Statistical Analysis Plan

Protocol Number: MT-5547-J01

A Phase 2/3 (Placebo-Controlled, Double-Blind, Comparative) Study on MT-5547 in Patients with Osteoarthritis Accompanied by Moderate to Severe Pain

For database at week 16
Version: 2.0
Date: 2 July 2020

For database at week 68
Version: 1.0
Date: 19 April 2021

NCT number: NCT03245008

Statistical Analysis Plan (For Database at Week 16)

List of Errata

Protocol No. MT-5547-J01

A Phase 2/3 (Placebo-Controlled, Double-Blind, Comparative) Study on MT-5547 in Patients with Osteoarthritis Accompanied by Moderate to Severe Pain

Prepared By:	Mitsubishi Tanabe Pharma Corporation
Version:	2.0
Date:	2-July-2020

NCT Number: NCT03245008

APPROVAL FORM

Statistical Analysis Plan(For Database at Week 16)

-List of Errata-

Protocol No.	MT-5547-J01
Protocol Title	A Phase 2/3 (Placebo-Controlled, Double-Blind, Comparative) Study on MT-5547 in Patients with Osteoarthritis Accompanied by Moderate to Severe Pain
Version / Date	2.0 / 2-July-2020

Authors:

Statistics Author	
Print Name:	[redacted]
Position:	Biostatistician

Approved by:

Statistic Approver	
Print Name:	[redacted]
Position:	Responsible Biostatistician
Signature:	Refer to email
Approval date:	Refer to email

1. INTRODUCTION

This list of errata is based on the statistical analysis plan for database at week 16 dated 10-February-2020.

2. LIST OF ERRATA

befroe change	after change
8.2.1. Primary Efficacy Endpoint **Sensitivity analyses** ● **Multiple Imputation (MI) analysis by Reason for Discontinued Treatment** The following SAS code will be used to generate the monotone missing data pattern: [redacted]	8.2.1. Primary Efficacy Endpoint **Sensitivity analyses** ● **Multiple Imputation (MI) analysis by Reason for Discontinued Treatment** The following SAS code will be used to generate the monotone missing data pattern: [redacted]
10. CHANGE FROM THE PROTOCOL There are currently no changes to analysis from protocol.	10. CHANGE FROM THE PROTOCOL Sensitivity analysis of the primary endpoint was changed from placebo multiple imputation analysis to multiple imputation analysis with values centered at the mean baseline value of the treatment group by reason for discontinued treatment. Because confirming the robustness against missing data using the multiple imputation method similar to the primary analysis in the overseas P3 study.

3. HISTORY OF ERRATA

Up to version 1.0 (24-April -2020)

befroe change	after change
4.2. Final Analysis This SAP for 16 week will be finalized before database lock in week 16. The efficacy analysis performed at week 16 is the final efficacy analyasis.	4.2. Final Analysis This SAP for 16 week will be finalized before database lock in week 16. The efficacy analysis performed at week 16 is the final efficacy analysis.
7.1.1.1. Demographic and Other Baseline Characteristics (3)Time since osteoarthritis diagnosis (years) If (month of date of consent > month of osteoarthritis diagnosis) or (month of date of consent = month of date osteoarthritis diagnosis, and day of date of consent >= day of date of osteoarthritis diagnosis), then Time since osteoarthritis diagnosis (years) = year of date of consent – year of date of osteoarthritis diagnosis if else case, then Time since osteoarthritis diagnosis (years) = year of date of consent – year of date osteoarthritis diagnosis – 1.	7.1.1.1. Demographic and Other Baseline Characteristics (3) Time since osteoarthritis diagnosis (years) If (month of date of consent > month of osteoarthritis diagnosis) or (month of date of consent = month of date osteoarthritis diagnosis, and day of date of consent >= day of date of osteoarthritis diagnosis), then Time since osteoarthritis diagnosis (years) = ~~year of~~ (date of consent – ~~year of~~ date of osteoarthritis diagnosis)/365.25* if else case, then Time since osteoarthritis diagnosis (years) = ~~year of~~ (date of consent – ~~year of~~ date osteoarthritis diagnosis – 1) /365.25*. * 1 year = 365.25 days
7.1.1.2. Treatment Duration and Compliance 1)Treatment Duration (days) = The date of last study drug injection prior to week 16 - the date of the first study drug injection + min(28 or (the date of last study drug infection - the date of Week 16 study drug injection))	7.1.1.2. Treatment Duration and Compliance 1)Treatment Duration (days) = The date of last study drug injection prior to week 16 - the date of the first study drug injection + min(28 or (the date of week 16 study drug infection - the date of last study drug injection prior to week 16))
7.1.2.6. SF-36 / EQ-5D-5L e.g) If answer is [1 1 1 1 1], Utility Index Score is=1+0+0=1. If answer is [1 1 3 2 1],	7.1.2.6. SF-36 / EQ-5D-5L e.g) If answer is [1 1 1 1 1], Utility Index Score is=1+0+0=1. If answer is [1 1 3 2 1],

Utility Index Score is=1+(-0.060924)+(-0.091131-0.068178)=0.779767.	Utility Index Score is=1+(-0.060924)+(-0.091131-0.044545)=0.8034.
8.1.4. Treatment Duration and Compliance Treatment compliance will be summarized by treatment group and combined MT-5547 on the SAF population.	8.1.4. Treatment Duration and Compliance Treatment compliance will be summarized by treatment group and combined MT-5547 on the FAS population.

[redacted]

APPROVAL FORM

Statistical Analysis Plan(For Database at Week 16)

-List of Errata-

Protocol No.	MT-5547-J01
Protocol Title	A Phase 2/3 (Placebo-Controlled, Double-Blind, Comparative) Study on MT-5547 in Patients with Osteoarthritis Accompanied by Moderate to Severe Pain
Version / Date	2.0 / 2-July-2020

Authors:

Statistics Author	
Print Name:	[redacted]
Position:	Biostatistician

Approved by:

Statistic Approver	
Print Name:	[redacted]
Position:	Responsible Biostatistician
Signature:	Refer to email
Approval date:	Refer to email

[redacted]

[redacted]

[redacted]

Statistical Analysis Plan (For Database at Week 16)

Protocol No. MT-5547-J01

A Phase 2/3 (Placebo-Controlled, Double-Blind, Comparative) Study on MT-5547 in Patients with Osteoarthritis Accompanied by Moderate to Severe Pain

Prepared By:	Mitsubishi Tanabe Pharma Corporation
Version:	1.0
Date:	10-February-2020

APPROVAL FORM

Statistical Analysis Plan

Protocol No.	MT-5547-J01
Protocol Title	A Phase 2/3 (Placebo-Controlled, Double-Blind, Comparative) Study on MT-5547 in Patients with Osteoarthritis Accompanied by Moderate to Severe Pain
Version / Date	1.0 / 10-February -2020

Authors:

Statistics Author	
Print Name:	[redacted]
Position:	Biostatistician

Approved by:

Statistic Approver	
Print Name:	[redacted]
Position:	Responsible Biostatistician
Signature:	[redacted]
Approval date:	[redacted]

TABLE OF CONTENTS

ABBREVIATIONS

Abbreviations	Definitions
AE	Adverse event
AESI	Adverse event of special interest
ALT	Alanine transaminase
ALP	Alkaline phosphatase
ANCOVA	Analysis of covariance
AST	Aspartate transaminase
ATC	Anatomical therapeutic chemical
BDR	Blinded data review
BLQ	Below limit of quantification
BMI	Body mass index
CI	Confidence interval
DA	Destructive arthropathy
DP	Decimal places
DMC	Data monitoring committee
EQ-5D-5L	EuroQol Group Questionnaire
FAS	Full analysis set
ITT	Intent-to-treat
K-L	Kellgren-Lawrence
LLOQ	Lower limit of quantitation
MAR	Missing at random
MedDRA	Medical dictionary for regulatory activities
MMRM	Mixed model repeated measures
NGF	Nerve growth factor
NRS	Numerical rating scale
NSAID	Nonsteroidal anti-inflammatory drug
OMERACT-OARSI	Outcome Measures in Rheumatology Arthritis Clinical Trials-Osteoarthritis Reseach Society International
PGA	Patient Global Assessment
PPS	Per protocol set
PT	Preferred term
q4w	Every 4 weeks
q8w	Every 8 weeks
QOL	Quality of life
RAND	All subjects randomized population
RPOA	Rapidly progressive osteoarthritis
SAE	Serious adverse event
SAF	Safety population
SAP	Statistical analysis plan
SD	Standard deviation
SF-36	Short form 36-item health survey
SIF	Subchondral insufficiency fracture
SOC	System organ class
TEAE	Treatment emergent adverse event

TESAE	Treatment emergent serious adverse events
ULN	Upper limit of normal range
WHO	World Health Organization
WOMAC	Western Ontario and McMaster Universities Osteoarthritis Index

1. INTRODUCTION

This statistical analysis plan (SAP) is based on the final protocol (v4.0) dated 31-July-2018. This protocol will plan to respectively perform the statistical analysis for the clinical trial data from screening to Week 16, from screening to Week 24 and from screening to Week 48.

For the clinical trial data from screening to Week 16, this SAP covers statistical analysis, tabulations of the study data to investigate the efficacy and safety of MT-5547 SC 1mg q4w and 1mg q8w compared to matching placebo. Therefore, this SAP will be finalized prior to database lock at Week16 to ensure the credibility of the study results by pre-specifying the statistical methods for the data analyses before un-blinding of treatment assignments. For the clinical trial data from screening to Week 24 and from screening to post-treatment observation period of Week 20 after the end of the treatment period, the other SAPs will be separately created.

Any statistical analysis details described in this document supersede any description of statistical analysis in the protocol.

2. STUDY OBJECTIVE AND ENDPOINTS

2.1. Study Objective(s)

The objective of this study is to verify the superiority of 16 weeks of MT-5547 treatment to placebo, as evidenced by the WOMAC pain subscale score (the efficacy outcome measure), in patients with osteoarthritis of the knee or hip. Additional objectives of the study are to investigate the efficacy, safety, and pharmacokinetics of MT-5547 in long-term use.

2.2. Study Endopoint(s)

2.2.1. Efficacy Endpoint(s)

(1) Primary efficacy endpoint
WOMAC pain subscale score (change from baseline at Week 16)

(2) Key secondary efficacy endpoint
WOMAC physical function subscale score (change from baseline at Week 16)

(3) Secondary efficacy endpoints
1) Patient global assessment (PGA) (change from baseline at each assessment time point)
2) WOMAC pain subscale score (change from baseline at each assessment time point)
3) WOMAC physical function subscale score (change from baseline at each assessment time point)
4) WOMAC stiffness subscale score (change from baseline at each assessment time point)
5) WOMAC total score (change from baseline at each assessment time point)
6) Proportions of subjects with 30%, 50%, 70% and 90% improvements in their WOMAC pain subscale scores compared to baseline (Weeks 16)
7) Proportions of subjects with 30%, 50%, 70% and 90% improvements in their WOMAC physical function subscale scores compared to baseline (Weeks 16)

8) Numerical Rating Scale (NRS) score for the mean pain on walking in the evaluated joint (change from baseline at each assessment time point)
9) SF-36 (change from baseline at each assessment time point)
10) EQ-5D-5L (change from baseline at each assessment time point)
11) Improvement rate based on the Outcome Measures in Rheumatology Arthritis Clinical Trials-Osteoarthritis Research Society International (OMERACT-OARSI) criteria (at Weeks 16)
12) Amount of rescue medication used (number of days on which rescue medication was used, and amount [number of tablets] used)

2.2.2. Safety Endpoint(s)

(1) Adverse events and adverse reactions
(2) Adverse events of special interest as defined in the protocol i.e. adjudicated arthropathy, sympathetic nervous system disorders and altered peripheral sensation.

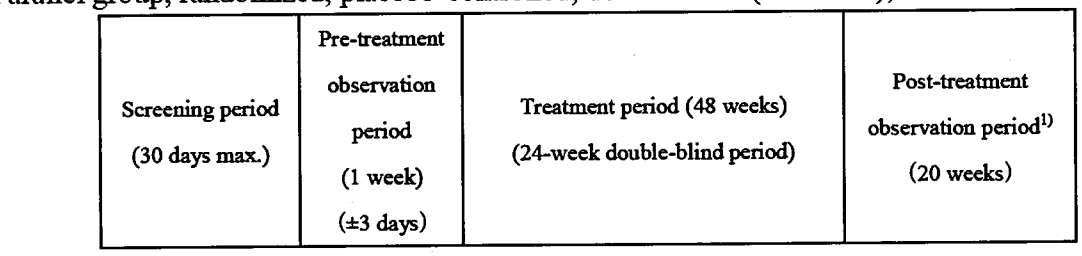
3. STUDY DESIGN

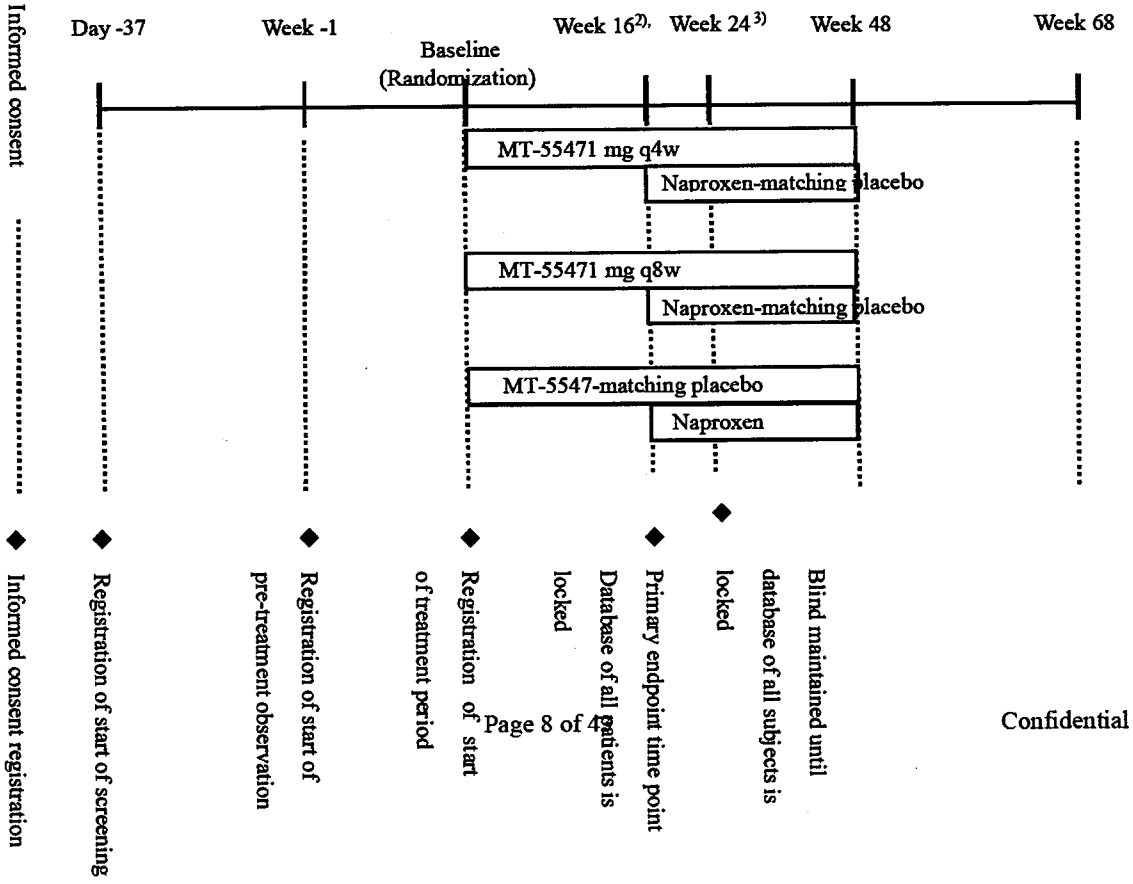
3.1. Study Design

Study Phase: Phase 2/3

Type of Study: Confirmatory Study

Parallel group, randomized, placebo-controlled, double-blind (24 weeks), multicenter





1) The post-treatment observation period will be the 20 weeks from the day after the end of the 48-week treatment period. Furthermore, for subjects discontinued from the study, the post-treatment observation period will be the 24 weeks from the administration of MT-5547 or MT-5547 placebo. If a joint replacement procedure is going to be performed, the postoperative test observation period will be the period from the day after the day of discontinuation until 20 weeks after the joint replacement procedure is performed.
2) From Week 16 to Week 48, the placebo group will receive an additional dose of naproxen tablets to reduce the dropout rate in the placebo group. The MT-5547 group will receive an additional dose of naproxen placebo tablets.
   For each subject, the data for up through Week 16 will be fixed prior to the Week 20 visit, and the fixed data will be prepared.
3) The blind will be maintained, including for the study sponsor, until the key code is opened after the data have been fixed for all subjects in Week 24. In addition, in order to eliminate any bias that could affect the study assessments, to the extent possible, the blind will be maintained with respect to the study site and the subject even after the opening of the study sponsor key code.

3.2. Schedule of Study Procedures

3.2.1. Test/Observation Schedule

(1) Test/Observation Schedule from informed consent to Week 48 of treatment period (if a joint replacement procedure is not going to be performed)

				Treatment period									Treatment period						
Visit times (visit window)	Informed consent	Screening Day -37 ~ -8	Pre-treatment of observation period Day -7 ~ -1 (±3)	Day1 Baseline	Week1 Day 8 (±3)	Week2 Day 15 (±3)	Week4 Day 29 (±7)	Week8 Day 57 (±7)	Week12 Day85 (±7)	Week16 Day 113 (±7)	Week 20 Day 141 (±7)	Week 24 Day 169 (±7)	Week 28 Day 197 (±7)	Week 32 Day 225 (±7)	Week 36 Day 253 (±7)	Week 40 Day 281 (±7)	Week 44 Day 309 (±7)	Week 48 Day 337 (±7)	At withdrawal[16]
Visit Number		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
Written informed consent	X																		
Eligibility confirmed		X	X	X															
Demographics		X		X															
Randomization[1]				X															
Study drug SC administration[2]				X			X	X	X	X	X	X	X	X	X	X	X		
Administration of naproxen										←								→	
Delivery of naproxen										X	X	X	X	X	X	X	X		
Return of naproxen											X	X	X	X	X	X	X	X	X
Delivery of acetaminophen			X	X		X	X	X	X	X	X	X	X	X	X	X	X		
Return of acetaminophen/check of amount taken[3]				X		X	X	X	X	X	X	X	X	X	X	X	X	X	X
Concomitant medications		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
[Efficacy]																			
NRS (average daily pain)[4]			X	←						→									
WOMAC[5]		X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
PGA of OA		X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
SF-36				X			X	X		X		X		X		X		X	X
EQ-5D-5L				X			X	X		X		X		X		X		X	X
[Safety]																			
Body weight (height: only at screening)		X								X								X	X
Vital signs[6]		X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
ECG		X																X	X
Physical examination		X										X		X				X	X
Orthostatic blood pressure		X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X
Injection site assessment (30 minutes after administration)				X			X	X	X	X	X	X	X	X	X	X	X		
Joint Pain Questionnaire		X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Autonomic Nerve Questionnaire		X		X			X	X	X	X	X	X	X	X	X	X	X	X	X
Neurological evaluation		X		X (brief version)		X (brief version)	X (brief version)	X (brief version)	X (brief version)	X (brief version)	X (brief version)	X (brief version)	X (brief version)	X (brief version)	X (brief version)	X (brief version)	X (brief version)	X	X
Radiology (bilateral knees, hips, shoulders)[7]		X								X								X	X
MRI testing (evaluated joint, contralateral joint, knee or hip joints with K-L ≥ 3)[7]		X																	
MRI testing (joint of joint replacement surgery)																			X
Adverse events		←																→	X
At joint pain worsening (imaging [X-ray, MRI] assessments)[8]					←													→	X
JOA Score[9]																			X
Laboratory testing[10]		X		X			X	X	X	X		X		X		X		X	X
Pregnancy test (for WOCBP)[11]		X		X			X	X	X	X	X	X	X	X	X	X	X	X	X
Bone density[12]		X																X	X
Urinalysis		X		X						X		X						X	X
[Pharmacokinetics]																			
PK measurements				X[15]	X	X	X[15]	X[15]	X[15]	X[15]		X[15]		X[15]		X[15]	X[15]	X	X
ADA antibody				X[15]						X[15]		X[15]						X	X
[Other]																			
[redacted]				[redacted]			[redacted]			[redacted]		[redacted]		[redacted]				[redacted]	[redacted]
[redacted]				[redacted]															

(2) Test/Observation Schedule of post-treatment observation period (if a joint replacement procedure is not going to be performed)

	Post-treatment observation period (20 weeks)[14]					
	Week 52	Week 56	Week 60	Week 64	Week 68	At withdrawal
Visit times (visit window)	Day 365 (±7)	Day 393 (±7)	Day 421 (±7)	Day 449 (±7)	Day 477 (±7)	
Visit Number	18	19	20	21	22	
Written informed consent						
Eligibility confirmed						
Demographics						
Randomization[1]						
Study drug SC administration[2]						
Administration of naproxen						
Delivery of naproxen						
Return of naproxen						
Delivery of acetaminophen						
Return of acetaminophen/check of amount taken[3]						
Concomitant medications	X	X	X	X	X	X
[Efficacy]						
NRS (average daily pain)[4]						
WOMAC[5]	X		X		X	X
PGA of OA	X		X		X	X
SF-36					X	X
EQ-5D-5L					X	X
[Safety]						
Body weight (height: only at screening)					X	X
Vital signs[6]	X				X	X
ECG						
Physical examination					X	X
Orthostatic blood pressure	X				X	X
Injection site assessment (30 minutes after administration)						
Joint Pain Questionnaire	X				X	X
Autonomic Nerve Questionnaire	X				X	X
Neurological evaluation	X (brief version)				X	X
Radiology (bilateral knees, hips, shoulders)[7]					X	X
MRI testing (evaluated joint, contralateral joint, knee or hip joints with K-L ≥ 3)[7]						
MRI testing (joint of joint replacement surgery)						X
Adverse events	←				→	X
At joint pain worsening (imaging [X-ray, MRI] assessments)[8]	←				→	X
JOA Score[9]						X
Laboratory testing[10]					X	X
Pregnancy test (for WOCBP)[11]	X				X	X
Bone density[12]						
Urinalysis					X	X
[Pharmacokinetics]						
PK measurements	X	X			X	X
ADA antibody					X	X
[Other]						
[redacted]					[redacted]	[redacted]
[redacted]						

1) Subjects will be randomized (enrolled at the start of the treatment period) after all of the assessments through the pre-treatment observation period have been completed.
2) Investigational drug or comparative drug will be administered subcutaneously after all of the safety assessments, efficacy assessments, and collection of blood samples for pharmacokinetic assessments have been completed, and subjects will be kept under supervision at the study site for 30 minutes after the administration of the study drug.
3) The dose of acetaminophen taken will be checked based on the treatment diary. Every night, subjects will record in their treatment diaries the amount (number of tablets) of acetaminophen they took that day.
4) Every night (and as a rule at the same time) from the start of the pre-treatment observation period until Week 16, subjects will use the IVRS to report the mean pain on walking for the previous 24 hours (NRS data). Subjects will receive instruction in how to use the IVRS on the day of the visit at the start of the pre-treatment observation period.
5) The WOMAC score at screening will be assessed for the right knee joint, the left knee joint, the right hip joint, and the left hip joint, and the evaluated joint will be determined based on the WOMAC pain subscale scores. At the other time points, the WOMAC score will be assessed for the evaluated joint only.
6) If the pulse is less than 45 bpm at an assessment time point after investigational drug or comparative drug administration, electrocardiography will be performed and the patient carefully examined for cardiac function abnormalities.
7) Screening:
The following imaging tests will be performed for subjects who meet the study eligibility criteria, and the images will be submitted to a central laboratory. Imaging tests will also be performed if the central laboratory requests that additional tests be conducted.
   - X- ray tests: Both knee joints, both hip joints, both shoulder joints
   - MRI tests: The evaluated joint, the joint opposite the evaluated joint, and any knee or hip joints with a K-L grade of 3 or more

MRI tests will be performed as needed by the (sub)investigator.
A check will be performed by the day specified below to make sure that the results of the central assessments meet the study enrollment criteria.
   - X-ray tests: By the start of the pre-treatment observation period
   - MRI tests: By baseline

If the results of the central assessments cannot be confirmed by the specified day, then re-screening will be performed.
Baseline and after baseline:
At week 16, 48 (or at the early termination of the treatment period) and at 68 weeks (or at the early termination of the post-treatment observation period), images will be submitted to the central laboratory. If the (sub)investigator determine that there is an abnormal finding compared to normal osteoarthritis, MRI tests will be performed as needed. The images will be submitted to a central laboratory. Imaging tests will also be performed if the central laboratory requests that additional tests be conducted, and the images submitted to the central laboratory(for more detailed information, see in the text "9.2.4.1.11 Imaging Tests").
Week 16:
The X-ray tests (both knee joints, both hip joints, both shoulder joints) will be performed by the (sub)investigator. The (sub)investigator will check the X-ray images before administered the investigational drug or comparative drug to determine whether or not investigational drug or comparative drug should be administered. If any abnormalities are found, investigational drug or comparative drug administration will be postponed. If the central laboratory rules out adjudicated arthropathy, investigational drug or comparative drug administration will be resumed.
Assessment at the early termination of the treatment period and the early termination of the post-treatment observation period
The X-ray tests (both knee joints, both hip joints, both shoulder joints) will be performed by the (sub)investigator. The images will be submitted to a central laboratory. If more than 30 days have passed since the last images were taken, images will be taken again. If not more than 30 days have passed, imaging tests will be performed if the (sub)investigator determines that it is necessary.

