 150 mg	Descriptive Statistics
23.1	Raw value and change from baseline in ESR at each visit (Baseline, Weeks 4, 8, 12, 12/ET, 28/ET)			X	Placebo, 100mg, 150mg	ANCOVA[3] analysis with each visit.
23.2	Raw value and change from baseline in ESR at each visit (Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT)			X	100 mg, 150 mg	Descriptive Statistics
24.1	Raw value and change from baseline in SGAP (100 mm VAS) at each visit (Baseline, Week 4, 8, 12, 12/ET, 28/ET)			X	Placebo, 100mg, 150mg	ANCOVA[3] analysis with each visit.
24.2	Raw value and change from baseline in SGAP (100 mm VAS) at each visit (Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT)			X	100 mg, 150 mg	Descriptive Statistics
25.1	Raw value and change from baseline in SGA (100 mm VAS) at each visit (Baseline, Week 4, 8, 12, 12/ET, 28/ET)			X	Placebo, 100mg, 150mg	ANCOVA[3] analysis with each visit.
25.2	Raw value and change from baseline in SGA (100 mm VAS) at each visit (Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT)			X	100 mg, 150 mg	Descriptive Statistics
26.1	Raw value and change from baseline in PGA (100 mm VAS) at each visit (Baseline, Week 4, 8, 12, 12/ET, 28/ET)			X	Placebo, 100mg, 150mg	ANCOVA[3] analysis with each visit.
26.2	Raw value and change from baseline in PGA (100 mm VAS) at each visit (Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT)			X	100 mg, 150 mg	Descriptive Statistics
27.1	Raw value and change from baseline in HAQ-DI at each visit (Baseline, Week 4, 8, 12, 12/ET, 28/ET)			X	Placebo, 100mg, 150mg	ANCOVA[3] analysis with each visit.
27.2	Raw value and change from baseline in HAQ-DI			X	100 mg, 150 mg	Descriptive Statistics

	at each visit (Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT)					
Time-to-event Variable						
28	Time to the first occurrence of ACR20/50/70-CRP up to Week 12	X			Placebo, 100mg, 150mg	Cox, Kaplan-Meier, Log-rank No multiplicity adjustment.
29	Time to the first occurrence of DAS28-CRP < 2.6 up to Week 12	X			Placebo, 100mg, 150mg	Cox, Kaplan-Meier, Log-rank No multiplicity adjustment.
30	Time to the first occurrence of SDAI score <= 3.3 up to Week 12	X			Placebo, 100mg, 150mg	Cox, Log-rank No multiplicity adjustment.
31	Time to the first occurrence of ACR/EULAR score for remission up to Week 12	X			Placebo, 100mg, 150mg	Cox, Log-rank No multiplicity adjustment.
32	Time to the first occurrence of CDAI score <= 2.8 up to Week 12	X			Placebo, 100mg, 150mg	Cox, Log-rank No multiplicity adjustment.
Additional (Graphical Analysis)						
33	Two-panel line graphs of DAS28 over time up to Week 12 for ACR20 responders and non-responders at Week 12/ET (responders in 1 panel, non-responders in other panel)	X	X		Placebo, 100mg, 150mg	Descriptive Statistics

1 Abbreviations are defined in Section I. All analyses will be based on the FAS.
2 See Sections 7.4.2 for modeling details on ANCOVA, Kaplan-Meier, Cox proportional hazards, and Log-rank test.
3 ANCOVA for change only. For raw value, only descriptive statistics will be presented.

10.3 Appendix 3: Overview of AE analysis

Table16, Table17 summarizes the analyses of AE occurrence, the analysis of SOC, PT analysis, with "X1", "X2", "X3", and "X4" indicating that the variable will be analyzed by each defined groups defined footnote. All analyses will be based on the SAF.

Table16. The overview of AE occurrence analysis

	Overall	Week 0 to Week 12	Week 12 to Week 28	Week 28 to Week 52
Number and percentage of subjects with TEAE	X4	X1	X2	X3
Number and percentage of subjects with drug related TEAE	X4	X1	X2	X3
Number and percentage of subjects with death	X4	X1	X2	X3
Number and percentage of subjects with serious TEAE	X4	X1	X2	X3
Number and percentage of subjects with drug related serious TEAE	X4	X1	X2	X3
Number and percentage of subjects with TEAE Grade 3 or Higher in Severity	X4	X1	X2	X3
Number and percentage of subjects with TEAE leading to study drug discontinuation	X4	X1	X2	X3
Number and percentage of subjects with drug related TEAE leading to study drug discontinuation	X4	X1	X2	X3
Number and percentage of subjects with Serious TEAE Leading to Permanent Discontinuation of Study Drug	X4	X1	X2	X3
Number and percentage of subjects with Drug-Related Serious TEAE Leading to Permanent Discontinuation of Study Drug	X4	X1	X2	X3

X1: Placebo, 100mg, 150mg
X2: Placebo, 100mg, 150mg, Placebo to 100 mg at Week 12, Placebo to 150 mg at Week12
X3: 100mg, 150mg, Placebo to 100 mg at Week 12, Placebo to 150 mg at Week 12, Placebo to 100 mg at Week 28, Placebo to 150 mg at Week 28
X4: 100mg, 150mg

Table17. The overview of the number of SOC/PT analysis

	Overall	Week 0 to Week 12	Week 12 to Week 28	Week 28 to Week 52
TEAEs	X4	X1	X2	X3

Drug related TEAEs, defined as any TEAE with possible, probable or missing relationship to study drug as assessed by the investigator.	X4	X1	X2	X3
Serious TEAEs	X4	X1	X2	X3
Drug related serious TEAEs		X1		
TEAEs by severity based on NCI-CTCAE	X4	X1	X2	X3
Drug related TEAEs by severity		X1		
Grade 3 or higher TEAEs by severity	X4	X1		
TEAEs by relationship to study drug		X1		
TEAEs leading to death	X4	X1	X2	X3
TEAEs leading to permanent discontinuation of study drug	X4	X1	X2	X3
Drug related TEAEs leading to permanent discontinuation of study drug		X1		
Serious TEAEs leading to permanent discontinuation of study drug		X1		
Drug related serious TEAEs leading to permanent discontinuation of study drug		X1		
TEAEs excluding serious adverse events that equal to or Exceed a threshold of 5% in Any Treatment Group	X4			
Most common TEAEs (> 5% in any treatment group)	X4	X1	X2	X3
Most common serious TEAEs (> 5% in any treatment group)	X4	X1	X2	X3
TEAEs by subgroups defined in section 7.8.		X1		
The number and percentage of subjects with TEAEs of special interest	X4	X1		
Serious Infections	X4	X1		
Malignancies	X4	X1		
Herpes Zoster Related Disease (Herpes Zoster and Varicella)	X4	X1		
Herpes Zoster	X4	X1		
Varicella	X4	X1		
Infections That Require Intravenous Anti-infectious Therapy	X4	X1		
Subgroup				
Sex		X1		
Age Group		X1		
Prior Anti-TNF DMARD Use		X1		
Prior Biologic DMARD Use		X1		
Duration of RA (years)		X1		