If joint replacement surgery is required, the MRI tests of the joint of surgery will be performed (for more detailed information, see in the text "9.2.4.1.13 Assessments if a Joint Replacement Procedure Is Scheduled").

8) If a subject experiences a sudden worsening of pain that would not occur in the normal course of osteoarthritis, the (sub)investigator will postpone investigational drug or comparative drug administration, and perform imaging test (X-ray and MRI) of the affected joint. Imaging test (X-ray and MRI) of other joints will be performed as needed by the (sub)investigator. Imaging tests will also be performed if the central laboratory requests that additional tests be conducted. All images will be submitted to the central laboratory. The results of the central assessment will be checked by the (sub)investigator. The investigational drug or comparative drug administration will be postponed until the central laboratory rules out adjudicated arthropathy. If the result of the central assessment is adjudicated arthropathy, then study treatment must be discontinued. If adjudicated arthropathy is ruled out by the central assessment, the (sub)investigator will determine whether or not study treatment should be continued.
9) Assessmets of the joint by JOA score at discontinuation will be performed for subjects who required joint replacement surgery during the study.
10) Hematology and blood biochemistry tests will be performed. HbA1c, HBs antigen, HBs antibody, HCV antibody, and HIV antibody will be measured only at screening.
11) At screening, the pregnancy test will be performed using serum. At all other time points, the pregnancy test will be performed using urine. Furthermore, pregnancy tests will not be performed for postmenopausal females who have been amenorrheic for at least 1 year or for females who have undergone a surgical hysterectomy or bilateral ovariectomy. If a positive result is obtained on a urine pregnancy test, a serum pregnancy test will be performed. If a negative result is obtained on a serum pregnancy test, the study can be continued for the patient. If the serum pregnancy test is also positive, study treatment will be discontinued.
12) Bone density will be measured only at those study sites that are capable of performing measurements using the DEXA method. If it can not be measured at screening priod, it may be measured during the pre-treatment observation period.
13) [redacted]
14) The Week 64 assessments may be performed by telephone.
15) The blood samples will be collected before the investigational drug or comparative drug is administered.
16) In the case of early termination of the treatment period, follow-up evaluation of the same item as at the time of visit of the corresponding completed subjects below is carried out 4, 8, 12, 16, 20 and 24 weeks after the last dose of the investigational drug or comparative drug. However, blood sample collection for pharmacokinetic evaluation is performed only 24 weeks after the last dose (blood sampl collection for PK measurement is not necessary 4, 8 and 12 weeks after the the last dose corresponding to treatment period 48, 52 and 56 weeks of the completed subjects). The visit window at each evaluation time point is plus or minus 7 days.

Early termination of the treatment period	4 weeks after the last dose	8 weeks after the last dose	12 weeks after the last dose	16 weeks after the last dose	20 weeks after the last dose	24 weeks after the last dose	the early termination of the post-treatment observation period
Completed of the treatment period	Week 48	Week 52	Week 56	Week 60	Week 64	Week 68	the early termination of the post-treatment observation period

When assessments at discontinuation is included in visit window of each assessment time point, the same ones as those normally assessed at discontinuation. After assessment at discontinuation, the same ones as those normally assessed at the last dose of investigational drug or comparative drug.
If more than 30 days have passed since the last images were taken, images will be taken again. If not more than 30 days have passed, imaging tests will be performed if the (sub)investigator determines that it is necessary.
The tests at discontinuation and post-joint replacement procedure test/observation will be performed for

subjects who required joint replacement surgery during the study. The all images will be submitted to a central laboratory (for more detailed information, see "10 (3) Post-joint replacement procedure test/observation schedule). If the assessments at discontinuation are not performed before the joint replacement procedure is performed, then images from before the procedure will be obtained and sent to the central laboratory.

17) The screening period is from "the screening start date" to the day before "the pre-treatment of observation period start date", and the screening start date and the end date are defined along with the allowable range of "the pre-treatment of observation period" (for example, when the pre-treatment of observation period is 4 days, The maximum screening period is 30 days from Day - 34 to Day - 5, the pre-treatment of observation period is 10 days, the screening period is the maximum from Day - 40 to Day - 11).

At each study visit after the baseline date, the assessments that are performed by the subjects themselves (WOMAC, PGA, joint pain survey, SF-36, EQ-5D-5L, survey of autonomic symptoms) will be performed before all of the other assessments (including the assessments that are performed by the (sub)investigator) (except for cases where imaging tests are performed on a different day).

(3) Post-joint replacement procedure test/observation schedule

Follow-up day (visit window)	After surgery[1]	Long-term[1]
	Follow-up survey, 4 weeks post-operative	Follow-up survey 2, 20 weeks post-operative
	29 days post-operative (±5 days)	141 days post-operative (±7 days)
Concomitant medications	X	X
[Safety]		
Vital signs	X	X
Joint Pain Questionnaire	X	X
JOA score[2]	X	X
Radiology (bilateral knees, hips, shoulders)	X	X
Adverse events	←	→
At joint pain worsening (imaging [X-ray, MRI] assessments)3	←	→

1) Information about the procedure, including prosthesis replacement and/or the extent of surgery wound healing, will be collected.
2) The condition of the joint following joint replacement surgery will be assessed using the JOAScore.
3) If a sudden worsening of pain that does not occur in the normal progression of osteoarthritis occurs, the (sub)investigator will perform imaging test (X-ray and MRI) of the affected joint. Imaging test (X-ray and MRI) of other joints will be performed as needed by the (sub)investigator. Imaging tests will also be performed if the central laboratory requests that additional tests be conducted. All images will be submitted to the central laboratory.

3.3. Sample Size and Power Considerations

The original target sample size is 568 treated subjects (142 subjects per group)

(MT-5547 1 mg q4w group: 142 subjects; MT-5547 3 mg q4w group: 142 subjects; MT-5547 6 mg q8w group: 142 subjects; placebo group: 142 subjects)

After initiation of this study, based on the recommendations from the IDMC, the 3 mg q4w and 6 mg q8w groups were removed and 1 mg q8w group was newly added to the treatment

groups.

The amended target sample size is 506 treated subjects, including the subjects who were randomized to 1 mg q4w/placebo groups before the amendment (approximately 40 subjects per group) and the subjects who are randomized to 1 mg q4w/1 mg q8w/placebo groups with the ratio of 1:1:1 after the amendment (142 subjects per group).

(MT-5547 1 mg q4w group: 182 subjects; MT-5547 1 mg q8w group: 142 subjects; placebo group: 182 subjects)

In addition, 57 patients each were enrolled in the MT-5547 3 mg q4w and 6 mg q8w groups prior to the protocol amendment (Ver. 04.00.00000) and were discontinued from study drug based on the recommendations from the IDMC. Thus, enrollment of 563 subjects is planned for this study in total.

The original rationale at the initiation of the study

Because this study is being conducted as a placebo-controlled study, in order to detect a statistical difference between the MT-5547 groups and the placebo group, the mean intergroup difference relative to placebo in both the change from baseline in the WOMAC pain subscale score and the change from baseline in the physical function subscale score at Week 16 for the MT-5547 3 mg SC q4w, 6 mg SC q8w, and 1 mg SC q4w groups was set at 1.1, and the standard deviation 2.5, on the basis of the results obtained in the overseas P2a study (R475-PN-0901) and the overseas P2b study (R475-PN-1227) (Table 3.3-1). The effect relative to placebo in the P3 study of 5 mg and 10 mg of tanezumab, another NGF inhibitor, was 0.81 to 1.21 in the WOMAC pain subscale score and 0.98 to 1.25 [24), 25)] in the WOMAC physical function subscale score, and it therefore appears that a mean difference between the MT-5547 groups and placebo of 1.1 is a realistic assumption. It also appears that the minimum clinically significant difference in each WOMAC score will be around 0.67 to 0.75 [26), 27)]. The aforementioned intergroup difference of 1.1 exceeds the minimum clinically significant difference, and therefore seems to be an appropriate value to use for the intergroup difference. Based on a one-sided significance level of 2.5% and multiplicity adjustment based on a gate keeping method using a graphical approach, the power for MT-5547 3 mg SC q4w, 6 mg SC q8w, and 1 mg SC q4w will be 90%, 88%, and 88%, respectively, for the primary endpoint, and 82%, 79%, and 79%, respectively, for the key secondary endpoint. A 2-sample t-test will be used for the tests in each step. Based on this assumption, the number of subjects who will complete 16 weeks of treatment in each group will be 120.

Furthermore, because the proportion of missing data for the primary endpoint and the key secondary endpoint at Week 16 is around 15% according to Table 3.3-1, it is being assumed that the proportion of missing data at Week 16 will be 15%, and the number of patients per

group is therefore being set at 142.

The rationale after the changes in treartment groups

Assuming the absolute treatment difference of 1.1 between the MT-5547 1 mg SC q8w groups and placebo in both the change from baseline in the WOMAC pain subscale score and the change from baseline in the physical function subscale score at Week 16 with an associated standard deviation of 2.5 based on the original rationale at the initiation of the study, the number of patients per group after the protocol amendment (Ver. 04.00.00000) is planned to 142 subjects per group. The structure of statistical hypotheses have been changed due to the removal of 3 mg q4w / 6 mg q8w groups and the addition of 1 mg q8w. It is noted that approximately 40 subjects per group were randomized to 1 mg q4w / placebo groups before the amendment. For evaluation on 1 mg q4w, since subjects are randomized to 1 mg q4w / placebo at the same time with the same ratio of 1:1 regardless of pre- or post-amendment, analyses will be performed with pooled comparable data including pre- and post-amendment data on 1 mg q4w / placebo (1 mg q4w group: 182 subjects; placebo group: 182 subjects). For evaluation on 1 mg q8w, analyses will be performed with only the post-amendment data (1 mg q4w group: 142 subjects; 1 mg q8w group: 142 subjects; placebo group: 142 subjects). For four null hypotheses, based on multiplicity adjustment based on a gate keeping method using a graphical approach (Figure 8.2.1-1), the power for MT-5547 1 mg SC q4w and 1 mg SC q8w will be 97% and 92% for the primary endpoint, 94% and 87% for the key secondary endpoint, respectively.

Table 3.3-1Change From Baseline at Week 16 in the WOMAC Pain and Physical Function Subscale Scores

	REGN Study R475-PN-1227					REGN Study R475-PN-0901			
	Placebo (N=83)	MT-5547 1 mg (N=85)	MT-5547 3 mg (N=84)	MT-5547 6mg (N=85)	MT-5547 9 mg (N=84)	Placebo (N=55)	MT-5547 0.03 mg/kg (N=53)	MT-5547 0.1 mg/kg (N=53)	MT-5547 0.3 mg/kg (N=54)
Change from baseline in WOMAC pain subscale score									
Baseline	6.43	6.33	6.35	6.10	6.53	5.9	5.7	6.1	6.4
n[1)] Mean[1)] (SD)[1)]	71 -2.43 (2.38)	75 -3.49 (2.06)	78 -3.39 (2.44)	77 -3.07 (2.34)	79 -3.81 (2.49)	44 -2.4 (2.18)	47 -2.7 (1.89)	44 -3.4 (2.53)	41 -3.2 (2.24)
LSM[2)]	-2.25	-3.35	-3.33	-3.03	-3.65	-	-	-	-
Diff. vs. Pbo[3)]	-	-1.10	-1.08	-0.78	-1.40	-	-0.6	-1.1	-0.8

P value	-	0.0025	0.0029	0.0304	0.0001	-	0.1486	0.0090	0.0488
Change from baseline in WOMAC physical function subscale score									
Baseline	6.15	6.11	6.09	5.94	6.20	5.9	5.9	6.2	6.2
n[1)] Mean[1)] (SD)[1)]	70 -2.12 (2.26)	75 -3.21 (2.23)	78 -3.28 (2.29)	76 -2.97 (2.45)	80 -3.51 (2.50)	44 -2.3 (2.30)	47 -2.9 (1.78)	44 -3.4 (2.28)	41 -3.1 (2.18)
LSM[2)]	-1.98	-3.08	-3.27	-3.03	-3.41	-	-	-	-
Diff. vs. Pbo[3)]	-	-1.10	-1.29	-1.06	-1.43	-	-0.8	-1.1	-0.9
P value	-	0.0019	0.0003	0.0029	< 0.0001	-	0.0693	0.0071	0.0245

1) n: number of subjects at Week 16; mean: arithmetic mean at Week 16; SD: standard deviation at Week 16. However, in REGN Study R475-PN-1227, the n includes WOMAC data from after treatment discontinuation.

2) LSM: MMRM least squares mean

3) Diff. vs. Pbo: Difference for each group relative to the placebo group in the MMRM least squares mean

4. PLANNED ANALYSIS

4.1. Interim Analysis

No interim analysis was planned and hence no multiplicity adjustment for interim analysis is needed.

4.2. Final Analysis

This SAP for 16 week will be finalized before database lock in week 16. The efficacy analysis performed at week 16 is the final efficacy analyasis. The analysis for week 16 will be conducted using database up to 16 week after database lock in week 24.

4.3. Independent Data Monitoring Board

An independent data monitoring board will periodically assess unblinded data, and will make recommendations to the study sponsor regarding the appropriateness of continuing the study and about the need for protocol amendments.

The independent data monitoring board will include independent statistical and medical experts. For more details, the overseas independent data monitoring committee procedures (DMC Charter) will be followed.

5. ANALYSIS POPULATIONS

The analysis of efficacy will be performed in the full analysis set (FAS). The principal

analysis set for efficacy will be the modified FAS (mFAS). Secondary analyses of the primary efficacy endpoint and key secondary efficacy endpoint will also be performed in the per-protocol set (PPS). The analysis of safety will be performed in the safety analysis set (SAF). The analysis of pharmacokinetics will be performed in the pharmacokinetic analysis set.

The analysis sets are defined below. For the contents of the data review meeting, refer to a separate materials.

(1) Efficacy analysis sets

1) Randomized subjects (RAND)

2) FAS

The FAS will consist of all randomized subjects except for the following subjects.

- Subjects who have not received study drug even once
- Subjects with no WOMAC pain subscale score (the primary endpoint) at baseline
- Subjects with no WOMAC pain subscale score (the primary endpoint) at any time point after randomization

3) Modified FAS (mFAS)

The mFAS will consist of the FAS, except for the following subjects.

- Subjects who are randomized to 3 mg q4w group
- Subjects who are randomized to 6 mg q8w group

4) PPS (This set will be used for only statistical analysis for the data obtained from screening to Week 16)

The PPS will consist of the mFAS, except for the following subjects.

- Subjects who did not meet the inclusion criteria
- Subjects who met the exclusion criteria
- Subjects who did not comply with the rules on prohibited concomitant drugs/therapies
- Subjects who did not complete the period from study treatment initiation through the assessment of the WOMAC pain subscale score (the primary endpoint) at Week 16
- Subjects who did not received four injections of study drug

(2) SAF

The SAF will consist of all randomized subjects except for the following subjects

- Subjects who did not receive study drug even once
- Subjects for whom absolutely no post-randomization safety data are available

6. STATISTICAL CONSIDERATIONS

6.1. Descriptive Statistics

(1) Non-PK related

Continuous data will be summarized descriptively using the number in the analysis set (N), the number of observations (n), mean, standard deviation (SD), median, minimum and maximum. Categorical data will be summarized using frequency counts and percentages. The denominator for the percentages will be the total number of subjects in the treatment group and analysis population being presented, unless otherwise specified.

6.2. Statistical Tests

Unless otherwise specified, all formal statistical tests of treatment effects will be done at two-sided significance level of 0.05. Point estimates will be accompanied with two-sided 95% CIs where applicable.

7. DATA CONVENTIONS

7.1. Analysis Variable Definitions

7.1.1. Study Subjects

7.1.1.1. Demographic and Other Baseline Characteristics

(1) BMI

BMI will be recalculated using the formula below and reported to 1dp.

BMI (kg/m^2) = weight at screening (kg) / {height at screening (m)}^{2}

(2) Age(years)

If (month of date of consent > month of date of birth) or (month of date of consent = month of date of birth and day of date of consent >= day of date of birth), then Age (years) = year of date of consent - year of date of birth

if else case, then Age (years) = year of date of consent - year of date of birth –1.

(3) Time since osteoarthritis diagnosis (years)

If (month of date of consent > month of osteoarthritis diagnosis) or (month of date of consent = month of date osteoarthritis diagnosis, and day of date of consent >= day of date of osteoarthritis diagnosis), then Time since osteoarthritis diagnosis (years) = year of date of consent – year of date of osteoarthritis diagnosis

if else case, then Time since osteoarthritis diagnosis (years) = year of date of consent –year of date osteoarthritis diagnosis – 1.

Impute date of osteoarthritis diagnosis as the earliest possible date (i.e. first day of month if day is unknown or 1st January if day and month are unknown).

7.1.1.2. Treatment Duration and Compliance

(1) Treatment Duration

1) Treatment Duration (days) = The date of last study drug injection prior to week 16 - the date of the first study drug injection + min(28 or (the date of last study drug infection - the date of Week 16 study drug injection))
2) Criteria for pre-defined limit
   Treatment duration period ≥1 day, ≥29 days, ≥57 days, ≥85 days, and ≥113 days

(2) Treatment Compliance

1) Treatment compliance will be calculated using the formula below and reported to 1dp.

Treatment compliance(%) = (Number of injections of study drug during treatment exposure period) / (Number of planned injections of study drug during exposure period on or before the time that the subject discontinues from the study) x 100%

2) Criteria for pre-defined limit

the number of injections of study drug: 1, 2, 3 and 4

7.1.2. Efficacy assessments

7.1.2.1. WOMAC pain subscale score

$$\text{WOMAC pain subscale score} = \frac{\text{sum of items from Section A 1 to Section A5 in questionnaire}}{5}$$

7.1.2.2. WOMAC physical function subscale score

$$\text{WOMAC physical function subscale score} = \frac{\text{sum of items from Section C 8 to Section C24 in questionnaire}}{17}$$

7.1.2.3. WOMAC stiffness subscale score

$$\text{WOMAC stiffness subscale score} = \frac{\text{Section B 6} + \text{Section B7 in questionnaire}}{2}$$

7.1.2.4. WOMAC total score

$$\text{WOMAC total score} = \frac{\text{sum of items from Section A1 to Section C24 in questionnaire}}{24}$$

7.1.2.5. NRS score

- Weekly average walking index joint pain
  Baseline is defined as the average of the non-missing values during 7 days prior to taking study drug. For each week, the average of the non-missing values during the 7 days on or prior to that week will be used.
- Daily walking index joint pain
  Baseline is defined as the last non-missing values prior to taking study drug.

7.1.2.6. SF-36 / EQ-5D-5L

- SF-36
  Physical, Mental and Role social component summary will be calculated considering factor coefficients based on the 1995 Japan National Survey. The derivation of 8 subscales and three component is detailed in appendix 13.1.
- EQ-5D-5L
  Utility Index Score will be calculated using the formula and Table below.[1]

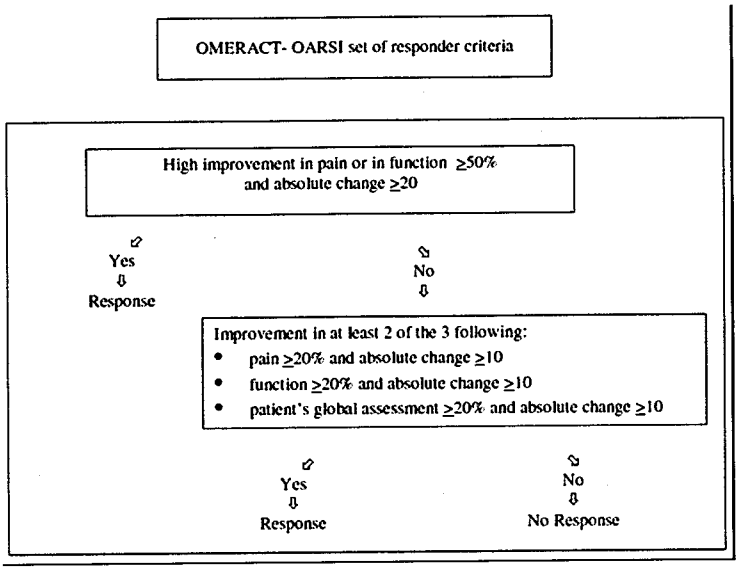
Utility Index Score
$= 1 + [Estimation\ of\ the\ constant\ term(if\ all\ answer\ is\ not\ 1)]$
$+ [Sum\ of\ "estimation\ of\ coefficients\ corresponding\ to\ answer\ other\ than\ 1"]$

Item	Level	Estimation	Standard Error	P-value
Intercept		−0.060924	0.013625	<0.0001
Mo	2	−0.063865	0.008996	<0.0001
	3	−0.112618	0.009287	<0.0001
	4	−0.179043	0.010231	<0.0001
	5	−0.242916	0.009425	<0.0001
Sc	2	−0.043632	0.008931	<0.0001
	3	−0.076660	0.009972	<0.0001
	4	−0.124265	0.010129	<0.0001
	5	−0.159659	0.008924	<0.0001
Ua	2	−0.050407	0.009205	<0.0001
	3	−0.091131	0.010005	<0.0001
	4	−0.147929	0.009744	<0.0001
	5	−0.174786	0.009115	<0.0001
Pd	2	−0.044545	0.008354	<0.0001
	3	−0.068178	0.010052	<0.0001
	4	−0.131436	0.008985	<0.0001
	5	−0.191203	0.009604	<0.0001
Ad	2	−0.071779	0.009701	<0.0001
	3	−0.110496	0.010863	<0.0001
	4	−0.168171	0.009850	<0.0001
	5	−0.195961	0.009164	<0.0001

Mo: mobility, Sc: self-care, Ua: usual activities,
Pd: pain/discomfort, Ad: anxiety/depression

e.g) If answer is [1 1 1 1 1], Utility Index Score is $= 1 + 0 + 0 = 1$. If answer is [1 1 3 2 1], Utility Index Score is $= 1 + (-0.060924) + (-0.091131 - 0.068178) = 0.779767$.