Number of Prior DMARDs Used		X1		
Number of Prior Biologic DMARDs Used		X1		
MTX Dose (mg/week) at Baseline (0 <- 8 mg/week, 8 <- 12 mg/week, >12 mg/week)		X1		
Concomitant Steroid at Baseline (No, Yes)		X1		
Prednisone Dose at Baseline (mg/day) (None, 0 <- 5 mg/day, > 5 mg/day)		X1		
Body Weight (kg) at Screening		X1		

X1: Placebo, 100mg, 150mg
X2: Placebo, 100mg, 150mg, Placebo to 100 mg at Week 12, Placebo to 150 mg at Week12
X3: 100mg, 150mg, Placebo to 100 mg at Week 12, Placebo to 150 mg at Week 12, Placebo to 100 mg at Week 28, Placebo to 150 mg at Week 28
X4: 100mg, 150mg

Patient-year analysis

	Week 0 to Week 12	Overall Period
TEAEs	X1	X4
TEAEs of special interest	X1	X4
Serious Infections	X1	X4
Malignancies	X1	X4
Herpes Zoster Related Disease (Herpes Zoster and Varicella)	X1	X4
Herpes Zoster	X1	X4
Varicella	X1	X4
Infections That Require Intravenous Anti-infectious Therapy*	X1	X4
Subgroup		
Sex		X4
Age Group		X4
Concomitant Steroid at Baseline (No, Yes)		X4
Prednisone Dose at Baseline (mg/day) (None, 0 <- 5 mg/day, > 5 mg/day)		X4

X1: Placebo, 100mg 150mg
X4: 100mg 150mg

*: subgroup analysis will not be conducted.

10.4 Appendix 4: Overview of Clinical Laboratory Evaluation, Vital Singes, and Electrocardiograms (ECGs) analyses

Table18 summarizes the analyses of AE occurrence, the analysis of SOC, PT, with "X1" , "X2", "X3" indicating that the variable will be analyzed by each defined groups defined footnote. All analyses will be based on the SAF.

Table18. The overview of Clinical Laboratory Evaluation, Vital Singes, and Electrocardiograms (ECGs) analyses

	Screening, Baseline, Weeks 4, 8, 12, 12/ET, 28/ET	Screening, Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, EOT
Quantitative Laboratory Test Results in SI Units, Hematology Actual and Change from Baseline	X1	X2
Quantitative Laboratory Test Results in SI Units, Biochemistry Actual and Change from Baseline	X1	X2
Quantitative Laboratory Test Results in SI Units, Urinalysis Actual and Change from Baseline	X1	X2
Qualitative Laboratory Test Results, Urinalysis 1	X1	X2
Qualitative Laboratory Test Results, Urinalysis 2	X1	X2
Qualitative Laboratory Test Results, Urinalysis 3	X1	X2
Qualitative Laboratory Test Results, Urinalysis 4	X1	X2
Qualitative Laboratory Test Results, Urinalysis 5	X1	X2
Quantitative Laboratory Test Results in SI Units, Fasting Lipids Profile Actual and Change from Baseline	X1	X2
Quantitative Laboratory Test Results in SI Units, Lipids Profile Including Values Under Non-fasting Condition at Baseline, Weeks 12, 12/ET, 28/ET Actual and Change from Baseline	X1	X2
Quantitative Laboratory Test Results in SI Units, T-Chol/HDL, LDL/HDL Ratio Actual and Change from Baseline	X1	X2
Quantitative Laboratory Test Results in SI Units, Erythrocyte Sedimentation Rate (ESR) Actual and Change from Baseline	X1	X2
Shift-from-Baseline Table for Laboratory Test Results, Hematology Baseline to Week 12/ET	X1	X3

Shift-from-Baseline Table for Laboratory Test Results, Hematology Baseline to Lowest Value up to Week 12	X1	X3
Shift-from-Baseline Table for Laboratory Test Results, Hematology Baseline to Highest Value up to Week 12	X1	X3
Shift-from-Baseline Table for Laboratory Test Results, Biochemistry Baseline to Week 12/ET	X1	X3
Shift-from-Baseline Table for Laboratory Test Results, Biochemistry Baseline to Lowest Value up to Week 12	X1	X3
Shift-from-Baseline Table for Laboratory Test Results, Biochemistry Baseline to Highest Value up to Week 12	X1	X3
Shift-from-Baseline Table for Laboratory Test Results, Urinalysis Baseline to Week 12/ET	X1	X3
Shift-from-Baseline Table for Laboratory Test Results, Urinalysis Baseline to Lowest Value up to Week 12	X1	X3
Shift-from-Baseline Table for Laboratory Test Results, Urinalysis Baseline to Highest Value up to Week 12	X1	X3
Shift-from-Baseline Table for Laboratory Test Results, Fasting Lipid Profile Baseline to Week 12/ET	X1	X3
Shift-from-Baseline Table for Laboratory Test Results, Fasting Lipid Profile Baseline to Lowest Value up to Week 12	X1	X3
Shift-from-Baseline Table for Laboratory Test Results, Fasting Lipid Profile Baseline to Week 12/ET	X1	X3
Shift-from-Baseline Table in Pre-Specified Reference Range Laboratory test Results by Visit, Selected Laboratory Variables		
Aspartate Aminotransferase (AST)	X1	X2
Alanine Aminotransferase (ALT)	X1	X2
Alkaline Phosphatase (ALP)	X1	X2
Total Bilirubin Category 1	X1	X2
Total Bilirubin Category 2	X1	X2
Low-Density Lipoprotein (LDL)	X1	X2
Hemoglobin (HGB) Category 1	X1	X2

Hemoglobin (HGB) Category 2	X1	X2
Hemoglobin (HGB) Category 3	X1	X2
Creatine Phosphokinase (CPK) Category 1	X1	X2
Creatine Phosphokinase (CPK) Category 2	X1	X2
Creatine Phosphokinase (CPK) Category 3	X1	X2
Creatinine	X1	X2
Absolute Neutrophil Count (ANC)	X1	X2
Lymphocytes Category 1	X1	X2
Lymphocytes Category 2	X1	X2
Platelets Cagetory1	X1	X2
Platelets Cagetory2	X1	X2
Shift-from-Baseline Table for Selected Laboratory Variables using NCI-CTCAE Toxicity Grade (version 4.0) Baseline to Week 12/ET	X1	X3
Liver Function Analysis Clinically Significant Values in Liver Function Tests	X1	X2
Summary of Moderate and Marked Liver Abnormalities by Visit	X1	X2
Liver Functions Plots	X1	X2
Vital Signs		
Vital Signs Actual and Change from Baseline	X1	X2
ECG Results by Local Investigator		
Interpretation of 12- Lead ECG Results, Assessment by Investigator Shift-from-Screening Table for 12-Lead ECG	X1	X2
Central ECG Results		
Summary of Central ECG Results Actual and Change From Baseline	X1	
Summary of Central ECG Results Change From Predose to Postdose at Week 4/ Week 8	X1	
Number and Percentage of Subjects with Categorized Central ECG Parameter Low Extreme Values	X1	
Number and Percentage of Subjects with Categorized Central ECG Parameter High Extreme Values	X1	
Number and Percentage of Subjects with Categorized Central ECG Parameter Low Extreme Values Change From Baseline	X1	
Number and Percentage of Subjects with Categorized Central ECG Parameter High Extreme Change Values Change From Baseline	X1	
Number and Percentage of Subjects with Categorized	X1	

Central ECG Parameter Low Extreme Change Values Change From Predose to Postdose at Week 4/ Week 8		
Number and Percentage of Subjects with Categorized Central ECG Parameter High Extreme Change Values Change From Predose to Postdose at Week 4/ Week 8	X1	

X1: Placebo, 100mg, 150mg
X2: 100mg, 150mg, Placebo to 100 mg at Week 12, Placebo to 150 mg at Week 12, Placebo to 100 mg at Week 28, Placebo to 150 mg at Week 28
X3: 100mg, 150mg

10.5 Appendix 5: Computation of HAQ-DI Score

The HAQ-DI is composed of 20 items in 8 categories (Dressing and Grooming, Arising, Eating, Walking, Hygiene, Reach, Grip, and Activities). Each category has at least two questions. Within each category, subjects report the amount of difficulty they have in performing the specific question items.

Table 19. Classification of HAQ Question and Checkbox for computation of HAQ-DI Score

Category	Question	Checkbox
1. Dressing and Grooming	1) Dress yourself, including shoelaces and buttons 2) Shampoo your hair	● 1 "aids or devices" check box: o Devices used for dressing (button hook, zipper pull, etc.) ● 1 "help from another person" checkbox: o Dressing and grooming
2. Arising	1) Stand up from a straight chair 2) Get in and out of bed	● 1 "aids or devices" checkbox: o Special or built up chair ● 1 "help from another person" checkbox: o Arising
3. Eating	1) Cut your own meat 2) Lift a full cup or glass to your mouth 3) Open a new milk carton	● 1 "aids or devices" checkbox: o Built up or special utensils ● 1 "help from another person" checkbox: o Eating
4. Walking	1) Walk outdoors on flat ground 2) Climb up five steps	● 4 "aids or devices" checkboxes: o Cane o Walker o Crutches o Wheelchair ● 1 "help from another person" checkbox: o Walking

Category	Question	Checkbox
5. Hygiene	1) Wash and dry your body 2) Take a tub bath 3) Get on and off the toilet	● 4 "aids or devices" checkboxes: o Raised toilet seat o Bathtub seat o Bathtub bar o Long-handled appliances in bathroom ● 1 "help from another person" checkbox: o Hygiene
6. Reach	1) Reach and get down a 5 pound object (such as a bag of sugar) from above your head 2) Bend down to pick up clothing from the floor	● 1 "aids or devices" checkbox: o Long-handled appliances for reach ● 1 "help from another person" checkbox: o Reach
7. Grip	1) Open car doors 2) Open previously opened jars 3) Turn faucets on and off	● 1 "aids or devices" checkbox: o Jar opener (for jars previously opened) ● 1 "help from another person" checkbox: o Gripping and opening things
8. Activities	1) Run errands and shop 2) Get in and out of a car 3) Do chores such as vacuuming or yard work	● 1 "help from another person" checkbox: o Errands and chores

For each question, there are four response options ranging from "Without Any Difficulty" to "Unable to Do", scored 0 - 3. Details are as follows.

Response	Score
Without Any Difficulty	0
With Some Difficulty	1
With Much Difficulty	2
Unable to Do	3

The patient must have a score for ≥ 6 of the 8 categories. If there are less than 6 categories completed, a HAQ-DI cannot be computed, whether the missing categories are due to missing values or they do not apply to the respondent. Individual questions within a category are not imputed. Therefore the maximum score in each category is based on non-missing questions, and a category score is missing when all questions within a category are missing.

There are three steps to compute the HAQ-DI score.

1. For each category, compute the category score by using the highest question score.

- For example, in the category "Eating" there are 3 questions, a subject responds with a 1, 2, and 0, respectively.

Category	Question	Subject Reported Response
Eating	1) Cut your own meat	1
	2) Lift a full cup or glass to your mouth	2
	3) Open a new milk carton	0

- The highest score is 2, so the category score is 2 for this subject.

2. Adjust for use of "aids or devices" and/or "help from another person" when indicated in the checkbox(es)
   - If the category score is < 2 but at least one "aids or devices" or "help from another person" box is checked, the category score is set equal to 2.
   - If the category score is < 2 and none of the "aids or devices" or "help from another person" boxes is checked, the category score remains.
   - If the category score is 2, it remains 2, and if a three, it remains a three.
3. Divide the summed category scores by the number of categories answered (must be a minimum of 6) to obtain a HAQ-DI score of 0 – 3, with higher scores indicating greater disability (3 = worst functioning).