7.1.2.7. Response rate based on the OMERACT-OARSI criteria



Note that the criteria in the diagram above are based on standardized score between 0 and 100. For this study, WOMAC pain and physical function subscale score are between 0 and 10, so the absolute change required for response is the required change in the diagram above divided by 10; PGA is 1, 2, 3, 4 or 5, so the absolute change required for response is at least 1 point.

7.1.2.8. Rescue Medication

Baseline amount of rescue medication use is defined as the average of the non-missing values during 7 days prior to taking study drug. For each week, the average of the non-missing values during the 7 days on or prior to that week will be used.

7.1.3. Safety Assessments

The on-treatment period is defined as the time from first dose of investigational product up to 28 days after the last dose of investigational product excluding adverse events after study drug sc administration at week 16.

7.1.3.1. Adverse Events

Adverse events will be coded according to the MedDRA version 20.1

(1) Treatment Emergent Adverse Events/ Treatment Emergent Serious Adverse Events (TEAEs/TESAEs)

An AE/SAE is classified as treatment emergent if it newly occurred after the first dose of study drug or if a pre-dose event increases in severity following the first dose of study drug.

(2) Adverse Drug Reaction

A TEAE is considered "adverse drug reaction" if it has been assessed as having a "reasonable

possibility" in relationship to the study drug.

(3) Adverse Event of Special Interest(AESI)

Adjudicated arthropathy, sympathetic nerve disorders and altered peripheral sensation are defined as adverse events of special interest (AESI) to facilitate characterizing the safety profile of MT-5547.

(4) Duration of Adverse Events

Duration of Adverse Events (days) = AE stop date – AE start date + 1

7.2. Analysis Visit Definitions

(1) General analysis visit definitions

Analysis visit	Nominal day	Window
Baseline	Day 1	<Day 1
Week 1	Day 8	Day 2 to 11
Week 2	Day 15	Day 12 to 21
Week 4	Day 29	Day 22 to 43
Week 8	Day 57	Day 44 to 71
Week 12	Day 85	Day 72 to 99
Week 16	Day 113	Day 100 to 127

(2) For NRS analisys visit definitions

Analysis visit	Window
Baseline	Day -7 to -1
Week 1	Day 1 to 7
Week 2	Day 8 to 14
Week 3	Day 15 to 21
Week 4	Day 22 to 28
Week 5	Day 29 to 35
Week 6	Day 36 to 42
Week 7	Day 43 to 49
Week 8	Day 50 to 56
Week 9	Day 57 to 63
Week 10	Day 64 to 70
Week 11	Day 71 to 77
Week 12	Day 78 to 84
Week 13	Day 85 to 91
Week 14	Day 92 to 98
Week 15	Day 99 to 105
Week 16	Day 106 to 112

The date of the first dose of study drug is defined as Day 1.

Unless otherwise specified, baseline will be the last observed value of the parameter of

interest prior to the first intake of study drug (this includes unscheduled visits). For other visits, if there are multiple data in a window, the closest data to nominal day will be used. If the distance to the nominal day is the same, the data of later date will be used.

7.3. Data Handling Convention for Missing Data

(1) Non-PK related

Efficacy:

WOMAC

WOMAC scores will be computed when one pain, one stiffness, or 1-3 physical function items are missing. The missing items will be imputed by the mean of available items within the same subscale. The scores will be set to missing if more items are missing.

NRS

If all values are missing for the 7 days, the value for that week is set to missing.

SF-36 / EQ-5D-5L

- SF-36 subscale scores will be computed if at least 50% of items are available. The missing items will be imputed by the mean of available items.
- EQ-5D-5L index will be set to missing if any of the 5 dimensions is missing.

Adverse events:

If severity or relationship is found to be missing the most severe occurrence will be imputed for the summary of interest.

For AE start missing or partial dates, the AE will be treated as TEAE if it cannot be determined to be a non-TEAE.

8. STATISTICAL METHODOLOGY

8.1. Study Subjects

8.1.1. Subject Disposition

Subject disposition will be summarized by treatment group, combined MT-5547 and overall on the randomized population.

8.1.2. Analysis Populations

Analysis populations will be summarized by treatment group, combined MT-5547 and overall on the randomized population.

8.1.3. Demographic and Other Baseline Characteristics

The following demographic and other baseline characteristics will be used.

	category	Descriptive
Sex	Male, Female	
Age(years)	<65, >=65	Yes
Height(cm)		Yes
Weight(kg)		Yes
BMI(kg/m2)	<25, >= 25 and < 30, >=30	Yes
Race	Japanese, other	
Time since OA diagnosis (years)	<1, >=1 and <5, >=5 and <10, >=10	Yes
Index Joint	Knee, Hip	
K-L score for index joint	1, 2, 3, 4	
WOMAC pain subscale score	<6, >=6	Yes
WOMAC Physical function subscale score	<6, >=6	Yes
PGA	1, 2, 3, 4, 5	Yes
Pain site	Right Knee, Left Knee, Right Hip, Left Hip, Right Shoulder, Left Shoulder	
Pretreatment drug	NSAIDs(Systemic use), NSAIDs(Topical use), Acetaminophen, Opioid, Monoamine reuptake inhibitor, Hyaluronic acid, Pregabalin, Herbal medicine, Other	
NSAIDs Inadequate Pain Relief	Yes, No	
NSAIDs Intolerance	Yes, No	
Ever taken Opioid for Pain due to OA	Yes, No	
Opioid Inadequate Pain Relief	Yes, No	
Opioid Intolerance	Yes, No	
Ever taken Acetaminophen for Pain due to OA	Yes, No	
Acetaminophen	Yes, No	

Inadequate Pain Relief		
Acetaminophen Intolerance	Yes, No	
NSAID + Opioid + Acetaminophen Inadequate Pain Relief	Yes, No	

Demographic and other baseline characteristics will be summarized by treatment group, combined MT-5547 and overall on the FAS population, PPS population and the SAF population.
Demographic and other baseline characteristics before the protocol amendment and after the protocol amendment will be summarized by treatment group, combined MT-5547 and overall on the mFAS population.

8.1.4. Treatment Duration and Compliance

Treatment duration will be summarized by treatment group and combined MT-5547 on the SAF population. Treatment compliance will be summarized by treatment group and combined MT-5547 on the SAF population.

8.2. Efficacy Assessments

Efficacy assessment will be summarized by treatment group on the mFAS population.

8.2.1. Primary Efficacy Endpoint

The change in the WOMAC pain subscale score from baseline to Week 16 will be summarized and plotted by treatment group.

Primary analysis method

The primary analysis will be tested using the estimator of mixed effect model for repeated measurement (MMRM) approach. The model will include the treatment group, measurement time point, treatment group and measurement time point interaction, stratification factor K-L grade (2-3 or 4), and stratification factor evaluated joint (knee or hip) as factors, and the baseline and baseline and measurement time point interaction as covariates. The within-subject correlation structure will be unstructured, and the denominator degrees of freedom for the test statistics of estimated parameters will be approximated using the Kenward-Roger method. In case the model will not converge with the unstructured covariance structure, Heterogeneous Autoregressive (1) (ARH[1]), Heterogeneous Compound Symmetry (CSH), Autoregressive (1) (AR[1]) or Compound Symmetry(CS) covariance structure will be used instead in this order.

- Evaluation on 1 mg q4w group

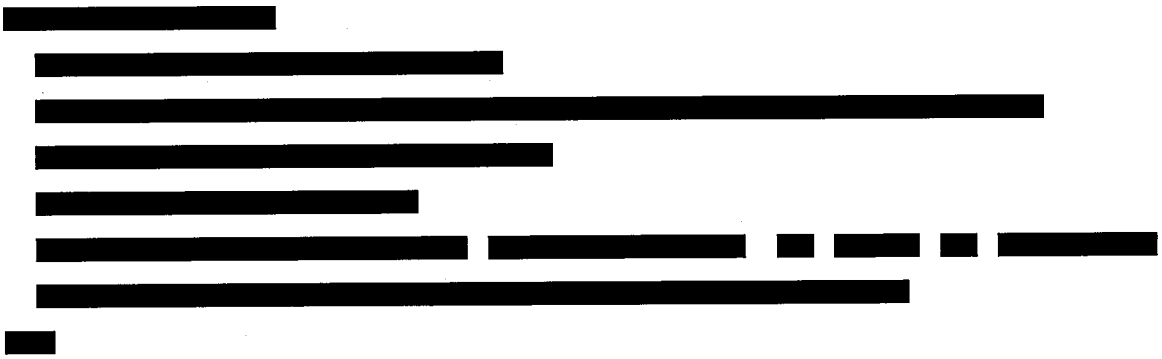
With pooling the data on 1 mg q4w and placebo obtained before and after the protocol amendment (Ver. 04.00.00000), the least square mean, standard error, P value, and 95% confidence interval of the difference between each of the MT-5547 1 mg q4w and the placebo group in the change in the WOMAC pain subscale score from baseline to Week 16 will be calculated using the MMRM approach.

- Evaluation on 1 mg q8w group

Using only the data obtained after the protocol amendment (Ver. 04.00.00000), the least square mean, standard error, P value, and 95% confidence interval of the difference between each of the MT-5547 1 mg q4w, MT-5547 1 mg q8w and the placebo groups in the change in the WOMAC pain subscale score from baseline to Week 16 will be calculated using the MMRM approach.

MMRM that assumes that the missing mechanism is missing at random (MAR) is an analysis method in which missing values are not replaced. This model will be used to estimate the treatment effect in the mFAS when subjects continue receiving treatment through Week 16.

The SAS code planned for the analysis is outlined below.



Sensitivity analyses

Sensitivity analysis will use the same estimator as for the primary analysis. Likelihood based model method on the mFAS under Missing Not at Random(MNAR) assumption will be performed using the same MMRM as specified for the primary analysis. The following additional analyses will be performed for the primary efficacy endpoint.

- **Multiple Imputation (MI) analysis by Reason for Discontinued Treatment**

The missing data for patients who discontinued treatment due to lack of efficacy or AEs will be imputed with values centered at the mean baseline value of the treatment group that patients were randomized to. The missing data for patients discontinued treatment due to other reasons will be imputed under missing-at-random assumption using the regression method with adjustment for covariates including treatment group, randomization strata, and baseline score. Intermittently missing data will be imputed using Markov Chain Monte Carlo method.

The primary efficacy variables will be analyzed using a multiple imputation approach with mixed-effect model for repeated measure (MMRM) based on the mFAS with adjustment for missing data due to treatment discontinuation for the reasons of lack of efficacy or AEs assuming the WOMAC scores would on average return to baseline values. Missing data up to week 16 will be imputed 50 times to generate 50 complete data sets by using the SAS procedure PROC MI following the 4 steps below:

Step 1: This methodology will structure data based on missing data patterns. The Pattern Mixture Model (PMM) method will be based on a missingness pattern having a monotone structure, i.e. if among the observations over time one data value is missing, all other values after this missing value will also be missing. For patients with intermittent missing values, before performing MI based on the PMM, it will be necessary to create a monotone missingness pattern. Intermittent missing values will be imputed using the Markov Chain Monte Carlo (MCMC) methodology which assumes a multivariate normal distribution over all variables included in the imputation model. The MI procedure in SAS will be used for this purpose and this first MI step is planned to be repeated 50 times, creating different datasets with a monotone missing data structure. Seed value of 1995 will be used in the MI procedure. The imputation is based on the MAR assumption, i.e. the missing data are assumed to follow the same model as the other patients in their respective treatment group.

The following SAS code will be used to generate the monotone missing data pattern:

[redacted]

[redacted]

Step 2: After this, the remaining missing data will be imputed using a method for monotone missingness, also based on the MAR assumption. Thus, for each of the created dataset with a monotone missing data pattern, the MI procedure in SAS will be used to impute missing values within the instrument range based on a sequential procedure reflecting the monotone missing data pattern. Patients with the first missing value occurring at Week 1 ('W1') will have their missing Week 1 value replaced by an imputed value from a regression model with treatment group ('TRTP'), K-L grade ('GRADE') , index joint ('INDEX') and baseline ('BASE') value as explanatory variables. In the next step, patients with their Week 2 ('W2') value missing will have their missing Week 2 value replaced by an imputed value from a regression model with treatment group, grade, index and baseline and the Week 1 value as explanatory variables. Similar procedure will be used to replace the missing values at Week 4, Week 8, Week 12 and Week 16.

The following SAS code will be used to make the imputation with the MAR assumption:

[redacted]

Step 3: The initially missing and now imputed data for patients discontinued from the study treatment due to lack of efficacy or AEs will be adjusted to be centered at the mean baseline value for that treatment group, i.e., final imputed score = imputed score under MAR - (mean change from baseline score at the post-baseline time point* for the treatment group based on patients on treatment with non-missing data at that time point).

* MMRM models similar to that described for the primary analysis will be run on each of the 50 generated imputed datasets and the differences between the treatment groups at the

post-baseline time point.

Step 4: The imputed datasets generated with the approach described above do contain only non-missing values and are used as input in the model for the sensitive analysis for the primary efficacy endpoint. Any score imputed outside the range of the WOMAC subscale score of 0-10 will be truncated to the nearest permissible value on the WOMAC scale. MMRM models similar to that described for the primary analysis will thus be run on each of the 50 generated imputed datasets and the differences between the treatment groups at Week 16 of the DBDD Maintenance period will be estimated (and export to data 'ESTIMATES'). Finally, the MIANALYZE procedure in SAS will be applied to combine the results from these analyses to derive an overall estimate of the treatment differences at W16 according to the following code. In addition to the estimates, corresponding 95% confidence intervals and p-values will be calculated.

[redacted]

- **Tipping Point Analysis with MNAR assumption**

Sensitivity analysis using a tipping point approach with multiple imputation will be performed to assess the robustness of the results due to data that may be missing not-at-random (MNAR).

This approach will introduce a sensitivity parameter, δ. Estimations will be performed using multiple imputation methodology. Missing data up to week 16 time point will be imputed 50 times to generate 50 complete datasets by using SAS procedure PROC MI for each δ following the 3 steps below:

Step 1 and 2: The same as step 1 and 2 for Multiple Imputation (MI) analysis by Reason for Discontinued Treatment. The imputation is based on the MNAR assumption, not based on the MAR assumption. Any score imputed outside the range of the WOMAC subscale score of 0-10 will be truncated to the nearest permissible value on the WOMAC scale.

The following SAS code will be used to make the imputation with the MNAR assumption:

[redacted]

[redacted]

Step 3: Each imputed data set will be analyzed using the MMRM model with treatment, randomization strata, and relevant baseline included in the model. For each δ the SAS MIANALYZE procedure will be used to generate valid statistical inferences by combining results from the 50 analyses using Rubin's formula. By progressively increasing δ, the sensitivity analysis will explore the tipping points, e.g., δ value when the p-value for a treatment comparison is above 0.05. Results will be then summarized using summary tables and graphs.

[redacted]

Multiplicity adjustment

Because this study is a confirmatory study, in order to keep the family wise type 1 error rate to a significance level of 5% (alpha), two-sided, in the tests of the following hypotheses regarding the primary endpoint and the key secondary efficacy endpoint, a gatekeeping method based on a graphical approach will be used to perform multiplicity adjustment, as shown in Figure 8.2.1-1 to verify the hypotheses (Bretz et al., 2009). This multiplicity adjustment will only be performed for the statistical analyses of data from the double-blind period from screening to Week 16 after treatment initiation.

- Null hypothesis H01: The change from baseline in the WOMAC pain subscale score at Week 16 in the 1 mg SC q4w group is the same as that in the placebo group
- Null hypothesis H02: The change from baseline in the WOMAC pain subscale score at

Week 16 in the 1 mg SC q8w group is the same as that in the placebo group

- Null hypothesis H03: The change from baseline in the WOMAC physical function subscale score at Week 16 in the 1 mg SC q4w group is the same as that in the placebo group
- Null hypothesis H04: The change from baseline in the WOMAC physical function subscale score at Week 16 in the 1 mg SC q8w group is the same as that in the placebo group

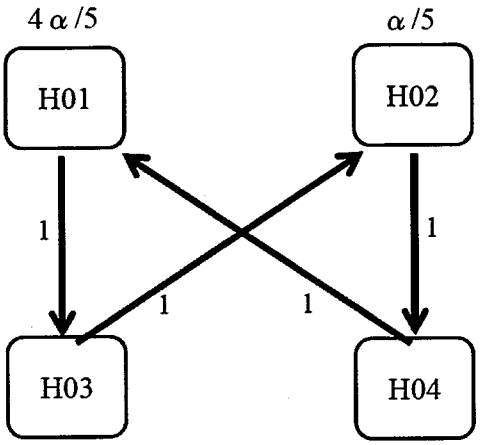


Figure 8.2.1-1: MT-5547 Multiple Test Procedure

Subgroup analysis

Subgroup analyses will be performed on the primary endpoint. Each subgroup is following at baseline timepoint:

- index joint (hip or knee)
- K-L score category(2-3 or 4)
- WOMAC pain (<6 or >=6)
- OA Duration(<5 or >=5)
- age (<65 or >=65)
- sex(Male, Female)
- BMI (<25 or >=25)
- Weight(<median or >=median)
- Opioid (Yes or No)

Each subgroup will be analyzed separately for the primary endpoint using MMRM model similar to the primary model.

Change from baseline to week 16 for the subgroup of index joint (hip or knee) and K-L score category(2-3 or 4) will be summarized by visit. Change from baseline to week 16 for the other subgroup will be summarized at week 16. In addition, forest plot depicting the treatment effect, p-values for interaction terms between the treatment group and each subgroup and 95% CI for each subgroup will be displayed.

Supplemental analysis

The PPS will be analyzed using the same method as for the primary analysis to confirm the robustness of the results of the primary analysis for the effects of the analysis set.

Supplementary analysis with respect to the data pooling on 1 mg q4w and placebo

The data on 1 mg q4w and placebo obtained before and after the protocol amendment (Ver. 04.00.00000) will be analyzed using the same method as for the primary analysis to confirm the data homogeneity and the robustness of the results of the primary analysis. Also, homogeneity in demographic and other baseline characteristics [Section 8.1.3] between data obtained before and after the protocol amendment will be confirmed.

Supplementary analysis with respect to response the site differences

Absolute values and changes from baseline will be summarized at week 16 by treatment group and site.

8.2.2. Key Secondary Efficacy Endpoints

The change in the WOMAC physical function subscale score from baseline to Week 16 will be analyzed on the mFAS and PPS population as the key secondary endpoint using the same primary analysis method (MMRM) and same subgroup analysis and supplementary analysis with respect the data pooling on 1mg q4w and placebo will be performed. Multiplicity adjustment described in Section 8.2.1.

8.2.3. Other Efficacy Endpoints

WOMAC pain subscale score, WOMAC physical function subscale score, WOMAC stiffness subscale score, WOMAC total score , WOMAC individual score and PGA

For the changes from baseline at each visit in the WOMAC pain subscale score, the WOMAC physical function subscale score, the WOMAC stiffness subscale score, the WOMAC total score, WOMAC individual score and PGA, the same primary analysis method (MMRM) described in Section 8.2.1 will be analyzed. And the WOMAC pain subscale score will also be summarized at each visit by treatment group on the FAS population. In addition,

change from baseline to week 16 in WOMAC pain subscale score, WOMAC physical function subscale score and PGA will be graphed by visit.

Weekly average and daily walking index joint pain using NRS

For the change from baseline in the weekly average and daily walking index joint pain using NRS, the same primary analysis method (MMRM) described in Section 8.2.1 will be analyzed. In addition, change from baseline to week 16 in weekly average walking index joint pain using NRS will be graphed by visit.

SF-36 / EQ-5D-5L

For the changes from baseline at each visit in SF-36 eight subscale summary scores and three component summary socres / EQ-5D-5L visual analogue scale(VAS) and utility index scores, the same primary analysis method (MMRM) described in Section 8.2.1will be analyzed.

Proportions of subjects with improvement of 30%, 50%, 70% and 90% at each visit compared to baseline in WOMAC pain subscale score, WOMAC physical function subscale score and PGA

A generalized linear mixed-effects model (GLMM) by logit type for response variable will be used to evaluate these binary response variables. The denominator degrees of freedom for the test statistics of estimated parameters will be approximated using the Kenward-Roger method. The estimated response rate for each treatment group and placebo group, P value, and 95% confidence interval of the odds ratio of each of the MT-5547 1 mg q4w and MT-5547 1 mg q8w groups compared to the placebo group in each binary response variables will be calculated at each visit using this GLMM approach.

- The covariates: baseline and the interaction of baseline and measurement time point.
- The fixed factors: Treatment group, measurement time point, the interaction of treatment group and measurement time point, K-L grade (2-3 or 4), and evaluated joint location (knee or hip)
- Random factors: Intercept for subject

The SAS code planned for the analysis is outlined below.

[redacted]

[redacted]

Response rate based on the OMERACT-OARSI criteria at Week 16 compared to baseline

Responder analyses for OMERACT-OARSI will be summarized at each visit using the same GLMM approach in previous model.

Proportion subjects who used rescue medication

The proportion of subjects taking rescue medication in baseline and up to week 16 will be summarized by treatment group. Number of days on rescue medication during treatment priod(Day1 to Day 28, Day 29 to Day 56, Day 57 to Day 84, Day 85 to Day 112) will be summarized by treatment group.Number of weekly average usage of rescue medication will be summarized at each week by treatment group.

8.3. Safety Assessments

Safety assessments will be made on the SAF population.

8.3.1. Adverse Events

Overall summary for the following will be conducted.

- Subjects with at least one TEAE
- Subjects with at least one adverse drug reaction
- Subjects with at least one TESAE
- Subjects with at least one serious adverse drug reaction
- Subjects with at least one AESI
- Subjects with at least one TEAE leading to discontinuation of study drug
- Subjects with TEAE leading to death

The following summaries also will be conducted.

- TEAEs by SOC and PT
- Adverse drug reactions by SOC and PT
- AESI by SOC and PT
- TESAEs by SOC and PT
- Serious adverse drug reactions by SOC and PT
- TEAEs leading to discontinuation of study drug by SOC and PT
- TEAEs by SOC, PT and severity
- Adverse drug reactions by SOC, PT and severity

Each of the summaries will be done by treatment group at the subject level - multiple occurrences of the same event within a subject will be counted once in the summaries by SOC and PT; multiple occurrences of the same event within a subject will be counted once in the

maximum severity category (severe > moderate > mild) and/or maximum drug relationship category (reasonable possibility/no reasonable possibility) and/or the earliest duration.

9. DATA PRESENTATION CONVENTIONS

9.1. Number of Digits to Report

(1) Non-PK related

Statistic	Specification	Apply to
Minimum, Maximum	Same number of DPs as the data provided in the datasets	All original (i.e. non-derived)
	see section 7.3	All derived data
Mean, Median, SD, SE, Confidence intervals	One more DP than above	All
Percentages[*1]	1 DP	All
Ratios	3 DPs	All
p-values[*2]	3 DPs	All

[*1] Percentages: use 1 place beyond the decimal point, except for the following cases:
If the percentage is equal to 0, then leave blank, do not use (0)
If the percentage is equal to 100, then use "(100)" without a decimal
[*2] p-values: use 3 places beyond the decimal point, except for the following cases:
If the p-value is less than 0.001, then use p<0.001

9.2. Treatments to Report

Treatment	For TFLs
MT-5547 1 mg q8w	MT-5547 1 mg q8w
MT-5547 1 mg q4w	MT-5547 1 mg q4w
MT-5547 1 mg q4w and MT-5547 1 mg q8w	MT-5547 Combined
MT-5547 3 mg q4w	MT-5547 3 mg q4w
MT-5547 6 mg q8w	MT-5547 6 mg q8w
Placebo	Placebo
Placebo group and all active dose levels	Total

9.3. Analysis Visits to Report

(1) Non-PK related

Efficacy:

Analysis Visit	Apply to			
	WOMAC and PGA	SF-36 and EQ-5D-5L	Weekly average NRS and Rescue	Daily NRS

			Medicine	
Baseline	X	X	X	X
Day 1				X
Day 2				X
Day 4				X
Day 6				X
Week 1	X		X	X
Week 2	X		X	X
Week 3			X	X
Week 4	X	X	X	X
Week 5			X	X
Week 6			X	X
Week 7			X	X
Week 8	X	X	X	X
Week 9			X	X
Week 10			X	X
Week 11			X	X
Week 12	X		X	X
Week 13			X	X
Week 14			X	X
Week 15			X	X
Week 16	X	X	X	X

10. CHANGE FROM THE PROTOCOL

There are currently no changes to analysis from protocol.

11. SOFTWARE

All statistical analyses will be performed using SAS version 9.3 or higher.

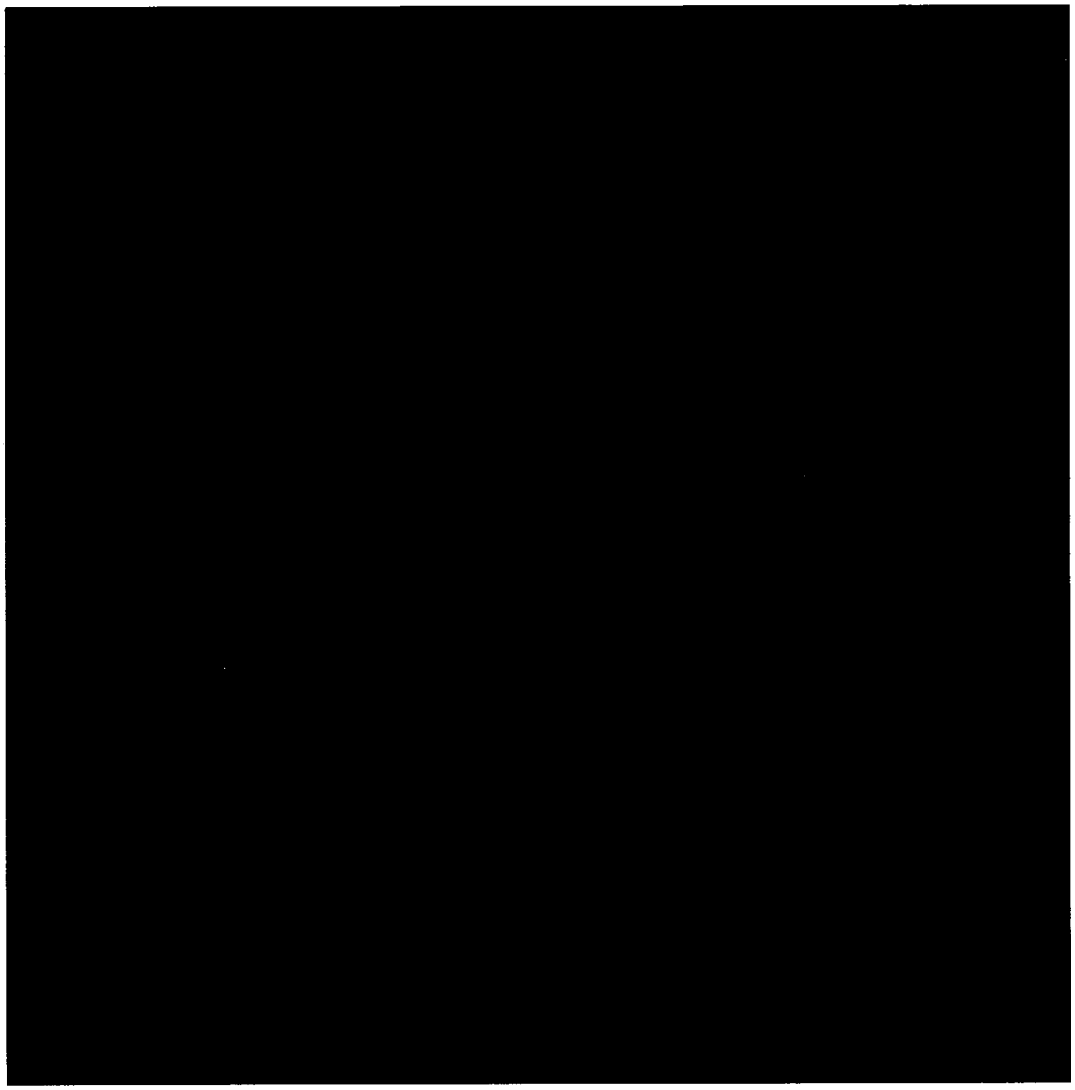
12. REFERENCES

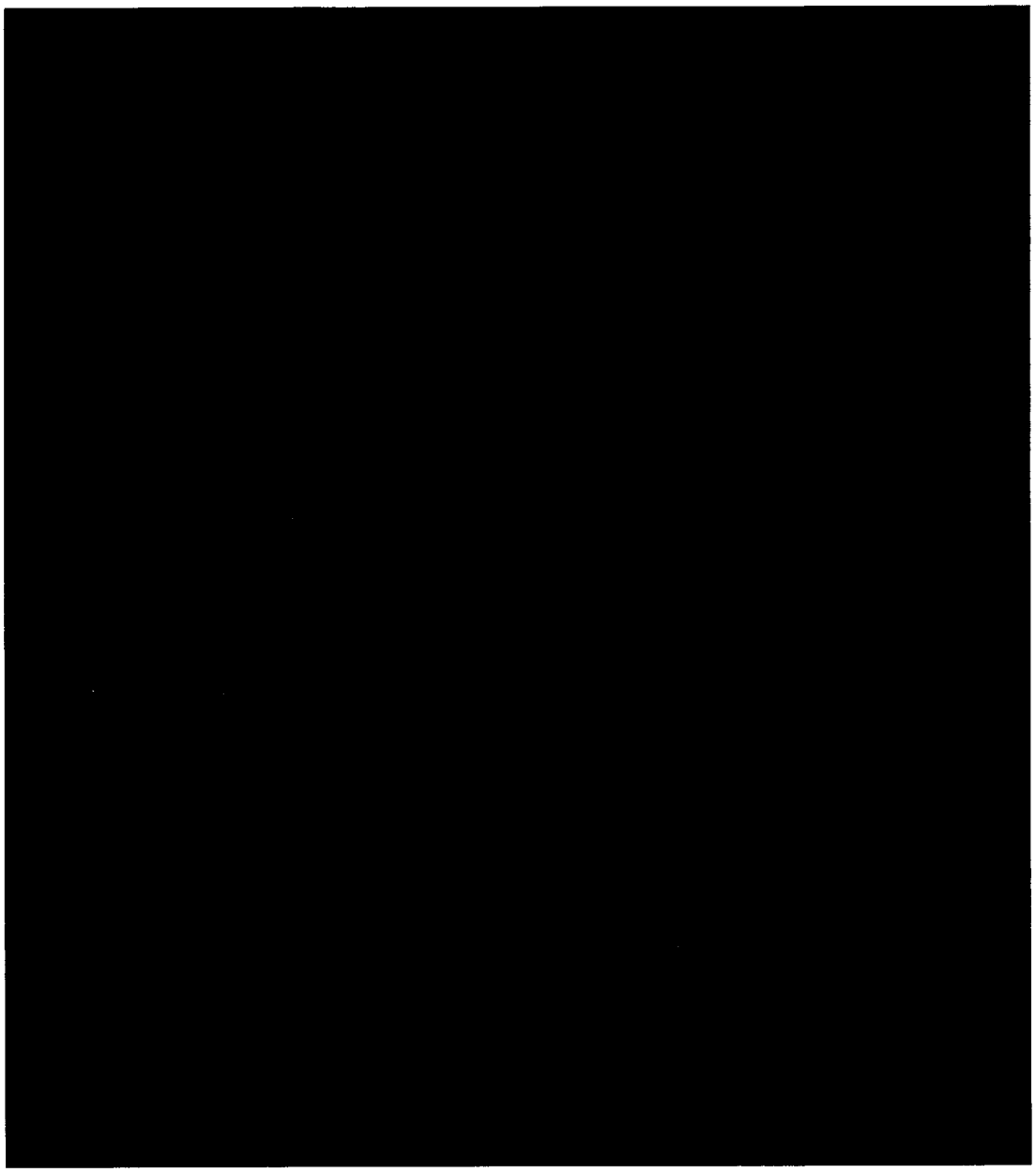
[1] Shinya I, Takeru S, Ataru I, Shinichi N, Takashi F, et al. Developing a Japanese version of the EQ-5D-5L value set. J. Natl. Inst. Public Health.2015;64 (1) : 47-55.

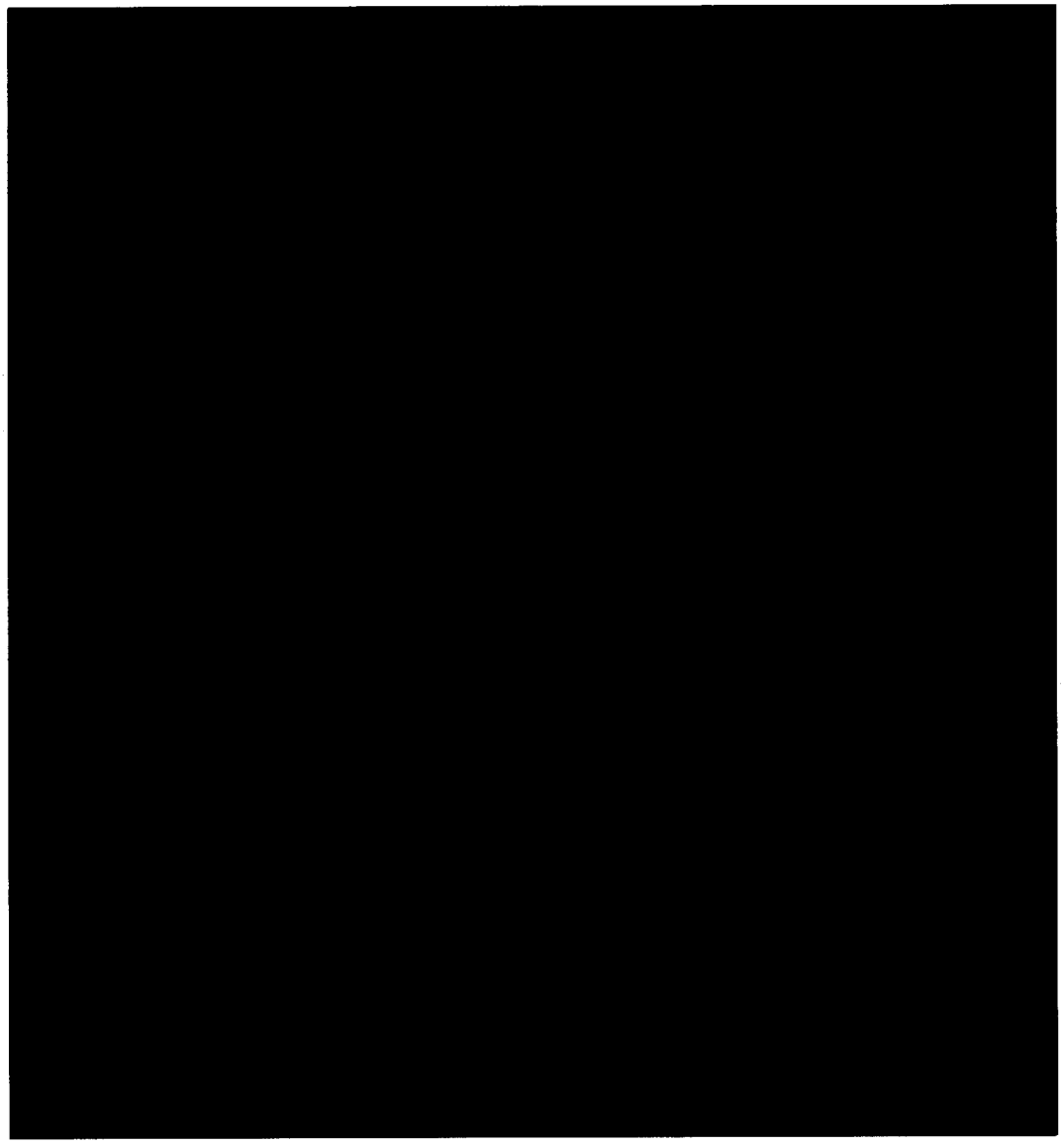
[2] Fukuhara S, Suzukamo Y. Manual of SF-36v2 Japanese version: Institute for Health Outcomes & Process Evaluation research. Kyoto: 2004. Ver3/2011.

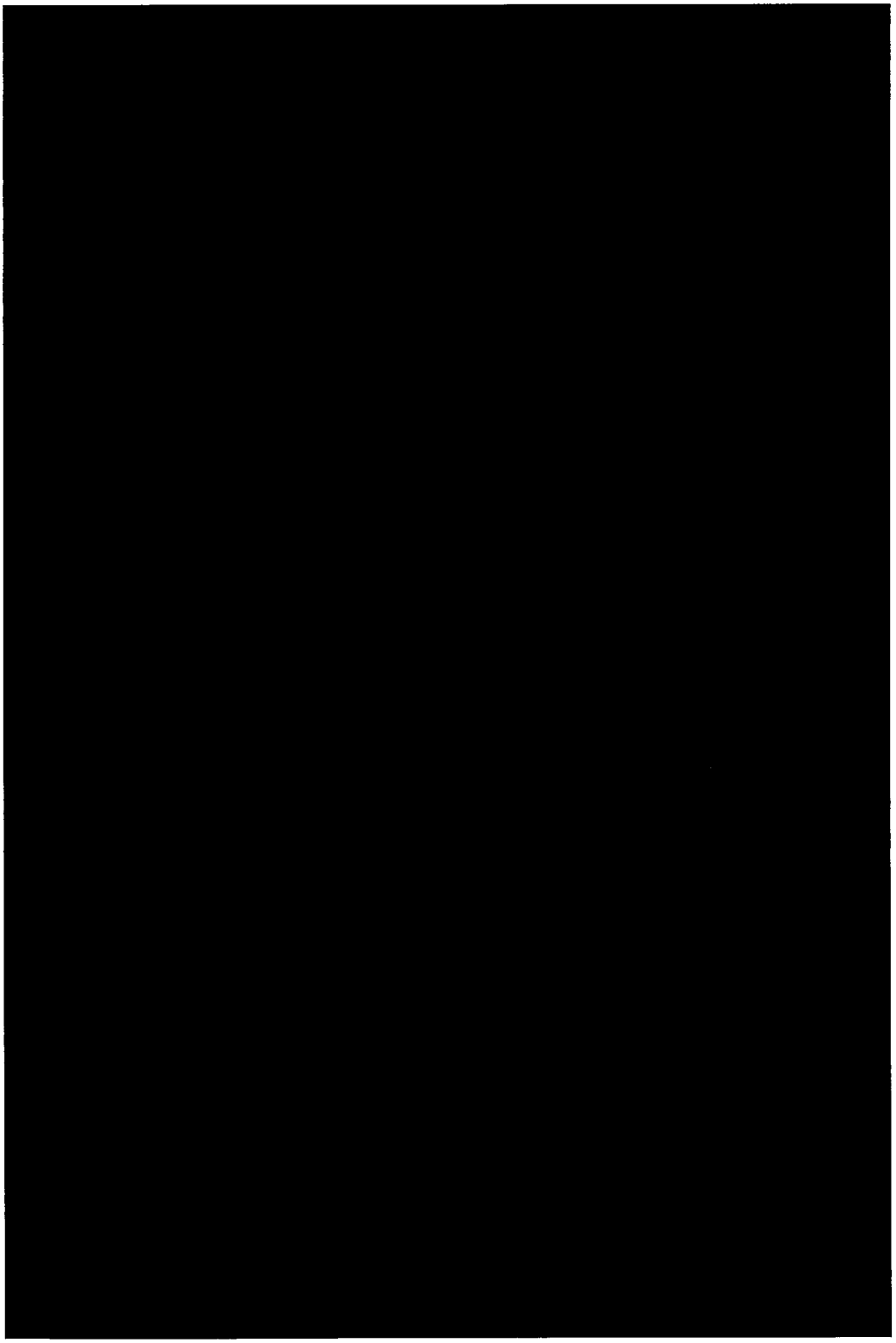
13. APPENDIX

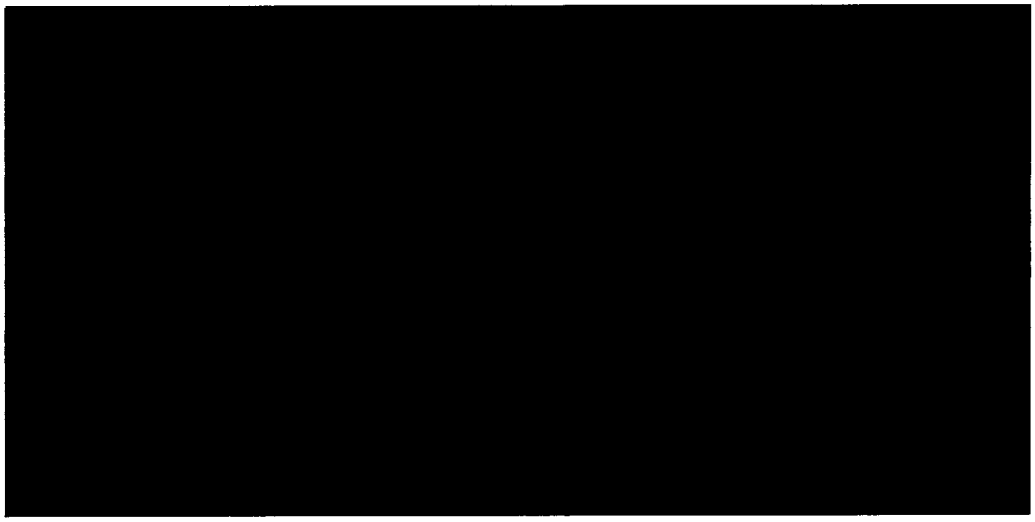
13.1. SF-36v2 Scoring Algorithm[2]











Statistical Analysis Plan (For Database at Week 68)

Protocol No. MT-5547-J01

A Phase 2/3 (Placebo-Controlled, Double-Blind, Comparative) Study on MT-5547 in Patients with Osteoarthritis Accompanied by Moderate to Severe Pain

Prepared By:	Mitsubishi Tanabe Pharma Corporation
Version:	1.0
Date:	19-April-2021

APPROVAL FORM

Statistical Analysis Plan(For Database at Week 68)

Protocol No.	MT-5547-J01
Protocol Title	A Phase 2/3 (Placebo-Controlled, Double-Blind, Comparative) Study on MT-5547 in Patients with Osteoarthritis Accompanied by Moderate to Severe Pain
Version / Date	1.0 / 19-April-2021

Authors:

Statistics Author	
Print Name:	[redacted]
Position:	Biostatistician
Clinical Pharmacokinetics Author	
Print Name:	[redacted]
Position:	Pharmacokinetics

Approved by:

Statistic Approver	
Print Name:	[redacted]
Position:	Responsible Biostatistician
Signature:	Refer to email
Approval date:	Refer to email
Clinical Pharmacokinetic Approver	
Print Name:	[redacted]
Position:	Responsible Pharmacokinetics
Signature:	Refer to email
Approval date:	Refer to email

TABLE OF CONTENTS

ABBREVIATIONS

Abbreviations	Definitions
ADA	anti-drug antibodies
AE	adverse event
AESI	Adverse event of special interest
ALP	alkaline phosphatase
ALT	alanine transaminase
ANCOVA	analysis of covariance
AST	aspartate transaminase
ATC	anatomical therapeutic chemical
BAP	bone specific alkaline phosphatase
BDR	blinded data review
BLQ	below limit of quantification
BMI	body mass index
CI	confidence interval
CRP	C-reactive protein
CV	coefficient of variation
DA	destructive arthropathy
DBP	diastolic blood pressure
DMC	data monitoring committee
DP	decimal places
ECG	12-lead electrocardiogram
EQ-5D-5L	EuroQol-5-Domain-5-Level health questionnaire
FAS	full analysis set
K-L	Kellgren-Lawrence
LLOQ	lower limit of quantification
MAR	missing at random
MedDRA	medical dictionary for regulatory activities
MMRM	mixed effect model for repeated measures
NAb	neutralizing anti-drug antibody
NGF	nerve growth factor
NRS	numerical rating scale
NSAID	nonsteroidal anti-inflammatory drug
OA	osteoarthritis
[redacted]	[redacted]
OMERACT-OARSI	Outcome Measures in Rheumatology Arthritis Clinical Trials-Osteoarthritis Reseach Society International
PD	pharmacodynamics
PGA	patient global assessment
[redacted]	[redacted]
PK	pharmacokinetics
PR	pulse rate
PT	MedDRA preferred term
Q4W	every 4 weeks
Q8W	every 8 weeks

QOL	quality of life
RAND	all subjects randomized population
SAE	serious adverse event
SAF	safety analysis set
SAP	statistical analysis plan
SBP	systolic blood pressure
SC	subcutaneous
SD	standard deviation
SF-36	Short Form 36-item health survey
SIF	Subchondral insufficiency fracture
SOC	system organ class
TEAE	treatment emergent adverse event
TESAE	treatment emergent serious adverse events
[redacted]	[redacted]
ULN	upper limit of normal range
WHO	World Health Organization
WOMAC	Western Ontario and McMaster Osteoarthritis Index

1. INTRODUCTION

This statistical analysis plan (SAP) is based on the final protocol (v4.0) dated 31-July-2018. This protocol will plan to respectively perform the statistical analysis for the clinical trial data from screening to Week 16, from screening to Week 24 and from screening to Week 68.

For the clinical trial data from screening to Week 68, this SAP covers statistical analysis, tabulations and listings of the study data to investigate the efficacy and safety of MT-5547 SC 1mg q4w and 1mg q8w compared to matching placebo. Therefore, this SAP will be finalized prior to database lock at Week 68 to ensure the credibility of the study results by pre-specifying the statistical methods for the data analyses before un-blinding of treatment assignments. For the clinical trial data from screening to Week 16 and from screening to Week 24 was separately created.

Any statistical analysis details described in this document supersede any description of statistical analysis in the protocol.

2. STUDY OBJECTIVE AND ENDPOINTS

2.1. Study Objective(s)

The objective of this study is to verify the superiority of 16 weeks of MT-5547 treatment to placebo, as evidenced by the WOMAC pain subscale score (the efficacy outcome measure), in patients with osteoarthritis of the knee or hip. Additional objectives of the study are to investigate the efficacy, safety, and pharmacokinetics of MT-5547 in long-term use.