10.6 Appendix 6: Computation of SF-36v2® Score

The SF-36v2® will be scored for the 8 scales according to the standard SF-36v2® scoring algorithm (0-100 scale) explained in the SF-36v2® Japanese Manual (Fukuhara et al., 2011). The physical component score (PCS), mental component score (MCS), role/social component score (RCS) will be scored according to the standard SF-36v2® scoring algorithm (0-100 scale) explained in the same Manual. A higher score indicates a better health state. Prior to the analysis, the responses will be scored according to the following four steps.

1. Item recoding for the 10 items which require recoding
2. Computing raw scale scores by summing across items within the same scale (raw scale scores)
3. Transforming the raw scale scores into a 0-100 scale (transformed scale scores)
4. Normalizing the transformed scale scores with a mean of 50 and a standard deviation of 10 in the general Japanese population (norm-based scale scores)

Questions, coding and scoring of the 8 scales (Physical Functioning, Role-Physical, Bodily Pain, General Health, Vitality, Social Functioning, Role-Emotional, Metal Health), PCS and MCS are presented below.

Table 20.1 Physical Functioning: Verbatim Items and Scoring Information

Question Number	Verbatim Items
3.	The following questions are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much?
3a.	*Vigorous activities*, such as running, lifting heavy objects, participating in strenuous sports
3b.	*Moderate activities*, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf
3c.	Lifting or carrying groceries
3d.	Climbing *several* flights of stairs
3e.	Climbing *one* flight of stairs
3f.	Bending, kneeling, or stooping
3g.	Walking *more than a mile*
3h.	Walking *several hundred yards*
3i.	Walking *one hundred yards*
3j.	Bathing or dressing yourself

Precoded and Final Values for Items 3a through 3j		
Response Choices	**Precoded Item Value**	**Final Item Value**
Yes, limited a lot	1	1
Yes, limited a little	2	2
No, not limited at all	3	3

Table 20.2. Role-Physical: Verbatim Items and Scoring Information

Question Number	Verbatim Items
4.	During the past 4 weeks, how much of the time have you had any of the following problems with your work or other regular daily activities as a result of your physical health?
4a.	Cut down the *amount of time* you spent on work or other activities
4b.	*Accomplished less* than you would like
4c.	Were limited in the *kind* of work or other activities
4d.	Had *difficulty* performing the work or other activities (for example, it took extra effort)

Precoded and Final Values for Items 4a through 4d		
Response Choices	**Precoded Item Value**	**Final Item Value**
All of the time	1	1
Most of the time	2	2
Some of the time	3	3
A little of the time	4	4
None of the time	5	5

Table 20.3. Bodily Pain: Verbatim Items and Scoring Information

Question Number	Verbatim Items
7.	How much bodily pain have you had during the past 4 weeks?
8.	During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)?

Precoded and Final Values for Item 7		
Response Choices	**Precoded Item Value**	**Final Item Value**
None	1	6.0
Very mild	2	5.4
Mild	3	4.2
Moderate	4	3.1
Severe	5	2.2
Very Severe	6	1.0

Scoring for Item 8 if both Items 7 and 8 are Answered			
Response Choices	**If Item 8 Precoded Item Value**	**And Item 7 Precoded Item Value**	**Then Final Item Value**
Not at all	1	1	6
	1	2 through 6	5
A little bit	2	1 through 6	4
Moderately	3	1 through 6	3
Quite a bit	4	1 through 6	2
Extremely	5	1 through 6	1

Scoring for Item 8 if Item 7 is Not Answered		
Response Choices	**Precoded Item Value**	**Then Final Item Value**
Not at all	1	6.0
A little bit	2	4.75
Moderately	3	3.5
Quite a bit	4	2.25
Extremely	5	1.0

Table 20.4. General Health: Verbatim Items and Scoring Information

Question Number	Verbatim Items
1.	In general, would you say your health is :
11.	How TRUE or FALSE is each of the following statements for you?
11a.	I seem to get sick a little easier than other people
11b.	I am as healthy as anybody I know
11c.	I expect my health to get worse
11d.	My health is excellent

Precoded and Final Values for Items 1 and 11a through 11d			
Item 1	**Response Choices**	**Precoded Item Value**	**Final Item Value**
	Excellent	1	5.0
	Very good	2	4.4
	Good	3	3.4
	Fair	4	2.0
	Poor	5	1.0
Items 11a and 11c	**Response Choices**	**Precoded Item Value**	**Final Item Value**
	Definitely True	1	1
	Mostly True	2	2
	Don't Know	3	3
	Mostly False	4	4
	Definitely False	5	5
Items 11b and 11d	**Response Choices**	**Precoded Item Value**	**Final Item Value**
	Definitely True	1	5
	Mostly True	2	4
	Don't Know	3	3
	Mostly False	4	2
	Definitely False	5	1

Table 20.5. Vitality: Verbatim Items and Scoring Information

Question Number	Verbatim Items
9.	These questions are about how you feel and how things have been with you during the past 4 weeks. For each question, please give the one answer that comes closest to the way you have been feeling. How much of the time during the past 4 weeks …
9a.	Did you feel full of life?
9e.	Did you have a lot of energy?
9g.	Did you feel worn out?
9i.	Did you feel tired?

Precoded and Final Values for Items 9a, 9e, 9g and 9i			
Item 9a and 9e	**Response Choices**	**Precoded Item Value**	**Final Item Value**
	All of the time	1	5
	Most of the time	2	4
	Some of the time	3	3
	A little of the time	4	2
	None of the time	5	1
Items 9g and 9i	**Response Choices**	**Precoded Item Value**	**Final Item Value**
	All of the time	1	1
	Most of the time	2	2
	Some of the time	3	3
	A little of the time	4	4
	None of the time	5	5

Table 20.6. Social Functioning: Verbatim Items and Scoring Information

Question Number	Verbatim Items
6.	During the past 4 weeks, to what extent has your physical health or emotional problems interfered with your normal social activities with family, friends, neighbors, or groups?
10.	During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)?