2.2. Study Endopoint(s)

2.2.1. Efficacy Endpoint(s)

(1) Primary efficacy endpoint
WOMAC pain subscale score (change from baseline at Week 16)

(2) Key secondary efficacy endpoint
WOMAC physical function subscale score (change from baseline at Week 16)

(3) Secondary efficacy endpoints
1) Patient global assessment (PGA) (change from baseline at each assessment time point)
2) WOMAC pain subscale score (change from baseline at each assessment time point)
3) WOMAC physical function subscale score (change from baseline at each assessment time point)
4) WOMAC stiffness subscale score (change from baseline at each assessment time point)
5) WOMAC total score (change from baseline at each assessment time point)
6) Proportions of subjects with 30%, 50%, 70% and 90% improvements in their WOMAC pain subscale scores compared to baseline (at each assessment time point)

7) Proportions of subjects with 30%, 50%, 70% and 90% improvements in their WOMAC physical function subscale scores compared to baseline (at each assessment time point)
8) Proportions of subjects with 30%, 50% and 70% improvements in their PGA compared to baseline (at each assessment time point)
9) Numerical Rating Scale (NRS) score for the mean pain on walking in the evaluated joint (change from baseline at each assessment time point)
10) SF-36 (change from baseline at each assessment time point)
11) EQ-5D-5L (change from baseline at each assessment time point)
12) Improvement rate based on the Outcome Measures in Rheumatology Arthritis Clinical Trials-Osteoarthritis Research Society International (OMERACT-OARSI) criteria (at each assessment time point)
13) Amount of rescue medication used (number of days on which rescue medication was used, and amount [number of tablets] used)

2.2.2. Safety Assessment(s)

(1) Adverse events and adverse reactions (clinically significant adverse events were identified as being adjudicated arthropathy, sympathetic nervous system disorders and altered peripheral sensation)
(2) General clinical laboratory tests
(3) Body weight
(4) Vital signs
(5) ECG
(6) Physical examination
(7) Joint pain questionnaire
(8) Survey of autonomic symptoms
(9) Assessment of orthostatic blood pressure
(10) Neurological assessments
(11) Bone density
(12) Imaging Tests
(13) Injection site assessments
(14) Assessments of subjects undergoing joint replacement

2.2.3. Pharmacokinetics Assessment(s)

MT-5547 concentration in serum

Anti-MT-5547 antibody in serum

2.2.4. Pharmacodynamics Endpoint(s)

Efficacy: WOMAC pain subscale score, WOMAC physical function subscale score, PGA

Safety: alkaline phosphatase (ALP), C-reactive protein (CRP), bone specific alkaline phosphatase (BAP)

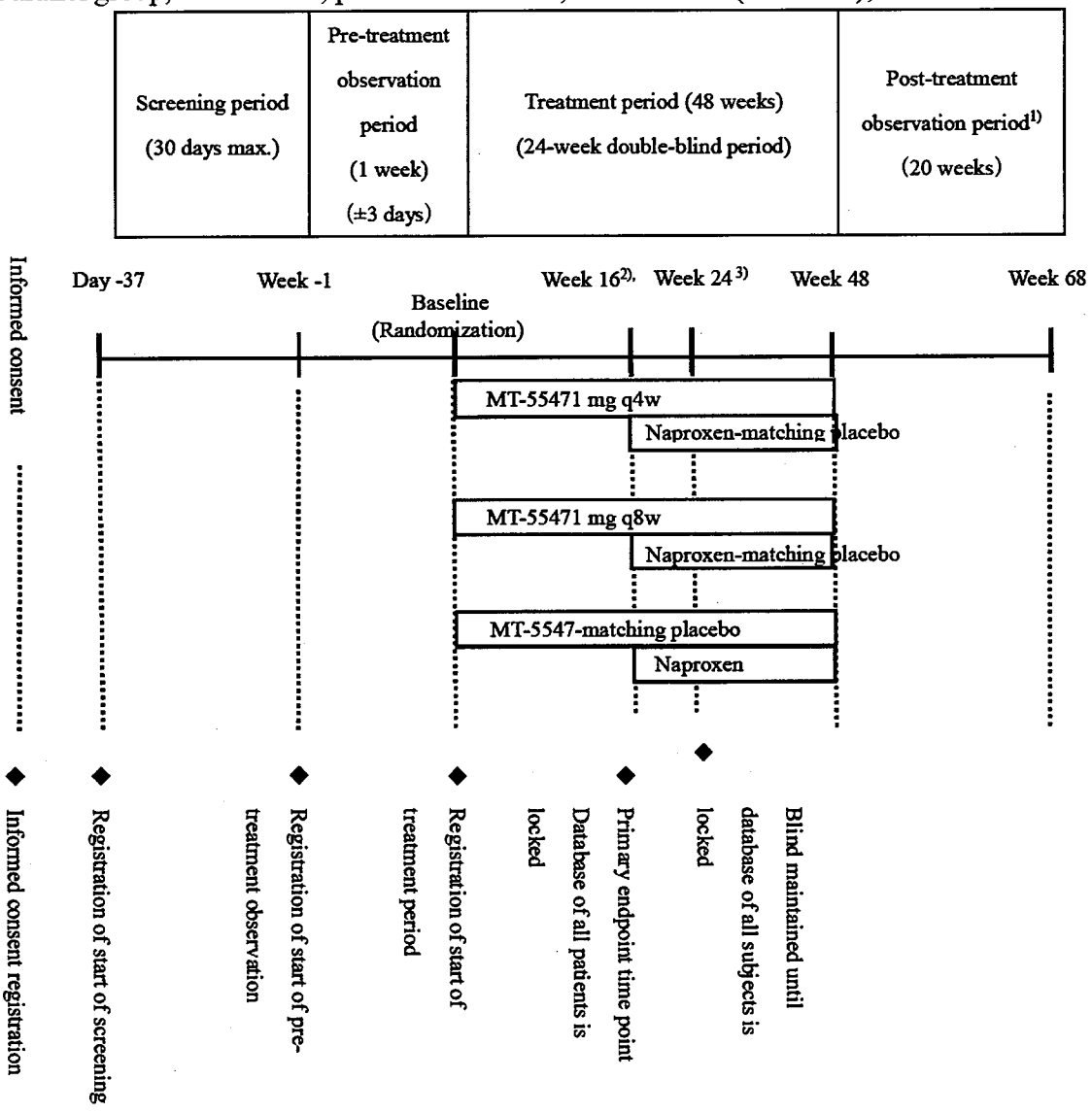
3. STUDY DESIGN

3.1. Study Design

Study Phase: Phase 2/3

Type of Study: Confirmatory Study

Parallel group, randomized, placebo-controlled, double-blind (24 weeks), multicenter



1) The post-treatment observation period will be the 20 weeks from the day after the end of the 48-week treatment period. Furthermore, for subjects discontinued from the study, the post-treatment observation period will be the 24 weeks from the administration of MT-5547 or MT-5547 placebo. If a joint replacement procedure is going to be performed, the postoperative test observation period will be the period from the day after the day of discontinuation until 20 weeks after the joint replacement procedure is performed.
2) From Week 16 to Week 48, the placebo group will receive an additional dose of naproxen tablets to reduce the dropout rate in the placebo group. The MT-5547 group will receive an additional dose of naproxen placebo tablets.
   For each subject, the data for up through Week 16 will be fixed prior to the Week 20 visit, and the fixed data will be prepared.
3) The blind will be maintained, including for the study sponsor, until the key code is opened after the data have been fixed for all subjects in Week 24. In addition, in order to eliminate any bias that could affect the study assessments, to the extent possible, the blind will be maintained with respect to the study site and the subject even after the opening of the study sponsor key code.

3.2. Schedule of Study Procedures

3.2.1. Test/Observation Schedule

(1) Test/Observation Schedule from informed consent to Week 48 of treatment period (if a joint replacement procedure is not going to be performed)

				Treatment period									Treatment period						
Visit times (visit window)	Informed consent	Screening Day -37 ~ -8	Pre-treatment of observation period Day -7 ~ -1 (±3)	Day1 Baseline	Week1 Day 8 (±3)	Week2 Day 15 (±3)	Week4 Day 29 (±7)	Week8 Day 57 (±7)	Week12 Day85 (±7)	Week16 Day 113 (±7)	Week 20 Day 141 (±7)	Week 24 Day 169 (±7)	Week 28 Day 197 (±7)	Week 32 Day 225 (±7)	Week 36 Day 253 (±7)	Week 40 Day 281 (±7)	Week 44 Day 309 (±7)	Week 48 Day 337 (±7)	At withdrawal[16]
Visit Number		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
Written informed consent	X																		
Eligibility confirmed		X	X	X															
Demographics		X		X															
Randomization[1]				X															
Study drug SC administration[2]				X			X	X	X	X	X	X	X	X	X	X	X		
Administration of naproxen										←								→	
Delivery of naproxen										X	X	X	X	X	X	X	X		
Return of naproxen											X	X	X	X	X	X	X	X	X
Delivery of acetaminophen			X	X		X	X	X	X	X	X	X	X	X	X	X	X		
Return of acetaminophen/check of amount taken[3]				X		X	X	X	X	X	X	X	X	X	X	X	X	X	X
Concomitant medications		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
[Efficacy]																			
NRS (average daily pain)[4]			X	←						→									
WOMAC[5]		X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
PGA of OA		X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
SF-36				X			X	X		X		X		X		X		X	X
EQ-5D-5L				X			X	X		X		X		X		X		X	X
[Safety]																			
Body weight (height: only at screening)		X								X								X	X
Vital signs[6]		X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
ECG		X																X	X
Physical examination		X										X		X				X	X
Orthostatic blood pressure		X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X
Injection site assessment (30 minutes after administration)				X			X	X	X	X	X	X	X	X	X	X	X		
Joint Pain Questionnaire		X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Autonomic Nerve Questionnaire		X		X			X	X	X	X	X	X	X	X	X	X	X	X	X
Neurological evaluation		X		X (brief version)		X (brief version)	X (brief version)	X (brief version)	X (brief version)	X (brief version)	X (brief version)	X (brief version)	X (brief version)	X (brief version)	X (brief version)	X (brief version)	X (brief version)	X	X
Radiology (bilateral knees, hips, shoulders)[7]		X								X								X	X
MRI testing (evaluated joint, contralateral joint, knee or hip joints with K-L ≥ 3)[7]		X																	
MRI testing (joint of joint replacement surgery)																			X
Adverse events		←																→	X
At joint pain worsening (imaging [X-ray, MRI] assessments)[8]					←													→	X
JOA Score[9]																			X
Laboratory testing[10]		X		X			X	X	X	X		X		X		X		X	X
Pregnancy test (for WOCBP)[11]		X		X			X	X	X	X	X	X	X	X	X	X	X	X	X
Bone density[12]		X																X	X
Urinalysis		X		X						X		X						X	X
[Pharmacokinetics]																			
PK measurements				X[15]	X	X	X[15]	X[15]	X[15]	X[15]		X[15]		X[15]		X[15]	X[15]	X	X
ADA antibody				X[15]						X[15]		X[15]						X	X
[Other]																			
[redacted]				[redacted]			[redacted]			[redacted]		[redacted]		[redacted]				[redacted]	[redacted]
[redacted]				[redacted]															

(2) Test/Observation Schedule of post-treatment observation period (if a joint replacement procedure is not going to be performed)

	Post-treatment observation period (20 weeks)[14]					
	Week 52	Week 56	Week 60	Week 64	Week 68	At withdrawal
Visit times (visit window)	Day 365 (±7)	Day 393 (±7)	Day 421 (±7)	Day 449 (±7)	Day 477 (±7)	
Visit Number	18	19	20	21	22	
Written informed consent						
Eligibility confirmed						
Demographics						
Randomization[1]						
Study drug SC administration[2]						
Administration of naproxen						
Delivery of naproxen						
Return of naproxen						
Delivery of acetaminophen						
Return of acetaminophen/check of amount taken[3]						
Concomitant medications	X	X	X	X	X	X
[Efficacy]						
NRS (average daily pain)[4]						
WOMAC[5]	X		X		X	X
PGA of OA	X		X		X	X
SF-36					X	X
EQ-5D-5L					X	X
[Safety]						
Body weight (height: only at screening)					X	X
Vital signs[6]	X				X	X
ECG						
Physical examination					X	X
Orthostatic blood pressure	X				X	X
Injection site assessment (30 minutes after administration)						
Joint Pain Questionnaire	X				X	X
Autonomic Nerve Questionnaire	X				X	X
Neurological evaluation	X (brief version)				X	X
Radiology (bilateral knees, hips, shoulders)[7]					X	X
MRI testing (evaluated joint, contralateral joint, knee or hip joints with K-L ≥ 3)[7]						
MRI testing (joint of joint replacement surgery)						X
Adverse events	←				→	X
At joint pain worsening (imaging [X-ray, MRI] assessments)[8]	←				→	X
JOA Score[9]						X
Laboratory testing[10]					X	X
Pregnancy test (for WOCBP)[11]	X				X	X
Bone density[12]						
Urinalysis					X	X
[Pharmacokinetics]						
PK measurements	X	X			X	X
ADA antibody					X	X
[Other]						
[redacted]					[redacted]	[redacted]
[redacted]						

1) Patients will be randomized (enrolled at the start of the treatment period) after all of the assessments through the pre-treatment observation period have been completed.
2) Investigational drug or comparative drug will be administered subcutaneously after all of the safety assessments, efficacy assessments, and collection of blood samples for pharmacokinetic assessments have been completed, and patients will be kept under supervision at the study site for 30 minutes after the administration of the study drug or comparative drug.
3) The dose of acetaminophen taken will be checked based on the treatment diary. Every night, patients will record in their treatment diaries the amount (number of tablets) of acetaminophen they took that day.
4) Every night (and as a rule at the same time) from the start of the pre-treatment observation period until the day before Week 16, patients will use the IVRS to report the mean pain on walking for the previous 24 hours (NRS data). Patients will receive instruction in how to use the IVRS on the day of the visit at the start of the pre-treatment observation period.
5) The WOMAC score at screening will be assessed for the right knee joint, the left knee joint, the right hip joint, and the left hip joint, and the index joint will be determined based on the WOMAC pain subscale scores. At the other time points, the WOMAC score will be assessed for the index joint only.
6) If the pulse is less than 45 bpm at an assessment time point after investigational drug or comparative drug administration, electrocardiography will be performed and the patient carefully examined for cardiac function abnormalities.
7) Screening:
The following imaging tests will be performed for patients who meet the study eligibility criteria, and the images will be submitted to a central adjudication office. Imaging tests will also be performed if the central adjudication office requests that additional tests be conducted.
· X- ray tests: Both knee joints, both hip joints, both shoulder joints
· MRI tests: The index joint, the joint opposite the index joint, and any knee or hip joints with a K-L grade of 3 or more
MRI tests will be performed as needed by the (sub) investigator.
A check will be performed by the day specified below to make sure that the results of the central assessments meet the study enrollment criteria.
· X-ray tests: By the start of the pre-treatment observation period
· MRI tests: By baseline
If the results of the central assessments cannot be confirmed by the specified day, then re-screening will be performed.
After baseline:
At week 16, 48 (or at discontinuation of the treatment period) and at 68 weeks (or at the early discontinuation of the post-treatment observation period), images will be submitted to the central adjudication office. If the (sub) investigator determine that there is an abnormal finding compared to normal osteoarthritis, MRI tests will be performed as needed. The images will be submitted to a central adjudication office. Imaging tests will also be performed if the central adjudication office requests that additional tests be conducted, and the images submitted to the central adjudication office (for more detailed information, see in the text "9.2.4.1.11 Imaging Tests").
Week 16:
The X-ray tests (both knee joints, both hip joints, both shoulder joints) will be performed by the (sub) investigator. The images will be submitted to a central adjudication office. The (sub) investigator will check the X-ray images before administered the investigational drug or comparative drug to determine whether or not investigational drug or comparative drug should be administered. If any abnormalities are found, investigational drug or comparative drug administration will be postponed. If the central adjudication office rules out adjudicated arthropathy, investigational drug or comparative drug administration will be resumed.
At discontinuation of the treatment period and discontinuation of the post-treatment observation period
The X-ray tests (both knee joints, both hip joints, both shoulder joints) will be performed by the (sub) investigator. The images will be submitted to a central adjudication office. If more than 30 days have passed since the last images were taken, images will be taken again. If not more than 30 days have passed, imaging tests will be performed if the (sub) investigator determines that it is necessary.
If joint replacement surgery is required, the MRI tests of the joint of surgery will be performed (for more detailed information, see in the text "9.2.4.1.13 Post-joint Replacement Procedure Assessment").
8) If a patient experiences a sudden worsening of pain that would not occur in the normal course of osteoarthritis, the (sub) investigator will postpone investigational drug or comparative drug administration, and perform imaging test (X-ray and MRI) of the affected joint. Imaging test (X-ray and MRI) of other joints will be performed as needed by the

(sub) investigator. Imaging tests will also be performed if the central adjudication office requests that additional tests be conducted. All images will be submitted to the central adjudication office. The results of the central assessment will be checked by the (sub) investigator. The investigational drug or comparative drug administration will be postponed until the central adjudication office rules out adjudicated arthropathy. If adjudicated arthropathy is ruled out by the central assessment, then study treatment will be resumed. If the result of the central assessment is adjudicated arthropathy, then study treatment must be discontinued.

9) Assessments of the joint by JOA score at discontinuation will be performed for patients who required joint replacement surgery during the study.
10) Hematology and blood biochemistry tests will be performed. HbA1c, HBs antigen, HBs antibody, HCV antibody, and HIV antibody will be measured only at screening.
11) At screening, the pregnancy test will be performed using serum. At all other time points, the pregnancy test will be performed using urine. Furthermore, pregnancy tests will not be performed for postmenopausal females who have been amenorrheic for at least 1 year or for females who have undergone a surgical hysterectomy or bilateral ovariectomy. If a positive result is obtained on a urine pregnancy test, a serum pregnancy test will be performed. If a negative result is obtained on a serum pregnancy test, the study can be continued for the patient. If the serum pregnancy test is also positive, study treatment will be discontinued.
12) Bone density will be measured only at those study sites that are capable of performing measurements using the DEXA method. If it can not be measured at screening period, it may be measured during the pre-treatment observation period.
13) [redacted]
14) The Week 64 assessments may be performed by telephone.
15) The blood samples will be collected before the investigational drug or comparative drug is administered.
16) If a subject discontinues during the treatment period, a post-treatment observations will be performed at 4, 8, 12, 16, 20 and 24 weeks after the administration of the last dose of the investigational drug or control drug for the same parameters as those evaluated at the corresponding visits for completers. However, blood sample collection for pharmacokinetic evaluation is performed only 24 weeks after the last dose (blood sample collection for PK measurement is not necessary 4, 8 and 12 weeks after the last dose, which correspond to Weeks 48, 52 and 56 of the treatment period, respectively, for completers). The allowable time window for each evaluation time point is plus or minus 7 days.
The assessments at 12, and 20 weeks after the last dose administered to subjects discontinuing in the treatment period may be performed by telephone.
If the day of discontinuation from the treatment period and the day of discontinuation from the post-treatment observation period are the same (if the patient discontinued from the treatment period without proceeding to the post-treatment observation period), the assessments that are to be performed at discontinuation from the treatment period will be performed.

Discontinuation from the treatment period	4 weeks after the last dose	8 weeks after the last dose	12 weeks after the last dose	16 weeks after the last dose	20 weeks after the last dose	24 weeks after the last dose	Discontinuation from the post-treatment observation period
Completed of the treatment period	Week 48	Week 52	Week 56	Week 60	Week 64	Week 68	Discontinuation from the post-treatment observation period

If the assessments at discontinuation are included in the allowable time window for each assessment time point, the parameters that are to be assessed at discontinuation will be assessed. Subsequently, at each time point after the last dose, the corresponding post-treatment assessments will be performed.
If not more than 30 days have passed since the last images were taken, imaging will not have to be performed again, but may be performed if the (sub) investigator determines that it is necessary. The tests at discontinuation and post-joint replacement procedure test/observation will be performed for patients who required joint replacement surgery during the study. Imaging test (X-ray and MRI) will be submitted to a central adjudication office (for more detailed information, see "9.1.1 (3) Post-joint replacement procedure test/observation schedule). If the assessments at discontinuation are not performed before the joint replacement procedure is performed, then images from before the procedure will be obtained and sent to the central adjudication office.
17) The screening period is from "the screening start date" to the day before "the pre-treatment observation period start date," and the screening start date and end date are defined to be consistent with the allowable time window of "the pre-

treatment observation period" (for example, if the pre-treatment observation period is 4 days, the maximum screening period is the 30 days from Day-34 to Day-5, and if the pre-treatment observation period is 10 days, the maximum screening period is the 30 days from Day-40 to Day-11).

At each study visit after the baseline date, the assessments that are performed by the patients themselves (WOMAC, PGA, joint pain questionnaire, SF-36, EQ-5D-5L, survey of autonomic symptoms) will be performed before all of the other assessments (including the assessments that are performed by the (sub) investigator) (except for cases where imaging tests are performed on a different day).

(3) Post-joint replacement procedure test/observation schedule

Follow-up day (visit window)	After surgery[1]	Long-term[1]
	Follow-up survey, 4 weeks post-operative	Follow-up survey 2, 20 weeks post-operative
	29 days post-operative (±5 days)	141 days post-operative (±7 days)
Concomitant medications	X	X
[Safety]		
Vital signs	X	X
Joint Pain Questionnaire	X	X
JOA score[2]	X	X
Radiology (bilateral knees, hips, shoulders)	X	X
Adverse events	←	→
At joint pain worsening (imaging [X-ray, MRI] assessments)3	←	→

1) Information about the procedure, including prosthesis replacement and/or the extent of surgery wound healing, will be collected.
2) The condition of the joint following joint replacement surgery will be assessed using the JOA Score.
3) If a sudden worsening of pain that does not occur in the normal progression of osteoarthritis occurs, the (sub) investigator will perform imaging test (X-ray and MRI) of the affected joint. Imaging test (X-ray and MRI) of other joints will be performed as needed by the (sub) investigator. Imaging tests will also be performed if the central adjudication office requests that additional tests be conducted. All images will be submitted to the central adjudication office.

3.3. Sample Size and Power Considerations

The original target sample size is 568 treated subjects (142 subjects per group)

(MT-5547 1 mg q4w group: 142 subjects; MT-5547 3 mg q4w group: 142 subjects; MT-5547 6 mg q8w group: 142 subjects; placebo group: 142 subjects)

After initiation of this study, based on the recommendations from the IDMC, the 3 mg q4w and 6 mg q8w groups were removed and 1 mg q8w group was newly added to the treatment groups.

The amended target sample size is 506 treated subjects, including the subjects who were randomized to 1 mg q4w/placebo groups before the amendment (approximately 40 subjects per group) and the subjects who are randomized to 1 mg q4w/1 mg q8w/placebo groups with the ratio of 1:1:1 after the amendment (142 subjects per group).

(MT-5547 1 mg q4w group: 182 subjects; MT-5547 1 mg q8w group: 142 subjects; placebo group: 182 subjects)

In addition, 57 patients each were enrolled in the MT-5547 3 mg q4w and 6 mg q8w groups prior to the protocol amendment (Ver. 04.00.00000) and were discontinued from study drug based on the recommendations from the IDMC. Thus, enrollment of 563 subjects is planned for this study in total.

The original rationale at the initiation of the study

Because this study is being conducted as a placebo-controlled study, in order to detect a statistical difference between the MT-5547 groups and the placebo group, the mean intergroup difference relative to placebo in both the change from baseline in the WOMAC pain subscale score and the change from baseline in the physical function subscale score at Week 16 for the MT-5547 3 mg SC q4w, 6 mg SC q8w, and 1 mg SC q4w groups was set at 1.1, and the standard deviation 2.5, on the basis of the results obtained in the overseas P2a study (R475-PN-0901) and the overseas P2b study (R475-PN-1227) (Table 3.3-1). The effect relative to placebo in the P3 study of 5 mg and 10 mg of tanezumab, another NGF inhibitor, was 0.81 to 1.21 in the WOMAC pain subscale score and 0.98 to 1.25 [24), 25)] in the WOMAC physical function subscale score, and it therefore appears that a mean difference between the MT-5547 groups and placebo of 1.1 is a realistic assumption. It also appears that the minimum clinically significant difference in each WOMAC score will be around 0.67 to 0.75 [26), 27)]. The aforementioned intergroup difference of 1.1 exceeds the minimum clinically significant difference, and therefore seems to be an appropriate value to use for the intergroup difference. Based on a one-sided significance level of 2.5% and multiplicity adjustment based on a gate keeping method using a graphical approach, the power for MT-5547 3 mg SC q4w, 6 mg SC q8w, and 1 mg SC q4w will be 90%, 88%, and 88%, respectively, for the primary endpoint, and 82%, 79%, and 79%, respectively, for the key secondary endpoint. A 2-sample t-test will be used for the tests in each step. Based on this assumption, the number of subjects who will complete 16 weeks of treatment in each group will be 120.

Furthermore, because the proportion of missing data for the primary endpoint and the key secondary endpoint at Week 16 is around 15% according to Table 3.3-1, it is being assumed that the proportion of missing data at Week 16 will be 15%, and the number of patients per group is therefore being set at 142.

The rationale after the changes in treartment groups

Assuming the absolute treatment difference of 1.1 between the MT-5547 1 mg SC q8w groups and placebo in both the change from baseline in the WOMAC pain subscale score and

the change from baseline in the physical function subscale score at Week 16 with an associated standard deviation of 2.5 based on the original rationale at the initiation of the study, the number of patients per group after the protocol amendment (Ver. 04.00.00000) is planned to 142 subjects per group. The structure of statistical hypotheses have been changed due to the removal of 3 mg q4w / 6 mg q8w groups and the addition of 1 mg q8w. It is noted that approximately 40 subjects per group were randomized to 1 mg q4w / placebo groups before the amendment. For evaluation on 1 mg q4w, since subjects are randomized to 1 mg q4w / placebo at the same time with the same ratio of 1:1 regardless of pre- or post- amendment, analyses will be performed with pooled comparable data including pre- and post- amendment data on 1 mg q4w / placebo (1 mg q4w group: 182 subjects; placebo group: 182 subjects). For evaluation on 1 mg q8w, analyses will be performed with only the post-amendment data (1 mg q4w group: 142 subjects; 1 mg q8w group: 142 subjects; placebo group: 142 subjects). For four null hypotheses, based on multiplicity adjustment based on a gate keeping method using a graphical approach (Figure 8.2.1-1 in section 8.2.1 of SAP for database at week 16), the power for MT-5547 1 mg SC q4w and 1 mg SC q8w will be 97% and 92% for the primary endpoint, 94% and 87% for the key secondary endpoint, respectively.

Table 3.1-1Change From Baseline at Week 16 in the WOMAC Pain and Physical Function Subscale Scores

	REGN Study R475-PN-1227					REGN Study R475-PN-0901			
	Placebo (N=83)	MT-5547 1 mg (N=85)	MT-5547 3 mg (N=84)	MT-5547 6mg (N=85)	MT-5547 9 mg (N=84)	Placebo (N=55)	MT-5547 0.03 mg/kg (N=53)	MT-5547 0.1 mg/kg (N=53)	MT-5547 0.3 mg/kg (N=54)
Change from baseline in WOMAC pain subscale score									
Baseline	6.43	6.33	6.35	6.10	6.53	5.9	5.7	6.1	6.4
n[1)] Mean[1)] (SD)[1)]	71 -2.43 (2.38)	75 -3.49 (2.06)	78 -3.39 (2.44)	77 -3.07 (2.34)	79 -3.81 (2.49)	44 -2.4 (2.18)	47 -2.7 (1.89)	44 -3.4 (2.53)	41 -3.2 (2.24)
LSM[2)]	-2.25	-3.35	-3.33	-3.03	-3.65	-	-	-	-
Diff. vs. Pbo[3)]	-	-1.10	-1.08	-0.78	-1.40	-	-0.6	-1.1	-0.8
P value	-	0.0025	0.0029	0.0304	0.0001	-	0.1486	0.0090	0.0488
Change from baseline in WOMAC physical function subscale score									
Baseline	6.15	6.11	6.09	5.94	6.20	5.9	5.9	6.2	6.2

n[1)]	70	75	78	76	80	44	47	44	41
Mean[1)]	-2.12	-3.21	-3.28	-2.97	-3.51	-2.3	-2.9	-3.4	-3.1
(SD)[1)]	(2.26)	(2.23)	(2.29)	(2.45)	(2.50)	(2.30)	(1.78)	(2.28)	(2.18)
LSM[2)]	-1.98	-3.08	-3.27	-3.03	-3.41	-	-	-	-
Diff. vs. Pbo[3)]	-	-1.10	-1.29	-1.06	-1.43	-	-0.8	-1.1	-0.9
P value	-	0.0019	0.0003	0.0029	< 0.0001	-	0.0693	0.0071	0.0245

1) n: number of subjects at Week 16; mean: arithmetic mean at Week 16; SD: standard deviation at Week 16. However, in REGN Study R475-PN-1227, the n includes WOMAC data from after treatment discontinuation.

2) LSM: MMRM least squares mean

3) Diff. vs. Pbo: Difference for each group relative to the placebo group in the MMRM least squares mean

4. PLANNED ANALYSIS

4.1. Interim Analysis

A First-Step Analysis of the data was conducted using database in 24 week. No alpha adjustment is necessary, as the week 16 efficacy analysis is the final primary analysis for efficacy.

4.2. Final Analysis

This SAP for 68 week will be finalized before database lock in 68 week. The analysis for week 68 will be conducted after database lock in 68 week.

4.3. Independent Data Monitoring Board

An independent data monitoring board will periodically assess unblinded data, and will make recommendations to the study sponsor regarding the appropriateness of continuing the study and about the need for protocol amendments.

The independent data monitoring board will include independent statistical and medical experts. For more details, the overseas independent data monitoring committee procedures (DMC Charter) will be followed.

5. ANALYSIS POPULATIONS

The analysis of efficacy will be performed in the full analysis set (FAS). The principal analysis set for efficacy will be the modified FAS (mFAS). The analysis of safety will be performed in the safety analysis set (SAF). The analysis of pharmacokinetics will be performed in the pharmacokinetic analysis set.

The analysis sets are defined below. For the contents of the data review meeting, refer to a separate materials.

(1) Efficacy analysis sets

1) Randomized subjects (RAND)

2) FAS

The FAS will consist of all randomized subjects except for the following subjects.

- Subjects who have not received study drug even once
- Subjects with no WOMAC pain subscale score (the primary endpoint) at baseline
- Subjects with no WOMAC pain subscale score (the primary endpoint) at any time point after randomization

3) Modified FAS (mFAS)

The mFAS will consist of the FAS, except for the following subjects.

- Subjects who are randomized to 3 mg q4w group
- Subjects who are randomized to 6 mg q8w group

(2) SAF

The SAF will consist of all randomized subjects except for the following subjects

- Subjects who did not receive study drug even once
- Subjects for whom absolutely no post-randomization safety data are available

(3) Self-injection analysis set(SIAS)

The self-injection analysis set will consist of the mFAS, except for the following subjects.

- Subjects who never performed self-injection

(4) Pharmacokinetic analysis set (PK)

The pharmacokinetic analysis set will consist of subjects who have received investigational drug at least once and who had at least one non-missing post-dose serum drug concentration data.

(5) Anti-drug antibody (ADA) analysis set (ADA)

The ADA analysis set will consist of all treated subjects who received investigational drug or comparative drug at least once and who had at least one non-missing post-dose ADA result.

(6) Neutralizing anti-drug antibody (NAb) analysis set (NAb)

The NAb population includes all treated patients who received any amount of study treatment (drug or placebo; Safety Analysis Set), and either tested negative for ADA or tested positive for ADA with at least one non-missing NAb result (either "NAb negative" or "NAb positive") following the first dose of study drug or placebo.

(7) Concentration-Response analysis set (CR)

The CR polulation other than CRP will include all patiens who received any study drug and who had at least one non-missing serum MT-5547 concentration following the first dose of study drug and at least one non-missing WOMAC pain subscale score, WOMAC physical function subscale score, alkaline phosphatase (ALP), or bone

specific alkaline phosphatase (BAP) value, as applicable for each C-R assessment. Placebo group patients who received any study treatment (i.e., placebo) and who had at least one non-missing WOMAC pain subscale score, WOMAC physical function subscale score, ALP or BAP value will also included as a reference for each C-R assessment other than CRP. For CRP, the CR polulation will include all patiens who received any study drug and who had at least one non-missing serum MT-5547 concentration following the first dose of study drug and at least one non-missing CRP value which measured by Latex method. The data after the discontinuation will be excluded in CR population. For CRP, Placebo group patients who received any study treatment (i.e., placebo) and who had at least one non-missing CRP value which measured by Latex method will also included as a reference.

6. STATISTICAL CONSIDERATIONS

6.1. Descriptive Statistics

(1) Non-PK related

Continuous data will be summarized descriptively using the number in the analysis set (N), the number of observations (n), mean, standard deviation (SD), median, minimum and maximum. Categorical data will be summarized using frequency counts and percentages. The denominator for the percentages will be the total number of subjects in the treatment group and analysis population being presented, unless otherwise specified.

(2) PK related

Serum concentrations will be summarized descriptively using N, n, mean, SD, median, Q1, Q3, minimum and maximum by clinical study day (as defined in the protocol) and treatment group.

6.2. Statistical Tests

All formal statistical tests of treatment effects will not be done. Point estimates will be accompanied with two-sided 95% CIs where applicable.

7. DATA CONVENTIONS

7.1. Analysis Variable Definitions

7.1.1. Study Subjects

7.1.1.1. Demographic and Other Baseline Characteristics

(1) BMI

BMI will be recalculated using the formula below and reported to 1dp.

BMI (kg/m^2) = weight at screening (kg) / {height at screening (m)}^{2}

(2) Age(years)

If (month of date of consent > month of date of birth) or (month of date of consent = month of date of birth and day of date of consent >= day of date of birth), then Age (years) = year of date of consent - year of date of birth

if else case, then Age (years) = year of date of consent - year of date of birth –1.

(3) Time since osteoarthritis diagnosis (years)

Time since osteoarthritis diagnosis will be calculated using the formula below and reported to 1dp.

If (month of date of consent > month of osteoarthritis diagnosis) or (month of date of consent = month of date osteoarthritis diagnosis, and day of date of consent >= day of date of osteoarthritis diagnosis), then Time since osteoarthritis diagnosis (years) = (date of consent – date of osteoarthritis diagnosis) / 365.25*

if else case, then Time since osteoarthritis diagnosis (years) = (date of consent – date osteoarthritis diagnosis – 1) / 365.25*.

* 1 year = 365.25 days

Impute date of osteoarthritis diagnosis as the earliest possible date (i.e. first day of month if day is unknown or 1st January if day and month are unknown).

(4) Renal function (stage 1, stage 2, stage 3, stage 4, stage 5)

The stages of renal function are defined by the FDA guidance as follows[1] :

Stage	Description[b]	eGFR[c] (mL/min/1.73m^2)
1	Control (normal) GFR	≥ 90
2	Mild decrease in GFR	60-89
3	Moderate decrease in GFR	30-59
4	Severe decrease in GFR	15-29
5	End Stage Renal Disease (ESRD)	<15 not on dialysis Requiring dialysis

7.1.1.2. Medical History

Medical history will be coded according to the MedDRA version 20.1.

7.1.1.3. Prior or Concomitant Medication

Medications will be coded according to the WHO Drug Dictionary (WHO-DD) Sep 2017 B3 version.

(1) Prior Medication

Prior medications are defined as medications starting prior to the first dose of study drug.

(2) Concomitant Medication

Medications that are on-going or start after the first intake of study drug are classified as "Concomitant".

Assesment period is defined as the time from first dose of investigational product up to 28 days after the last dose of investigational product.

(3) Post-treatment Medication

Medications that are start after on-treatment period are classified as "Post-treatment Medication".

7.1.1.4. Treatment Duration, Treatment and Post-Treatment Observation Duration and Compliance

(1) Treatment Duration

1) Treatment duration (days) = The date of last study drug injection - the date of the first study drug injection + 28.
2) Criteria for pre-defined limit
   Treatment duration period ≥1 day, ≥29 days, ≥57 days, ≥85 days, ≥113 days, ≥141 days, ≥169 days, ≥197 days, ≥225 days, ≥253 days, ≥281 days and ≥309 days

(2) Treatment and Post-Treatment Observation Duration

1) Treatment and Post-Treatment Observation duration (days) = Date of last study visit [up to end of the post treatment observation period] – date of first study drug dose) .
2) Criteria for pre-defined limit
   Treatment and Post-Treatment Observation duration period ≥1 day, ≥29 days, ≥57 days, ≥85 days, ≥113 days, ≥141 days and ≥169 days, ≥197 days, ≥225 days, ≥253 days, ≥281 days, ≥309 days, ≥337 days, ≥365 days, ≥393 days, ≥421 days and ≥449 days

(3) Treatment Compliance

1) Treatment compliance will be calculated using the formula below and reported to 1dp.

MT-5547

Treatment compliance(%) = (Number of injections of study drug during exposure period) / (Number of planned injections of study drug during exposure period on or before the time that the subject discontinues from the study) x 100%

Naproxen

Treatment compliance(%) = (Number of adherence of naproxen during exposure period) /

(Number of planned adherence of naproxen during exposure period on or before the time that the subject discontinues from the study) x 100%

2) Criteria for pre-defined limit

MT-5547

the number of injections of study drug: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12

the number of self-injection of study drug: 1, 2, 3, 4, 5 and 6

7.1.2. Efficacy assessments

7.1.2.1. WOMAC pain subscale score

$$\text{WOMAC pain subscale score} = \frac{\text{sum of items from Section A 1 to Section A5 in questionnaire}}{5}$$

7.1.2.2. WOMAC physical function subscale score

$$\text{WOMAC physical function subscale score} = \frac{\text{sum of items from Section C 8 to Section C24 in questionnaire}}{17}$$

7.1.2.3. WOMAC stiffness subscale score

$$\text{WOMAC stiffness subscale score} = \frac{\text{Section B 6+ Section B7 in questionnaire}}{2}$$

7.1.2.4. WOMAC total score

$$\text{WOMAC total score} = \frac{\text{sum of items from Section A1 to Section C24 in questionnaire}}{24}$$

7.1.2.5. SF-36 / EQ-5D-5L

- SF-36

  Physical, Mental and Role social component summary will be calculated considering factor coefficients based on the 1995 Japan National Survey. The derivation of 8 subscales and three component is detailed in appendix 1.

- EQ-5D-5L

  Utility Index Score will be calculated using the formula and Table below.[1]

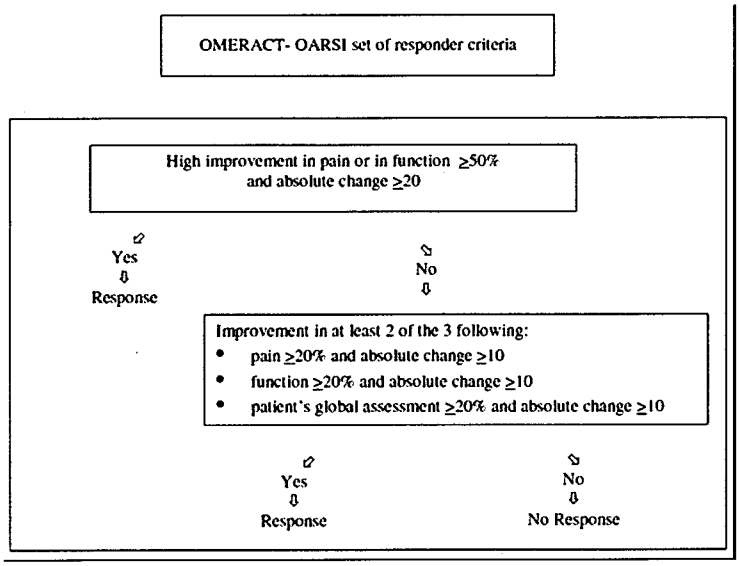
$$\text{Utility Index Score} = 1 + [\textit{Estimation of the constant term(if all answer is not 1)}] + [\textit{Sum of "estimation of coefficients corresponding to answer other than 1"}]$$

Item	Level	Estimation	Standard Error	P-value
Intercept		−0.060924	0.013625	<0.0001
Mo	2	−0.063865	0.008996	<0.0001
	3	−0.112618	0.009287	<0.0001
	4	−0.179043	0.010231	<0.0001
	5	−0.242916	0.009425	<0.0001
Sc	2	−0.043632	0.008931	<0.0001
	3	−0.076660	0.009972	<0.0001
	4	−0.124265	0.010129	<0.0001
	5	−0.159659	0.008924	<0.0001
Ua	2	−0.050407	0.009205	<0.0001
	3	−0.091131	0.010005	<0.0001
	4	−0.147929	0.009744	<0.0001
	5	−0.174786	0.009115	<0.0001
Pd	2	−0.044545	0.008354	<0.0001
	3	−0.068178	0.010052	<0.0001
	4	−0.131436	0.008985	<0.0001
	5	−0.191203	0.009604	<0.0001
Ad	2	−0.071779	0.009701	<0.0001
	3	−0.110496	0.010863	<0.0001
	4	−0.168171	0.009850	<0.0001
	5	−0.195961	0.009164	<0.0001

Mo: mobility, Sc: self-care, Ua: usual activities,
Pd: pain/discomfort, Ad: anxiety/depression

e.g) If answer is [1 1 1 1 1], Utility Index Score is $= 1 + 0 + 0 = 1$. If answer is [1 1 3 2 1], Utility Index Score is $= 1 + (-0.060924) + (-0.091131 - 0.044545) = 0.8034$.

7.1.2.6. Response rate based on the OMERACT-OARSI criteria



Note that the criteria in the diagram above are based on standardized score between 0 and 100. For this study, WOMAC pain and physical function subscale score are between 0 and 10, so the absolute change required for response is the required change in the diagram above divided by 10; PGA is 1, 2, 3, 4 or 5, so the absolute change required for response is at least 1 point.

7.1.2.7. Rescue Medication

Baseline amount of rescue medication is defined as the average of the non-missing values during 7 days prior to taking study drug. For each week, the average of the non-missing values between date collected WOMAC and next date collected WOMAC will be used.

7.1.3. Safety Assessments

The on-treatment period is defined as the time from first dose of investigational product up to 28 days after the last dose of investigational product.

The post-treatment period is defined as the time after the on-treatment period to the end of the post treatment observation period or early termination date.

7.1.3.1. Adverse Events

Adverse events will be coded according to the MedDRA version 20.1

(1) Treatment Emergent Adverse Events/ Treatment Emergent Serious Adverse Events (TEAEs/TESAEs)

An AE/SAE is classified as treatment emergent if it newly occurred after the first dose of study drug or if a pre-dose event increases in severity following the first dose of study drug during on-treatment period.

(2) Post-Treatment Adverse Events/ Treatment Serious Adverse Events

An AE/SAE is classified as post-treatment if it newly occurred after the on-treatment period or if the event after the on-treatment period increases in severity.

(3) Adverse Drug Reaction

A TEAE is considered "adverse drug reaction" if it has been assessed as having a "reasonable possibility" in relationship to the study drug.

(4) Adverse Event of Special Interest(AESI)

Adjudicated arthropathy(AA), sympathetic nervous system dysfunction, altered peripheral sensation, destructive arthropathy(DA) are defined as adverse events of special interest (AESI) for the sake of investigating the safety profile of MT-5547.

(5) Identified and Potential Adverse Events

Following are defined as identified and potentcial adverse event for the sake of investigating the safety profile of MT-5547.

- Arthralgia related event
  (PT) Arthralgia
  (PT) Osteoarthritis
  (PT) Arthritis
  (PT) Joint stiffness
  (PT) Joint swelling
  (PT) Joint effusion
  (PT) Synovitis

- Bone fractures related event
  Any PT containing "fracture" but excluding Tooth Fracture and Subchondral Insufficiency fracture (that have been identified to be AAs)

- Joint swelling related event
  (PT) Joint Swelling
  (PT) Joint Effusion

- Altered peripheral sensation related event
  (HLGT) Peripheral Neuropathies
  (HLT) Paraesthsias and dysaesthesias
  (HLT) Sensory abnormalities NEC
  (PT) Myalgia
  (PT) Hyperreflexia
  (PT) Hyporeflexia
  (PT) Areflexia

- Sympathetic nervous system dysfunction related event
  (HLT) Autonomic nervous system disorders
  (HLT) Cardiac conduction disorders
  (PT) Syncope

- Hypersensitivity / anaphylaxis related event
  (SMQ) Angioedema
  (SMQ) Anaphylactic reaction
  (SMQ) Hypersensitivity
  (PT) Erythema
  (PT) Prurigo
  (PT) Pruritus
  (PT) Pruritus Generalised
  (PT) Rash Papular

(6) Duration of Adverse Events
Duration of Adverse Events (days) = AE stop date – AE start date + 1
(7) Incidence Rate (per 100 patient-years)
For patients with events, number of patient years is calculated up to the date of the first event and for patients without event, it corresponds to the length of the study observation period.
Incidence Rate will be calculated using the formula.
Incidence Rate = the number of patients with event × 100 / sum of observation periods for all patients (days) × 365.25*
* 1 year = 365.25 days

7.1.3.2. Laboratory Tests

Listing of reference value in each laboratory parameters is detailed in appendix 3
(1) The Estimated Glomerular Filtration Rate (eGFR)
eGFR will be calculated as follows and reported to 1 dp.
$eGFR = 141 \times \min(S_{Cr}/\kappa, 1)^{\alpha} \times \max(S_{Cr}/\kappa, 1)^{-1.209} \times 0.993^{Age} \times 1.018$ [if female] × 1.159 [if Black]
S_{Cr} (standardized serum creatinine) = mg/dL
$\kappa = 0.7$ (females) or 0.9 (males)
$\alpha = -0.329$ (females) or -0.411 (males)
min = indicates the minimum of S_{Cr}/κ or 1

max = indicates the maximum of S_{Cr}/κ or 1

Age = years

(2) Laboratory values below the limit of quantification

1/2 LLOQ (lower limit of quantification) will be used for BLQ (below the limit of quantification) data in summary statistics. ULOQ (upper limit of quantification) will be used for over the limit of quantification data in summary statistics.

(3) Criteria for pre-defined

Potentially clinically significant values (PCSV) ranges will be applied to the laboratory test values as applicable (see Appendix 2 for PCSV definitions).

7.1.3.3. Vital Signs

(1) Criteria for pre-defined

Potentially clinically significant values (PCSV) ranges will be applied to the vital sign parameter values as applicable (see Appendix 2 for PCSV definitions).

7.1.3.4. ECG

(1) Criteria for pre-defined

Potentially clinically significant values (PCSV) ranges will be applied to the ECG parameter values as applicable (see Appendix 2 for PCSV definitions).

(2) QTcF

QTcF will be calculated as follows and reported to 1 dp.

$QTcF\ (msec) = QT/RR^{1/3}$

$RR = 60/HR$

7.1.4. Pharmacokinetics Evaluation

7.1.4.1. MT-5547 Concentrations in Serum

For the calculation of the summary statistics, concentration values reported as below the limit of quantification (BLQ) will be set to 0. In the log-scaled figures, concentrations below the LLOQ imputed as LLOQ/2. The lower limit of quantification (LLOQ) is 7.8 ng/mL in neat human serum.