Precoded and Final Values for Items 6 and 10			
Item 6	**Response Choices**	**Precoded Item Value**	**Final Item Value**
	Not at all	1	5
	Slightly	2	4
	Moderately	3	3
	Quite a bit	4	2
	Extremely	5	1
Item 10	**Response Choices**	**Precoded Item Value**	**Final Item Value**
	All of the time	1	1
	Most of the time	2	2
	Some of the time	3	3
	A little of the time	4	4
	None of the time	5	5

Table 20.7. Role-Emotional: Verbatim Items and Scoring Information

Question Number	Verbatim Items
5.	During the past 4 weeks, how much of the time have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)?
5a.	Cut down on the *amount of time* you spent on work or other activities
5b.	*Accomplished less* than you would like
5c.	Did work or other activities *less carefully than usual*

Precoded and Final Values for Items 5a through 5c			
Item 5a through 5c	**Response Choices**	**Precoded Item Value**	**Final Item Value**
	All of the time	1	1
	Most of the time	2	2
	Some of the time	3	3
	A little of the time	4	4
	None of the time	5	5

Table 20.8. Mental Health: Verbatim Items and Scoring Information

Question Number	Verbatim Items
9.	These questions are about how you feel and how things have been with you during the past 4 weeks. For each question, please give the one answer that comes closest to the way you have been feeling. How much of the time during the past 4 weeks …
9b.	Have you been very nervous?
9c.	Have you felt so down in the dumps that nothing could cheer you up?
9d.	Have you felt calm and peaceful?
9f.	Have you felt downhearted and depressed?
9h.	Have you been happy?

Precoded and Final Values for Items 9b, 9c, 9d, 9f and 9h			
Items 9b, 9c and 9f	**Response Choices**	**Precoded Item Value**	**Final Item Value**
	All of the time	1	1
	Most of the time	2	2
	Some of the time	3	3
	A little of the time	4	4
	None of the time	5	5
Item 9d and 9h	**Response Choices**	**Precoded Item Value**	**Final Item Value**
	All of the time	1	5
	Most of the time	2	4
	Some of the time	3	3
	A little of the time	4	2
	None of the time	5	1

Table 20.9. Reported Health Transition: Verbatim Items and Scoring Information

Question Number	Verbatim Items
2.	Compared to one year ago, how would you rate your health in general now?

Precoded and Final Values for Item 2			
Item 2	**Response Choices**	**Precoded Item Value**	**Final Item Value**
	Much better now than one year ago	1	1
	Somewhat better now than one year ago	2	2
	About the same as one year ago	3	3
	Somewhat worse now than one year ago	4	4
	Much worse now than one year ago	5	5

Use Table 20 to compute simple algebraic sums of the presented final item scores.

Table 21. Formulas for Scoring and Transforming Scales

SF-36v2 Scale	Sum Final Item Values (after recoding items as in Tables 7.1-7.9)	Lowest and Highest Possible Raw Scores	Possible Raw Score Range
Physical Functioning	3a + 3b + 3c + 3d + 3e + 3f + 3g + 3h + 3i + 3j	10, 30	20
Role-Physical	4a + 4b + 4c + 4d	4, 20	16
Bodily Pain	7 + 8	2, 12	10
General Health	1 + 11a + 11b + 11c + 11d	5, 25	20
Vitality	9a + 9e + 9g + 9i	4, 20	16
Social Functioning	6 + 10	2, 10	8
Role-Emotional	5a + 5b + 5c	3, 15	12
Mental Health	9b + 9c + 9d + 9f + 9h	5, 25	20

Formula for transformation of raw scale scores to 0-100 scale scores

$$\text{Transformed Scale} = \frac{\text{(Actual raw score – lowest possible raw score)}}{\text{(Possible raw score range)}} \times 100$$

After calculating the transformed scale score, the next step is to compute a z-score transformation. A z-score for each scale is computed by subtracting the 2007 General Japanese Population Means for each SF-36v2® scale and dividing the difference by the corresponding scale standard deviation (Table 20) from the 2007 General Japanese Population Means and Standard Deviations Used to Derive SF-36v2® Z-score. Formulas are listed below.

Table22. 2007 General Japanese Population Means and Standard Deviations Used to Derive SF-36v2® Z-score

SF-36v2 Scale	Mean	Standard Deviation
Physical Functioning (PF)	89.13446	13.85045
Role-Physical (RP)	89.24007	18.80773
Bodily Pain (BP)	73.77098	22.39818
General Health (GH)	62.91007	18.76562
Vitality (VT)	62.82787	19.46255
Social Functioning (SF)	86.38347	19.40441
Role-Emotional (RE)	87.84637	20.01521
Mental Health (MH)	71.60598	18.62983

Step 1. Formulas for Z-score Standardization of SF-36v2® Scales

PF_Z = (PF – 89.13446) / 13.85045
RP_Z = (RP – 89.24007) / 18.80773
BP_Z = (BP – 73.77098) / 22.39818
GH_Z = (GH – 62.91007) / 18.76562
VT_Z = (VT – 62.82787) / 19.46255
SF_Z = (SF – 86.38347) / 19.40441
RE_Z = (RE – 87.84637) / 20.01521
MH_Z = (MH – 71.60598) / 18.62983

Means and standard deviations are from Table 20.

Step 2. Norm-based Transformation of SF-36v2® Z-scores

The next step involves transforming each SF-36v2® z-score to the norm-based (50, 10) scoring. This is accomplished by multiplying each z-score from Step 1 by 10 and adding the resulting product to 50. Formulas are listed below.

Norm-Based Physical Functioning (PF) = 50 + (PF_Z × 10)
Norm-Based Role-Physical (RP) = 50 + (RP_Z × 10)
Norm-Based Bodily Pain (BP) = 50 + (BP_Z × 10)
Norm-Based General Health (GH) = 50 + (GH_Z × 10)
Norm-Based Vitality (VT) = 50 + (VT_Z × 10)
Norm-Based Social Functioning (SF) = 50 + (SF_Z × 10)
Norm-Based Role-Emotional (RE) = 50 + (RE_Z × 10)
Norm-Based Mental Health (MH) = 50 + (MH_Z × 10)

PCS, MCS and RCS are scored in three steps as explained below:

Step 1. Z-score Standardization of SF-36v2® Scales

The first consists of standardizing each of the 8 SF-36v2® scales using a z-score transformation. This is the same as Step 1 used in the norm-based scoring of the 8 SF-36 scales.