7.1.4.2. Anti-Drug Antibodies in Serum

Anti-drug antibody variables will include ADA status (positive or negative) and titer as follows:

- Total subjects negative in the ADA assay analyzed at all time points
- Pre-existing immunoreactivity - a positive ADA response at baseline with all post-dose

ADA results negative, or a positive ADA response at baseline with all post-dose ADA responses less than 9-fold over baseline titer levels.

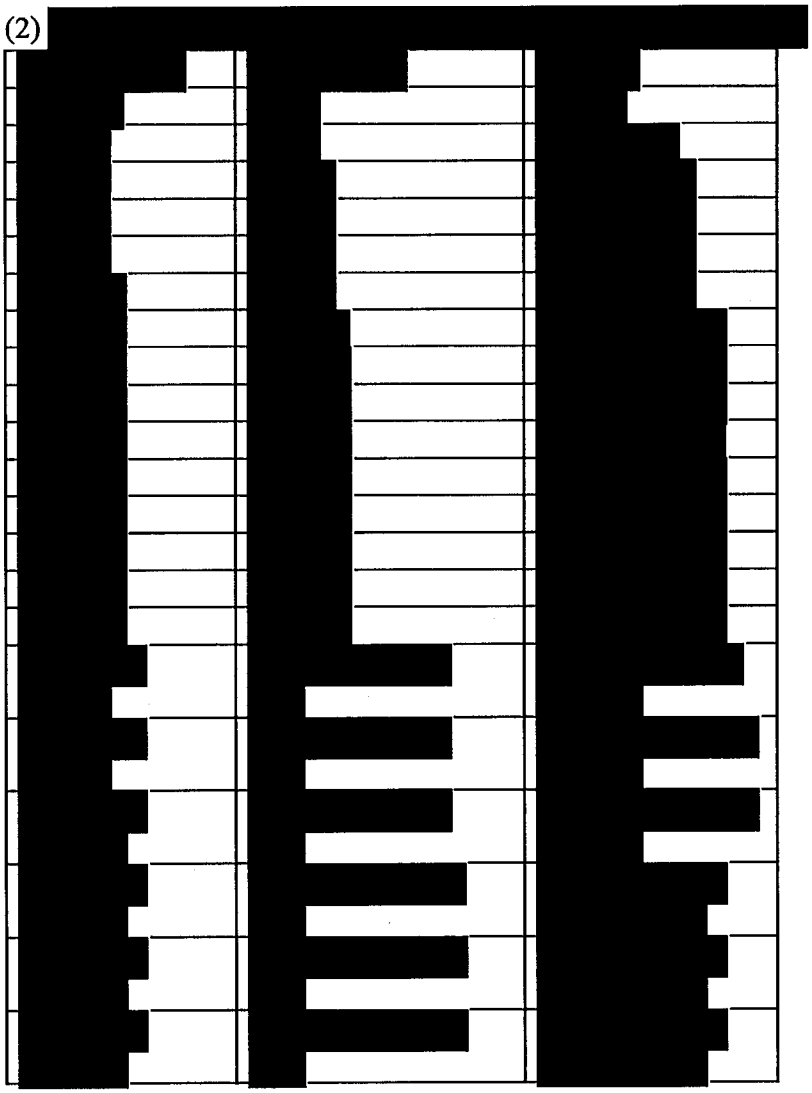
- Treatment emergent - defined as any post-dose positive ADA response when baseline results are negative
  - Persistent - A positive result in the ADA assay detected in 2 consecutive post baseline samples, separated by at least 16-week period [based on nominal sampling time], with no ADA-negative results in-between, regardless of any missing samples
  - Indeterminate - A positive result in the ADA assay at the last collection time point analyzed only, regardless of any missing samples
  - Transient - Not persistent or indeterminate regardless of any missing samples
- Treatment boosted - defined as any post-dose positive ADA response that is at least 9-fold over the baseline level when baseline is positive in the ADA assay
- Titer Values
  - Titer category: low (titer <1,000); moderate (1,000≤ titer ≤10,000); high (titer >10,000)
- Neutralizing ADA activity for samples positive in the ADA assay

7.2. Analysis Visit Definitions

(1) Analysis visit definitions for efficacy endpoints and PK endpoints (MT-5547 Concentrations, ADA and Concentration-Response Relationships)

Analysis visit	Nominal day	Window
Baseline	Day 1	<Day 1
Week 1	Day 8	Day 2 to 11
Week 2	Day 15	Day 12 to 21
Week 4	Day 29	Day 22 to 43
Week 8	Day 57	Day 44 to 71
Week 12	Day 85	Day 72 to 99
Week 16	Day 113	Day 100 to 127
Week 20	Day 141	Day 128 to 155
Week 24	Day 169	Day 156 to 183
Week 28	Day 197	Day 184 to 211
Week 32	Day 225	Day 212 to 239
Week 36	Day 253	Day 240 to 267
Week 40	Day 281	Day 268 to 295
Week 44	Day 309	Day 296 to 323
Week 48	Day 337	Day 324 to 351
Week 52	Day 365	Day 352 to 379
Week 56	Day 393	Day 380 to 407
Week 60	Day 421	Day 408 to 435
Week 64	Day 449	Day 436 to 463
Week 68	Day 477	Day 464 to 491

(2)



(3) For Radiology analysis visit definitions

Analysis visit	Nominal day	Window
Baseline	Day 1	<Day 1
Week 16	Day 113	Day 86 to 141
Week 48	Day 337	Day 310 to 365
Week 68	Day 477	Day 450 to 505

The date of the first dose of study drug is defined as Day 1.

Unless otherwise specified, baseline will be the last observed value of the parameter of interest prior to the first intake of study drug (this includes unscheduled visits). For other visits, if there

are multiple data in a window, the closest data to nominal day will be used. If the distance to the nominal day is the same, the data of later date will be used. For orthostatic blood pressure, if a retest is performed, the initial value is compared with the retest value, and the worst value is used.

7.3. Data Handling Convention for Missing Data

(1) Non-PK related

Efficacy:

WOMAC

WOMAC scores will be computed when one pain, one stiffness, or 1-3 physical function items are missing. The missing items will be imputed by the mean of available items within the same subscale. The scores will be set to missing if more items are missing.

SF-36 / EQ-5D-5L

- SF-36 subscale scores will be computed if at least 50% of items are available. The missing items will be imputed by the mean of available items.
- EQ-5D-5L index will be set to missing if any of the 5 dimensions is missing.

Rescue Medication

If there is no date collected WOMAC, nominal date will be used.

Safety:

Adverse events

If severity or relationship is found to be missing the most severe occurrence will be imputed for the summary of interest.

For AE start missing or partial dates, the AE will be treated as TEAE if it cannot be determined to be a non-TEAE.

Laboratery test

Regarding hyaline casts and epithelial cells, set "Null" for blanks(= not detected) with date data.

Orthostatic Blood Pressure

If a retest is performed, the initial value should be compared with the retest value, and the value with the worst value should be used.

Other safety:

For safety summaries, only observed data will be used. If a value is indeterminate or a reference

value because of a problem with a test sample, then this value will be handled as a missing value.

(2) PK related
For PK summaries, only observed data will be used. Missing PK data will not be imputed.
The PK data handling will be confirmed during blinded data review (BDR). PK data that are considered "invalid" will be flagged in the listing. Due to the nature of PK data, some issues may only be discovered after PK data are unblinded. Should new issues identified post unblinding, and new data handling rules would have to be applied, a separate PK data handling document will be produced to provide detailed rationale and decision making. If there is clear evidence that PK sample handling errors or other factors identified after data unblinding and these errors have led to unexpected erroneous data, then these erroneous data will be regarded as "invalid", full explanations will be given in the PK data handling document.

8. STATISTICAL METHODOLOGY

8.1. Study Subjects

8.1.1. Subject Disposition
Screened subject disposition will be summarized on the all subjects. Randmized subject disposition will be summarized by treatment group, combined MT-5547 and overall on the randomized population. Randomization details will be listed on the randomized population. Subject disposition will be listed on the randomized population.

8.1.2. Analysis Populations
Analysis populations will be summarized by treatment group, combined MT-5547 and overall on the randomized population. Analysis populations will be listed on the randomized population.

8.1.3. Site Enrollment
Investigator enrollment will be summarized by treatment group and overall on the randomized population.

8.1.4. Demographic and Other Baseline Characteristics
The following demographic and other baseline characteristics will be used.

	category	Descriptive
Sex	Male, Female	
Age(years)	<65, >=65	Yes
Height(cm)		Yes

Weight(kg)		Yes
BMI(kg/m2)	<25, >=25 and<30, >=30	Yes
Race	Japanese, other	
Time since OA diagnosis (years)	<1, >=1 and <5, >=5 and <10, >=10	Yes
Complication	Yes, No	
Index Joint	Knee, Hip	
K-L score	1, 2, 3, 4	
Renal function	stage 1, stage 2, stage 3, stage 4, stage 5	
WOMAC pain subscale score	<6 >=6	Yes
WOMAC Physical function subscale score	<6 >=6	Yes
PGA	1, 2, 3, 4, 5	Yes
Pain site	Right Knee, Left Knee, Right Hip, Left Hip, Right Shoulder, Left Shoulder	
Pretreatment drug	NSAIDs(Systemic use), NSAIDs(Topical use), Acetaminophen, Opioid, Monoamine reuptake inhibitor, Hyaluronic acid, Pregabalin, Herbal medicine, Other	
NSAIDs Inadequate Pain Relief	Yes, No	
NSAIDs Intolerance	Yes, No	
Ever taken Opioid for Pain due to OA	Yes, No	
Opioid Inadequate Pain Relief	Yes, No	
Opioid Intolerance	Yes, No	
Ever taken Acetaminophen for Pain due to OA	Yes, No	
Acetaminophen Inadequate Pain Relief	Yes, No	
Acetaminophen Intolerance	Yes, No	

NSAID + Opioid + Acetaminophen Inadequate Pain Relief	Yes, No	

Demographic and other baseline characteristics will be summarized by treatment group, and before the protocol amendment and after the protocol amendment on the mFAS population. Demographic and other baseline characteristics will be summarized by treatment group, combined MT-5547 and overall on the FAS population. Demographic and other baseline characteristics will be listed on the randomized population.

8.1.5. Medical History

Medical history will be summarized by treatment group, combined MT-5547 and overall on the FAS population. Medical history will be listed on the randomized population.

8.1.6. Radiograph eligibility data

Radiograph eligibility data at pre-treatment will be summarized by treatment group, combined MT-5547 and overall on the FAS population. Radiograph data will be listed on the randomized population.

8.1.7. Prior or Concomitant Medications

Prior medication used for OA and concomitant medication and post-treatment medication will be summarized separately by treatment group and combined MT-5547 on the FAS population. All prior and concomitant medication and concomitant therapy and post-treatment therapy will be listed on the randomized population.

8.1.8. Treatment Duration, Treatment and Post-Treatment Observation Duration and Compliance

Treatment duration and treatment and post-treatment observation duration will be summarized by treatment group and combined MT-5547 on the SAF population. Treatment compliance of MT-5547 and naproxen will be summarized by treatment group and combined MT-5547 on the FAS population. The number of MT-5547 self-injection will be summarized by the principal investigator or sub-investigator and the self-injections and by MT-5547 treatment group on the mFAS population.Treatment duration, treatment and post-treatment observation duration and compliance will be listed on the randomized population. Study drug interruptions will be listed on the randomized population.

8.2. Efficacy Assessments

Efficacy assessment will be summarized by treatment group on the mFAS population.

8.2.1. Other Efficacy Endpoints

WOMAC pain subscale score, WOMAC physical function subscale score, WOMAC

stiffness subscale score, WOMAC total score , WOMAC individual score and PGA

For the changes from baseline at each assessment time point in the WOMAC pain subscale score, the WOMAC physical function subscale score, the WOMAC stiffness subscale score, the WOMAC total score, WOMAC individual score and PGA, the analysis will be tested using the estimator of mixed effect model for repeated measurement (MMRM) approach. The model will include the treatment group, measurement time point, treatment group and measurement time point interaction, stratification factor K-L grade (2-3 or 4), and stratification factor evaluated joint (knee or hip) as factors, and the corresponding baseline and baseline and measurement time point interaction as covariates. The within-subject correlation structure will be unstructured, and the denominator degrees of freedom for the test statistics of estimated parameters will be approximated using the Kenward-Roger method. In case the model will not converge with the unstructured covariance structure, Heterogeneous Autoregressive (1) (ARH[1]), Heterogeneous Compound Symmetry (CSH), Autoregressive (1) (AR[1]) or Compound Symmetry(CS) covariance structure will be used instead in this order. MMRM that assumes that the missing mechanism is missing at random (MAR) is an analysis method in which missing values are not replaced. This model will be used to estimate the treatment effect in the mFAS when subjects continue receiving treatment through Week 68.

The sampel SAS code planned for the analysis is outlined below.

[redacted]

WOMAC score and PGA will be listed. In addition, change from baseline to week 68 in WOMAC pain subscale score, WOMAC physical function subscale score and PGA will be graphed by visit.

SF-36 / EQ-5D-5L

For the changes from baseline at each assessment time point in SF-36 eight subscale summary scores and three component summary scores / EQ-5D-5L visual analogue scale(VAS) and utility index scores, the same analysis method (MMRM) as above will be analyzed. SF-36 / EQ-5D-5L will be listed.

Proportions of subjects with improvement of 30%, 50%, 70% and 90% at each assessment time point compared to baseline in WOMAC pain subscale score, WOMAC physical function subscale score and improvement of 30%, 50% and 70% at each assessment time point compared to baseline in PGA

A generalized linear mixed-effects model (GLMM) by logit type for response variable will be used to evaluate these binary response variables. The denominator degrees of freedom for the test statistics of estimated parameters will be approximated using the Kenward-Roger method. The estimated response rate for each treatment group and placebo group and 95% confidence interval of the odds ratio of each of the MT-5547 1 mg q4w and MT-5547 1 mg q8w groups compared to the placebo group in each binary response variables will be calculated using this GLMM approach.

- The covariates: baseline and the interaction of baseline and measurement time point.
- The fixed factors: Treatment group, measurement time point, the interaction of treatment group and measurement time point, K-L grade (2-3 or 4), and evaluated joint location (knee or hip)
- Random factors: Measurement time point for subject specific with covariance structure of unstructured

The sample SAS code planned for the analysis is outlined below.

[redacted]

Response rate based on the OMERACT-OARSI criteria at each assessment time point compared to baseline

Responder analyses for OMERACT-OARSI will be summarized at each assessment time point using the same GLMM approach in previous model.

Proportion subjects who used rescue medication

The proportion of subjects taking rescue medication in baseline and up to week 48 will be summarized by treatment group. The proportion of days on rescue medication use during treatment priod will be summarized at each week by treatment group. Number of average usage of rescue medication will be summarized at each week by treatment group.

8.2.2. Assessment of self-injection

Analysis with respect to a comparison of the doses administered by the principal investigator or sub-investigator up through Treatment Week 24 and the self-injections performed after Treatment Week 24

The doses administered by the principal investigator or sub-investigator up through Treatment Week 24 and the self-injections performed after Treatment Week 24 will be analyzed using the same method[Section 8.2.1] as for the WOMAC pain subscale score.
Theses analysis will be summarized based on the self-injection population.

8.3. Safety Assessments

Safety assessments will be made on the SAF population.

8.3.1. Adverse Events

Overall summary of TEAEs / Post-treatment AEs for the following will be conducted.

- Subjects with at least one AE
- Subjects with at least one adverse drug reaction
- Subjects with at least one SAE
- Subjects with at least one serious adverse drug reaction
- Subjects with at least one AESI
- Subjects with at least one severe AE
- Subjects with at least one AE leading to discontinuation of study drug
- Subjects with AE leading to death

The following summaries also will be conducted.

- TEAEs / Post-treatment AEs AEs by SOC and PT
- TEAEs / Post-treatment AEs related to drug reactions by SOC and PT
- TEAESI / Post-treatment AESI by SOC and PT
- TESAEs / Post-treatment SAEs by SOC and PT
- TESAEs related to drug reactions by SOC and PT
- TEAEs leading to discontinuation of study drug by SOC and PT

- TEAEs / Post-treatment AEs by SOC, PT and severity
- TEAEs related to drug reactions by SOC, PT and severity
- TEAEs by SOC, PT and relationship to study drug
- TESAEs by SOC, PT and relationship to study drug
- Identified TEAEs / Post-Treatment AEs by SMQ, HLGT, HLT and PT

Each of the summaries will be done by treatment group and combined MT-5547 at the subject level - multiple occurrences of the same event within a subject will be counted once in the summaries by SOC and PT; multiple occurrences of the same event within a subject will be counted once in the maximum severity category (severe > moderate > mild) and/or maximum drug relationship category (reasonable possibility/no reasonable possibility) and/or the earliest duration.

The following will be listed.

- AE
- AESI
- AE leading to discontinuation of study drug
- SAE
- Death
- Joint replacements

adjudicated arthropathy

The number and percentage of patients with images requiring adjudicated arthropathy as well as the number and percentage of those patients with confirmed adjudicated arthropathy including all available follow-up data after baseline will be summarized by treatment group and combined MT-5547. Subtypes of AAs and outcomes of AAs will also be summarized. Time to event with AA will be analyzed using Kaplan-Meir approach and will be plotted by treatment group. Cox regression model will be used for descriptively comparing each treatment group to placebo by obtaining hazard ratio estimates along with 95% confidence intervals.

Incidence of AA will be summarized by treatment group and combined MT-5547.

Incidence rate ratio estimates and 2-sided 95% confidence interval will be computed using Generalized Estimating Equations with a log-linked Poisson model via the PROC GENMOD procedure if there are events in the placebo group. The model will specify the use of a Poisson probability distribution with an offset= log (patient-observation period days) and model terms for treatment-as-received and each randomization strata. If there is no event in the placebo group, incidence rate ratio will not be calculated and confidence intervals for the incidence rate difference will be computed using Wald method.

The incidence rate estimates for each arm, the estimates of the incidence rate ratio and incidence rate difference between MT-5547 and placebo will be summarized.

The SAS code planned for the analysis is outlined below.

[redacted]

Assessment of self-injection

—Analysis with respect to a comparison of the doses administered by the principal investigator or sub-investigator after treatment week 24 and the self-injections performed after treatment week 24

This anaylsis will be summarized using the data after treatment week 24 by the principal investigator or sub-investigator and the self-injections and by MT-5547 treatment group excluding MT-5547 3mg q4w and MT-5547 6mg q8w on SAF population.

Overall summary for the following will be conducted.

- Subjects with at least one TEAE
- Subjects with at least one adverse drug reaction
- Subjects with at least one TESAE
- Subjects with at least one serious adverse drug reaction
- Subjects with at least one AESI
- Subjects with at least one severe TEAE
- Subjects with at least one TEAE leading to discontinuation of study drug
- Subjects with TEAE leading to death
- Subjects with at least one abnormal NCS or CS for injection site assessment

The following summaries also will be conducted.

- TEAEs by SOC and PT

8.3.2. Laboratory Tests

Absolute values and changes from baseline will be summarized at each assessment time point by treatment group for the following laboratory tests parameters.

Laboratory Test	Parameters

Hematology	RBC, Hemoglobin, Hematocrit, MCHC, WBC, Platelet, Neutrophils, Lymphocytes, Monocytes, Basophils, Eosinophils
Biochemistry	AST, ALT, ALP, LDH, Gamma-GTP, Total protein, Albumin, Total cholesterol, Total bilirubin, BUN, Serum creatinine, eGFR, Na, K, Cl, Ca, HCO_3^-, Uric acid, Triglycerides, BAP, Inorganic phosphorus, CPK, CRP
Urinalysis	pH, Specific gravity , Creatinine, Phosphate (Qualitative value) : Glucose, Protein, Occult blood, Bilirubin, Nitrite, WBC, Sendiment WBC, Hyaline casts, Epithelial cells, Bacteria, Yeast, Ketones

Shift table of clinically relevant categories will be presented for the following laboratory tests parameters.
All data will be listed.

8.3.3. Vital Signs

Absolute values and changes from baseline will be summarized at each assessment time point by treatment group for the following parameters.

- Systolic Blood Pressure (mmHg)
- Diastolic Blood Pressure (mmHg)
- Pulse Rate (bpm)
- Body Temperature(°C)

The percentage of subjects with vital signs pre-defined limit will be summarized.
All data will be listed.

8.3.3.1. Assessment of Orthostatic Blood Pressure with standing value minus supine value

Absolute values and changes from baseline will be summarized at each assessment time point by treatment group for the following parameters.

- Systolic blood pressure 1 minute after standing
- Diastolic blood pressure 1 minute after standing
- Pulse rate 1 minute after standing
- Systolic blood pressure 3 minute after standing
- Diastolic blood pressure 3 minute after standing
- Pulse rate 3 minute after standing

8.3.4. ECGs

Absolute values and changes from baseline will be summarized at each assessment time point by treatment group for the following parameters.

- Heart Rate (bpm)

- PR (msec)
- RR (msec)
- QRS (msec)
- QT (msec)
- QTcF (msec)

The percentage of subjects with ECG values outside pre-defined limit will be summarized.
All data will be listed.

8.3.5. Physical Examinations

Physical examination will be listed for the following parameters.

- General Appearance
- Neurological
- Eyes
- Ears, nose, throat
- Cardiovascular
- Respiratory
- Abdominal
- Hepatic
- Gastrointestinal
- Musculoskeletal
- Dermatological

8.3.6. Bone density

Bone density will be summarized and listed for the following parameters.

- YAM(%) and T-score

- Femoral neck
- Total hip
- Lumber Vertebra

8.3.7. JOA score

JOA score will be listed. Spaghetti plots including all available follow-up data after baseline will be presented by treatment group and JR joint and the corresponding to joint.

8.3.8. Radiograph

Adjudicated Arthropathy
Overall summary for adjudicated arthropathy will be conducted.
The following summaries for adjudicated arthropathy also will be conducted.

- Subtype of adjudicated arthropathy
- Subtype of adjudicated arthropathy meeting destructive arthropathy

- Incidence rate of adjudicated arthropathy
- Time to first event of adjudicated arthropathy

As regards time to first event adjudicated arthropathy, cox regression model will be used for descriptively comparing each treatment group to placebo by obtaining hazard ratio estimates along with 95% confidence intervals.

Each of the summaries will be done by treatment group and combined MT-5547.

All data will be listed.

Jonint Space Width

The following summaries for joint space width will be conducted.

- Non arthropathy adjucation joint with adjudicated arthropathy
- All joint with non adjudicated arthropathy subjects

Joint space width for AA will be summarized by treatment group and combined MT-5547. All data will be listed. Each of the summaries will be done by treatment group and combined MT-5547.

8.3.9. Injection Site Assessments

Injection Site Assessments will be summarized at each assessment time point by treatment group. All data will be listed.

8.3.10. Autonomic Symptoms Questionnaire

Absolute values and changes from baseline will be summarized at each assessment time point by treatment group for the following tests.

- total symptoms score
- total symptom impact scores

All data will be listed.

8.3.11. Neurological Examinations

Neurological Examinations will be summarized by treatment group for the following tests.

- Cranial Nerve Exam
- Sensory Function
- Motor Function
- Reflexes
- Coordination

8.3.12. Joint Pain Questionnaire

Joint pain questionnaire will be listed.