Step 2. Aggregating Scales in Estimating Aggregate Physical, Mental, and Role/social Component Scores

After a z-score has been computed for scale, the second step involves computation of aggregate scores for the physical, mental and role/social components using the physical and mental factor score coefficients from 2002 survey as given in Table 22.

Table 23. 2002 Factor Score Coefficients Used to Derive PCS , MCS, and RCS Scale Scores

SF-36v2 Scale	PCS	MCS	RCS
Physical Functioning (PF)	0.67908	-0.20472	-0.13048
Role Physical (RP)	0.22298	-0.27243	0.40393
Bodily Pain (BP)	0.37244	0.14644	-0.21786
General Health (GH)	0.36992	0.33933	-0.41710
Vitality (VT)	-0.08420	0.46413	-0.13120
Social Functioning (SF)	-0.30769	0.06727	0.49261
Role Emotional (RE)	-0.14256	-0.15597	0.61022
Mental Health (MH)	-0.33155	0.44572	0.10326

Computation of an aggregate physical component score consists of multiplying each SF-36v2® scale z-score by its respective physical factor score coefficient and summing the eight products, as shown below. Similarly, an aggregate mental, and role/social component score is obtained by multiplying each SF-36v2® scale z-score by its respective mental factor score coefficient and summing the eight products.

Formulas for Aggregating Scales in Estimating Aggregate Physical, Mental, and Role/social Component Scores (Standard Form)

AGG_PHYS = (PF_Z × 0.67908) + (RP_Z × 0.22298) + (BP_Z × 0.37244) + (GH_Z × 0.36992) + (VT_Z × -0.08420) + (SF_Z ×-0.30769) + (RE_Z × -0.14256) + (MH_Z × -0.33155)

AGG_MENT = (PF_Z × -0.20472) + (RP_Z × -0.27243) + (BP_Z × 0.14644) + (GH_Z × 0.33933) + (VT_Z × 0.46413) + (SF_Z ×0.06727) + (RE_Z × -0.15597) + (MH_Z × 0.44572)

AGG_ROLE = (PF_Z × -0.13048) + (RP_Z × 0.40393) + (BP_Z × -0.21786) + (GH_Z × -0.41710) + (VT_Z × -0.13120) + (SF_Z × 0.49261) + (RE_Z × 0.61022) + (MH_Z ×0.10326)

Step 3. Formulas for T-score Transformation of Component Scores (Standard Form)

The third step involves transforming each component score to the norm-based (50, 10) scoring. This is accomplished by multiplying each aggregate component scale score by 10 and adding the resulting product to 50.

$$\text{Transformed Physical (PCS)} = 50 + (\text{AGG\_PHYS} \times 10)$$
$$\text{Transformed Mental (MCS)} = 50 + (\text{AGG\_MENT} \times 10)$$
$$\text{Transformed Role/Social (RCS)} = 50 + (\text{AGG\_ROLE} \times 10)$$

Missing Data in an Individual Questionnaire

In the event that data is missing for an individual item from the scales of the SF-36v2® Health Status Survey, the average value of the completed items in the corresponding scale will be used as an estimate of the missing item. If more than 50 percent of the items from a scale are missing for an individual questionnaire, the corresponding deficient scale(s) will be excluded from analyses.

Items in each scale are given below:

SF-36v2 Scale	No. Items	Items	Minimum No. Non-Missing Items for Imputation of Missing Values
PF	10	3a, 3b, 3c, 3d, 3e, 3f, 3g, 3h, 3i, 3j	5
RP	4	4a, 4b, 4c, 4d	2
BP	2	7, 8	1
GH	5	1, 11a, 11b, 11c, 11d	3
VT	4	9a, 9e, 9g, 9i	2
SF	2	6, 10	1
RE	3	5a, 5b, 5c	2
MH	5	9b, 9c, 9d, 9f, 9h	3

Example Showing the Mechanics of Imputation Technique:

Suppose for a subject the items 1, 3b, 3e, 3f, 4a, 4b, 11b, 11d, 5a and 5b are missing. First group the missing items according to the scale:

- GH: Items 1, 11b, 11d
- PF: Items 3b, 3e, 3f
- RP: Items 4a and 4b
- RE: Items 5a and 5b

GH consists of 5 items and 3 are missing. That is only 2 have non-missing scores. Since at least half of the items are not non-missing, missing items 1, 11a and 11d cannot be imputed. Therefore for this subject, the General Health scale score will be missing.

PF consists of 10 items. So, for this subject 7 of the items are non-missing. That is, at least half of the items are non-missing. Therefore replace the missing Items 3b, 3e and 3f scores

by the average score of the non-missing items, i.e., replace by (3a + 3c + 3d + 3g + 3h + 3i + 3j) / 7.

RP consists of 4 items and two are missing. So at least half are non-missing. Therefore replace the missing item scores 4a and 4b by (4c+4d) / 2.

RE consists of 3 and two are missing. That is, more than half of the item scores are missing. Therefore, missing scores 5a and 5b cannot be imputed for this subject and thus RE scale score for this subject will be missing.

10.7 Appendix 7: Computation of WPAI Scale Score

Work Productivity and Activity Impairment Questionnaire: WPAI

In this study, the following scoring method are applied. These scoring method are based on the specific health problem version of WPAI (WPAI:SHP).

Question Number	Contents
1.	1.Are you currently employed (working for pay)?
2.	2.During the past seven days, how many hours did you miss from work because of problems associated with your rheumatoid arthritis?
3.	3.During the past seven days, how many hours did you miss from work because of any other reason, such as vacation, holidays, time off to participate in this study?
4.	4.During the past seven days, how many hours did you actually work?
5.	5.During the past seven days, how much did your rheumatoid arthritis affect your productivity while you were working?
6.	6.During the past seven days, how much did your rheumatoid arthritis affect your ability to do your regular daily activities, other than work at a job?

Scores:
Multiply scores by 100 to express in percentages.

- Percent work time missed due to problem: Q2/(Q2+Q4)
- Percent impairment while working due to problem: Q5/10
- Percent overall work impairment due to problem:
  Q2/(Q2+Q4)+[(1-(Q2/(Q2+Q4))x(Q5/10)]
- Percent activity impairment due to problem: Q6/10

10.8 Appendix 8: AE of Special Interest

1. Serious Infections
   - AE which belongs to SOC of Infections and infestations (10021881) and regarded as serious.
2. Malignancies
   Following PT terms are included.
   Diffuse large B-cell lymphoma (10012818), Bladder cancer (10005003), Breast cancer (10006187), Carcinoma in situ (10061450), Colon cancer (10009944), Gastric cancer (10017758), Renal cancer (10038389), Squamous cell carcinoma (10041823), Small cell lung cancer stage unspecified (10041071), Extraskeletal chondrosarcoma (10015838), Lymphoma (10025310), Thyroid cancer (10066474).

3. Herpes Zoster Related Disease (Herpes Zoster and Varicella)
   Following PT terms are included.
   Herpes zoster (10019974), Herpes zoster iridocyclitis (10019980), Herpes zoster ophthalmic (10019983), Herpes zoster multi-dermatomal (10058428), Herpes zoster infection neurological (10061208), Herpes zoster oticus (10063491), Herpes zoster disseminated (10065038), Encephalitis post varicella (10014603), Varicella (10046980), Varicella post vaccine (10063522).
4. Herpes Zoster
   Following PT terms are included.
   Herpes zoster (10019974), Herpes zoster iridocyclitis (10019980), Herpes zoster ophthalmic (10019983), Herpes zoster multi-dermatomal (10058428), Herpes zoster infection neurological (10061208), Herpes zoster oticus (10063491), Herpes zoster disseminated (10065038)
5. Varicella
   Following PT terms are included.
   Encephalitis post varicella (10014603), Varicella (10046980), Varicella post vaccine (10063522).
6. Infections That Require Intravenous Anti-infectious Therapy
   All PTs which belong to SOC of Infections and Infestations (10021881) for which there is an intravenous concomitant medication (antibiotics, antivirals, antifungals, etc.) associated with that event for that patient.

10.9 Appendix 9: Standard Toxicity Grading for Laboratory Tests According to National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE) v4.0

Table 24. Standard Toxicity Grading for Laboratory Tests According to National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE) v4.0

Test Name	Test Code	Units [1]	Direction of Change / AE Term	CTCAE Grade [2]					NCI Ref.		
				0	1	2	3	4	Version	Page	Units
Alanine Aminotransferase(ALT/SGPT)	ALT	n/a	Increase/ Alanine aminotransferase increased	WNL	>ULN - ≤ 3.0×ULN	>3.0×ULN - ≤ 5.0×ULN	>5.0×ULN - ≤ 20.0×ULN	>20.0 ×ULN	4.0	107	n/a
Albumin	ALB	g/dL	Decrease/ Hypoalbuminemia	WNL	≥3 - <LLN	≥2 - <3	<2		4.0	116	g/dL
Alkaline Phosphatase	ALKP	n/a	Increase/ Alkaline phosphatase increased	WNL	>ULN - ≤ 2.5×ULN	>2.5×ULN - ≤ 5.0×ULN	>5.0×ULN - ≤ 20.0×ULN	>20.0 ×ULN	4.0	107	n/a
Amylase	AMY	n/a	Increase/ Serum amylase increased	WNL	>ULN - ≤ 1.5×ULN	>1.5×ULN - ≤ 2.0×ULN	>2.0×ULN - ≤ 5.0×ULN	>5.0× ULN	4.0	112	n/a
Aspartate Aminotransferase(AST/SGOT)	AST	n/a	Increase/ Aspartate aminotransferase increased	WNL	>ULN - ≤3.0×ULN	>3.0×ULN - ≤5.0×ULN	>5.0×ULN - ≤20.0× ULN	>20.0 ×ULN	4.0	107	n/a
Calcium Corrected [3]	CACOR	mg/dL	Increase/ Hypercalcemia	WNL	>ULN - ≤ 11.5	>11.5 - ≤12.5	>12.5 - ≤ 13.5	>13.5	4.0	115	mg/dL
Calcium Corrected [3]	CACOR	mg/dL	Decrease/ Hypocalcemia	WNL	≥8.0 - <LLN	≥7.0 - <8.0	≥6.0 - <7.0	<6.0	4.0	116	mg/dL
Total Cholesterol	TCHOL	mg/dL	Increase/ Cholesterol high	WNL	>ULN - ≤300	>300 -≤ 400	>400- ≤500	>500	4.0	109	mg/dL
Creatine Phosphokinase	CPK	n/a	Increase/ CPK increased	WNL	>ULN - ≤ 2.5×ULN	>2.5×ULN - ≤ 5×ULN	>5×ULN - ≤ 10×ULN	>10× ULN	4.0	109	n/a
Creatinine	CREAT	n/a	Increase/ Creatinine increased	WNL	>ULN - ≤ 1.5×ULN or >ULN- ≤1.5×baseline	>1.5×ULN - ≤ 3.0×ULN or >1.5×ULN - ≤3.0×baseline	>3.0×ULN - ≤6.0×ULN or >3.0×baseline	>6.0× ULN	4.0	109	n/a
Estimated glomerular filtration rate	GFR	mL/min per 1.73 m^2	Decrease/ Chronic kidney disease	*Not applicable* [8]	*Not applicable* [8]	30 - 59	15 - 29	<15	4.0	147	mL/min per 1.73 m^2

Gamma Glutamyl Transferase (GGT)	GGT	n/a	Increase/ GGT increased	WNL	>ULN - ≤2.5×ULN	>2.5×ULN - ≤5.0×ULN	>5.0×ULN - ≤20.0×ULN	>20.0 ×ULN	4.0	110	n/a
Glucose	GLUC	mg/dL	Decrease/ Hypoglycemia	WNL	≥55 - <LLN	≥40 - <55	≥30- <40	<30	4.0	117	mg/dL
Hemoglobin	HGB	g/dL	Increase/ Hemoglobin increased	WNL	Increase in >0 - 2 g/dL above ULN or above baseline if baseline is above ULN	Increase in >2 - 4 g/dL above ULN or above baseline if baseline is above ULN	Increase in >4 g/dL above ULN or above baseline if baseline is above ULN		4.0	111	g/dL
Hemoglobin	HGB	g/dL	Decrease/ Anemia	WNL	≥10 - <LLN	≥8 - <10.0	<8		4.0	3	g/dL
Leukocytes (WBC)	WBC	$10^3/\mu L$	Decrease/ White blood cell decreased	WNL	≥3.0 - <LLN	≥2.0 - <3.0	≥1.0 - <2.0	<1.0	4.0	113	$10^3/\mu L$
Lymphocytes, Absolute Units	LYMPHAB	$10^3/\mu L$	Decrease/ Lymphocyte count decreased	*Not applicable*[8]	*Not applicable*[8]	≥0.5 - <0.8	≥0.2 - <0.5	<0.2	4.0	111	$10^3/\mu L$
Lymphocytes, Absolute Units	LYMPHAB	$10^3/\mu L$	Increase/ Lymphocyte count increased	*Not applicable*[8]		>4 - 20	>20		4.0	111	$10^3/\mu L$
Magnesium	MG	mg/dL	Increase/ Hypermagnese-mia	WNL	>ULN - ≤3.0		>3.0 - ≤8.0	>8.0	4.0	115	mg/dL
Magnesium	MG	mg/dL	Decrease/ Hypomagnesemia	WNL	≥1.2 - <LLN	≥0.9 - <1.2	≥0.7 - <0.9	<0.7	4.0	117	mg/dL
Absolute Neutrophil Count	ANC	$10^3/\mu L$	Decrease/ Neutrophil count decreased	*Not applicable*[8]	*Not applicable*[8]	≥1.0 - <1.5	≥0.5 - <1.0	<0.5	4.0	112	$10^3/\mu L$
Phosphate (Phosphorus)	PHOS	mg/dL	Decrease/ Hypophosphatemia	WNL	*Not applicable*[8]	≥2.0 - <2.5	≥1.0 - <2.0	<1.0	4.0	117	mg/dL
Platelets	PLT	$10^3/\mu L$	Decrease/ Platelet count decreased	WNL	≥75.0 - <LLN	≥50.0 - <75.0	≥25.0 - <50.0	<25.0	4.0	112	$10^3/\mu L$
Potassium	K	mEQ/L	Increase/ Hyperkalemia	WNL	>ULN - ≤5.5	>5.5 - ≤6.0	>6.0 - ≤7.0	>7.0	4.0	115	mEQ/L[7]
Potassium[4]	K	mEQ/L	Decrease/ Hypokalemia	WNL		≥3.0 - <LLN	≥2.5 - <3.0	<2.5	4.0	117	mEQ/L[7]
Urine Protein[5]	UPROT		Increase/ Proteinuria	WNL	+	++	+++ or ++++		4.0	149	n/a
Sodium	NA	mEQ/L	Increase/ Hypernatremia	WNL	>ULN - ≤150	>150 - ≤155	>155 - ≤160	>160	4.0	116	mEQ/L[7]

Sodium	NA	mEQ/L	Decrease/ Hyponatremia	WNL	≥130 - <LLN		≥120 - <130	<120	4.0	117	mEQ/L [7]
Total Bilirubin	TBILI	n/a	Increase/ Blood bilirubin increased	WNL	>ULN - ≤1.5×ULN	>1.5×ULN - ≤3.0×ULN	>3.0×ULN - ≤10.0×ULN	>10.0 ×ULN	4.0	107	n/a
Triglycerides	TRIG	mg/dL	Increase/ Hypertriglyceridemia	WNL	>ULN [9] - ≤300	>300 - ≤500	>500 - ≤1000	>1000	4.0	116	mg/dL
Uric Acid [6]	UACID	mg/dL	Increase/ Hyperuricemia	WNL			>ULN - 10	>10	4.0	116	mg/dL

[1] "n/a" is specified if the criteria are based strictly on a comparison of the value to its reference range.

[2] WNL = Within Normal Limits; LLN = Lower Limit of Normal; ULN = Upper Limit of Normal.

[3] **Calcium Corrected** will be derived using the following formula: Calcium Corrected = 0.8 × (4.0 – Albumin) + Calcium, where 4.0 represents the average albumin level in g/dL, Albumin is measured in unit of g/dL, and Calcium is measured in unit of mg/dL.

[4] Hypokalemia Grade 2 is the same as Grade1 but has additional clinical criteria that cannot be assessed. Therefore, it was decided to grade all criteria falling within ≥3.0 - <LLN as grade 2.

[5] For Urine Protein, the maximum category is "++++", while NCI CTCAE version4.0 doesn't define "+++", "++++". So, these are defined as grade 3 per internal discussion.

[6] Hyperuricemia Grade 3 is the same as Grade1 but has additional clinical criteria that cannot be assessed. Therefore, it was decided to grade all criteria falling within >ULN – 10 mg/dL as grade 3.

[7] For hyperkalemia, hypokalemia, hypernatremia and hyponatremia, the unit in the NCI reference is mmol/L, which has been converted to mEQ/L by multiplying 1.

[8] For ANC, LYMPHAB, GFR, there are no reference range in central laboratory, so, NCI CTCAE grading scale is only defined for not-related reference range category. For PHOS, LLN=2.5mg/dL, while NCI Grade 1 is defined "≥2.5 - <LLN", so Grade 1 can't be defined.

[9] For TRIG, the upper limit of normal range is 149 mg/dL. Therefore, WNL for Grade 0 is up to 149 mg/dL. But the lower limit of Grade 1 is originally more than 150 mg/dL. So, the definition for the lower limit of grade 1 is changed from 150 mg/dL to ULN.

- Hyperuricemia (see pg. 116 of NCI document) cannot be assessed accurately without knowing whether or not the subject was experiencing physiologic consequences, so it is not included in the table.
- Hemoglobinuria (see pg. 148 of NCI document) cannot be assessed accurately without knowing clinical or diagnostic observations, so it is not included in the table.
- Troponin T (see pg. 109 of NCI document) cannot be assessed accurately without knowing manufacturer ranges, so it is not included in the table.

10.10 Appendix 10: Key Contributors and Approvers

List of Key Contributors and Approvers

Key Contributors

The following contributed to or reviewed this Statistical Analysis Plan as relevant to their indicated discipline or role.

Primary author (s)

[redacted]

Contributors and Reviewers

[redacted], Medical Advisor

[redacted], Statistical Advisor

[redacted]

Author and Approver Signatories

(E-signatures are attached at end of document)

The Statistical Analysis Plan was prepared by :

[redacted],
[redacted]

This Statistical Analysis Plan was approved by:

[redacted]
[redacted]
[redacted]

This Statistical Analysis Plan was approved by:

[redacted]
[redacted]
[redacted]