8.3.13. Joint Replacement

Number and percentage of patients with joint replacement including all available follow-up data after baseline will be summarized by treatment group and combined MT-5547. Time to event with joint replacement including all available follow-up data after baseline will be analyzed using Kaplan-Meir approach and will be plotted by treatment group.
All data will be listed.

8.4. Pharmacokinetics Evaluation

8.4.1 Analysis of MT-5547 Concentration in Serum

Serum MT-5547 concentrations will be summarized at each nominal sampling point for each treatment group. All serum concentrations will also be listed.
Plots of the mean (+SD) concentrations (linear scales or log scales) will be presented by nominal time. Plots of the individual concentrations (linear scales or log scales) will be presented over actual sampling time. When the scale is linear, concentrations below the LLOQ are set to zero. In the log-scaled figures, concentrations below the LLOQ imputed as LLOQ/2.

8.4.1. Analysis of Anti-Drug Antibody in Serum

Anti-MT-5547 antibody status and titers will be presented by nominal time.
The ADA variables described in Section 7.1.3.3 will be summarized using descriptive statistics by group/drug product in the ADA analysis set. Prevalence of pre-existing immunogenicity, treatment-emergent and treatment-boosted ADA will be assessed as absolute occurrence (N) and percent of patients (%), grouped by groups. For treatment emergent ADA, occurrence (N) and percent of patients (%) with persistent, transient and indeterminate ADA will be reported.
The influence of treatment-emergent or treatment-boosted ADA assay response on individual PK profiles will be evaluated.
Listings of ADA status presented by subject, time point, and group will be provided.
Listings of ADA positivity and titers presented by subject, time point, and group will also be provided.
Correlation analysis of safety versus ADA and NAb positivity status may be performed on the SAF. Assessment will focus on the following safety events:

- Hypersensitivity (SMQ : Hypersensitivity [Narrow])
- Anaphylaxis (SMQ: Anaphylaxis[Narrow])

Number (%) of patients with the above mentioned safety events may be summarized by ADA positivity status, during the TEAE period.

In addition, correlation analysis of key efficacy endpoints versus ADA and NAb status may be summarized.

8.4.2. Assessment of Concentration-Response Relationships

The scatter plot of each efficacy (WOMAC pain subscale score and WOMAC physical function subscale score) and safety variables (alkaline phosphatase, CRP and BAP value) vs. MT-5547 concentration at each time point by clinical study week will be presented for each dosing regimen, q4w (1 mg q4w and 3 mg q4w) and q8w (1 mg q8w and 6 mg q8w) and all dosing regimen (1 mg q8w, 1 mg q4w, 3 mg q4w and 6 mg q8w). Regarding the scatter plot of each efficacy (WOMAC pain subscale score and WOMAC physical function subscale score), the data after the discontinuation will be excluded. For safety variables, additional enlarged figures will be presented as needed.

In all plots, concentrations below the LLOQ were imputed to zero.

Additional analyses will be performed as necessary.

In accordance with the minutes of Blind Data Review Meeting conducted on 14th of Apr 2021 (BDRM), C-R analysis for CRP will include only the patients who have CRP values measured by Latex method throughout the study (Item No. 3 below).

Type	C-R analysis
Only measured as hsCRP	Not included
Baseline was measured as hsCRP, which was changed to CRP during the study.	Not included
Only measured as CRP	Included

8.4.3. [redacted]

9. DATA PRESENTATION CONVENTIONS

9.1. Number of Digits to Report

(1) Non-PK related

Statistic	Specification	Apply to
Minimum, Maximum	Same number of DPs as the data provided in the datasets	All original (i.e. non-derived)
	see section 7.3	All derived data
Mean, Median, SD, SE, Confidence intervals	One more DP than above	All
Percentages[*1]	1 DP	All
Ratios	3 DPs	All

*1

Percentages: use 1 place beyond the decimal point, except for the following cases:

If the percentage is equal to 0, then leave blank, do not use (0)

If the percentage is equal to 100, then use "(100)" without a decimal

(2) PK and [redacted] Concentration in Serum

Statistic	Specification
Individual value	3 significant digits
Minimum, Maximum	Same number of DPs as the individual value
Mean, SD, Median	Same number of DPs as the individual value

9.2. Treatments to Report

Treatment	For TFLs
MT-5547 1 mg q8w	MT-5547 1 mg q8w
MT-5547 1 mg q4w	MT-5547 1 mg q4w
MT-5547 1 mg q4w and MT-5547 1 mg q8w	MT-5547 Combined
MT-5547 3 mg q4w	MT-5547 3 mg q4w
MT-5547 6 mg q8w	MT-5547 6 mg q8w
Placebo	Placebo
Placebo group and all active dose levels	Total

9.3. Analysis Visits to Report

(1) Non-PK related

Efficacy:

Analysis Visit	Apply to	
	WOMAC and PGA	SF-36 and EQ-5D-5L
Baseline	X	X
Week 1	X	
Week 2	X	
Week 4	X	X
Week 8	X	X
Week 12	X	
Week 16	X	X
Week 20	X	
Week 24	X	X
Week 28	X	
Week 32	X	X
Week 36	X	
Week 40	X	X
Week 44	X	
Week 48	X	X
Week 52	X	
Week 56		
Week 60	X	
Week 64		

Week 68	X	X

Safety:

Analysis Visit	Apply to			
	Laboratory Tests	Injection Site Assessments	Survey of Autonomic Symptoms	Vital Signs, Joint pain questionnaire
Baseline	X	X	X	X
Week 1				X
Week 2				X
Week 4	X	X	X	X
Week 8	X	X	X	X
Week 12	X	X	X	X
Week 16	X	X	X	X
Week 20		X	X	X
Week 24	X	X	X	X
Week 28		X	X	X
Week 32	X	X	X	X
Week 36		X	X	X
Week 40	X	X	X	X
Week 44		X	X	X
Week 48	X		X	X
FU Week 4	X		X	X
FU Week 8			X	X
FU Week 12				
FU Week 16				
FU Week 20				
FU Week 24	X		X	X
End of last observation	X			

Analysis Visit	Apply to				
	ECGs, Bone dencity*	Physical Examinations*	Neurological Examinations, Orthostatic Hypotension	Radiograph	JOA score**
Baseline	X	X	X	X	
Week 1					
Week 2			X		
Week 4			X		
Week 8			X		
Week 12			X		
Week 16			X	X	

Week 20			X		
Week 24		X	X		
Week 28			X		
Week 32		X	X		
Week 36			X		
Week 40			X		
Week 44			X		
Week 48	X	X	X	X	
FU Week 4			X	X	
FU Week 8			X		
FU Week 12					
FU Week 16					
FU Week 20					
FU Week 24		X	X	X	
End of last observation	X				

Unscheduled visits, retests (same visit number assigned) and follow-up visits will not be displayed in by-visit summary tables, but will be included in the data listings.

*Baseline is the screening assessment time point

**Assessmets of the joint by JOA score at discontinuation will be performed for subjects who required joint replacement surgery during the study.

(2) PK related

MT-5547 Concentrations, ADA and Concentration-Response Relationships for efficacy valiables (WOMAC pain subscale score and WOMAC physical function subscale score) :

Analysis Visit	Apply to	
	MT-5547 Concentrations and Concentration-Response Relationships for efficacy valiables (WOMAC)	ADA
Baseline	X	X
Week 1	X	
Week 2	X	
Week 4	X	
Week 8	X	
Week 12	X	
Week 16	X	X
Week 24	X	X
Week 32	X	
Week 40	X	
Week 44	X	
Week 48	X	X
Week 52	X	

Week 56	X	
Week 68	X	X
End of treatment		X

Concentration-Response Relationships for safety valiables (ALP, CRP and BAP) and [redacted]:

Analysis Visit	Apply to	
	Concentration-Response Relationships for safety valiables (ALP, CRP and BAP)	[redacted]
Baseline	X	X
Week 1	X	
Week 2	X	
Week 4	X	X
Week 8	X	
Week 12	X	
Week 16	X	X
Week 24	X	X
Week 32	X	X
Week 40	X	
Week 44	X	
Week 48	X	X
FU Week 4	X	X
FU Week 8		
FU Week 12		
FU Week 16		
FU Week 20		
FU Week 24	X	X

10. CHANGE FROM THE PROTOCOL

There are currently no changes to analysis from protocol.

11. SOFTWARE

All statistical analyses will be performed using SAS version 9.3 or higher.

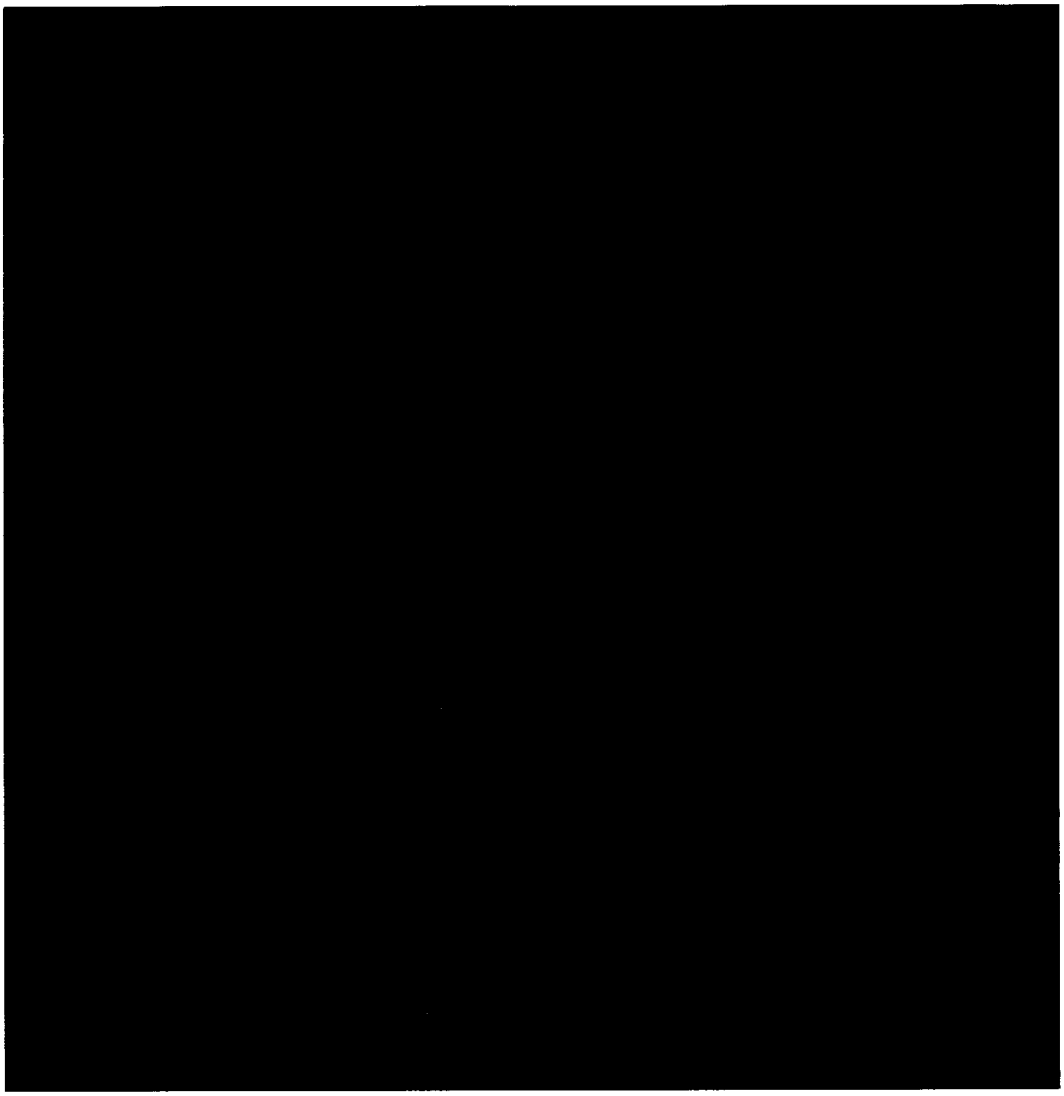
12. REFERENCES

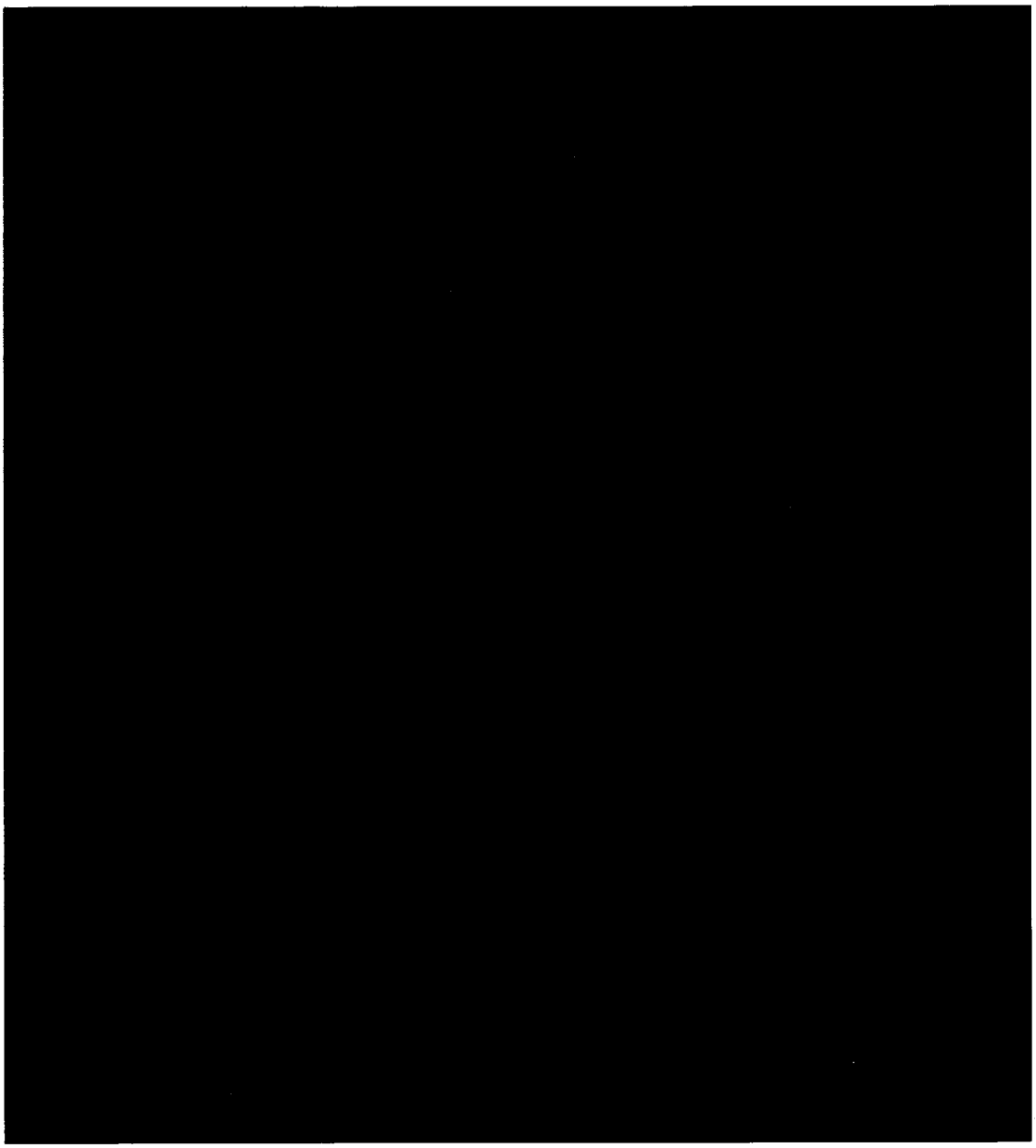
[1] Shinya I, Takeru S, Ataru I, Shinichi N, Takashi F, et al. Developing a Japanese version of the EQ-5D-5L value set. J. Natl. Inst. Public Health.2015;64 (1) : 47-55.

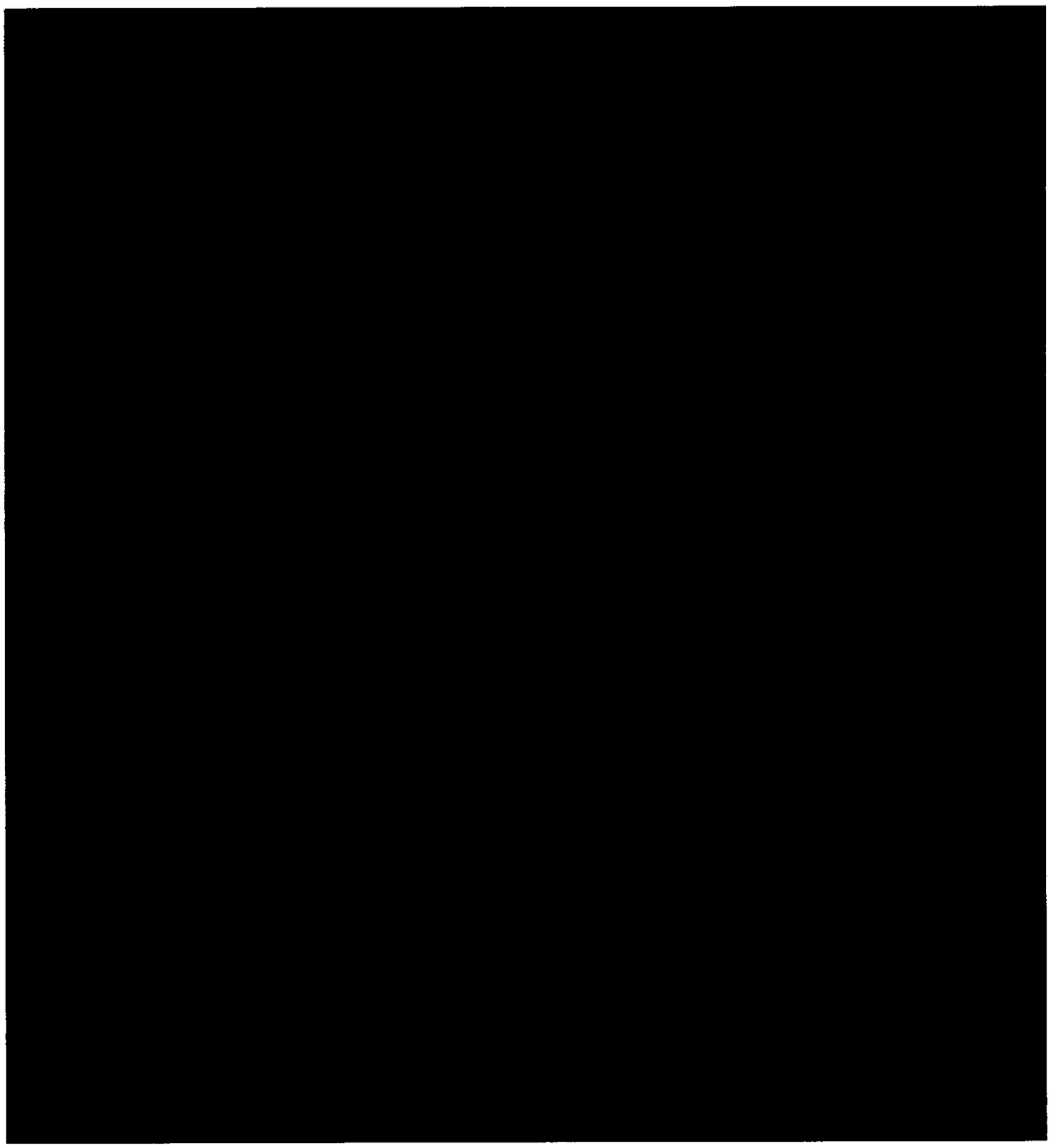
[2] Fukuhara S, Suzukamo Y. Manual of SF-36v2 Japanese version: Institute for Health Outcomes & Process Evaluation research. Kyoto: 2004. Ver3/2011.

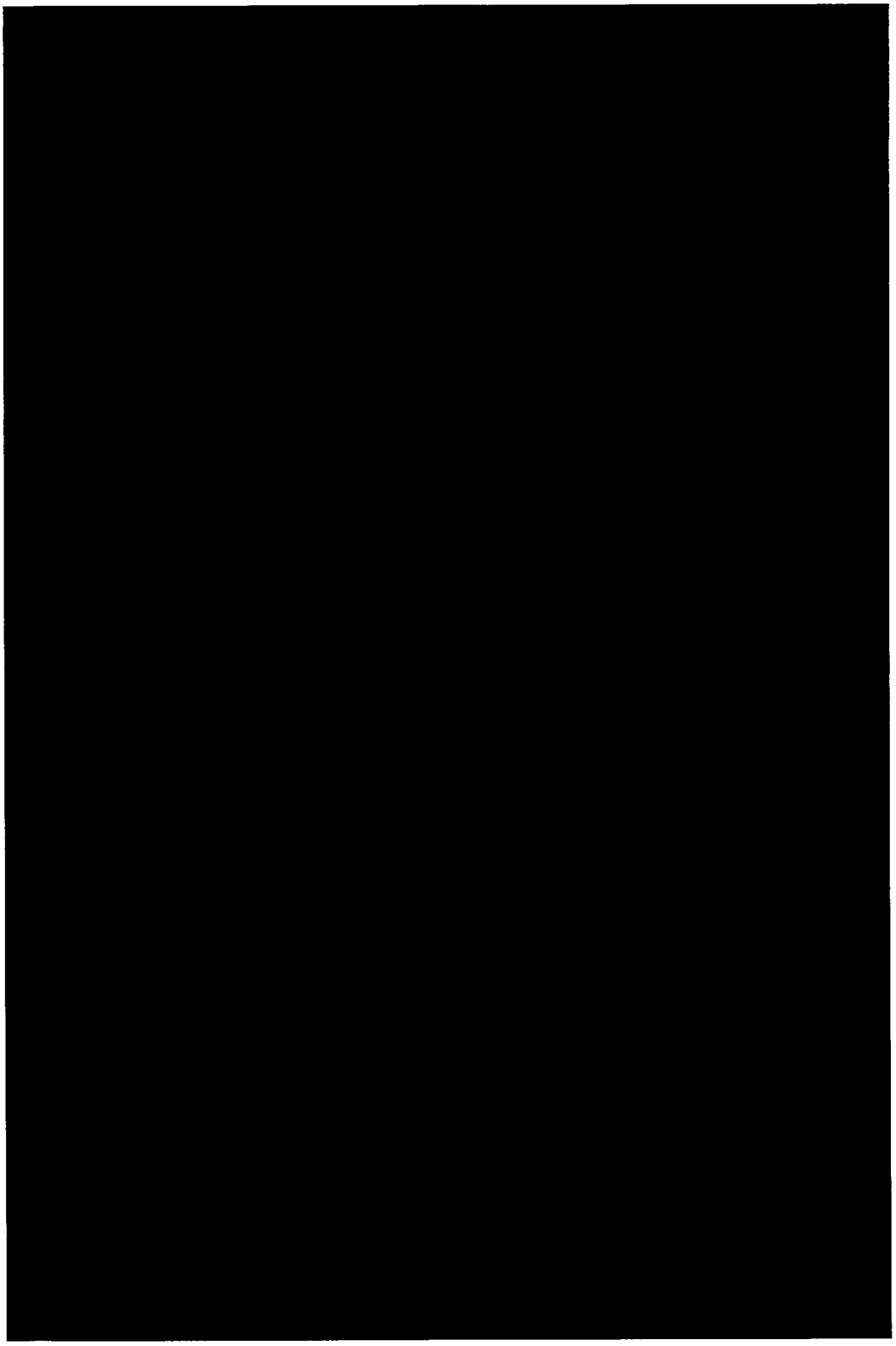
APPENDIX

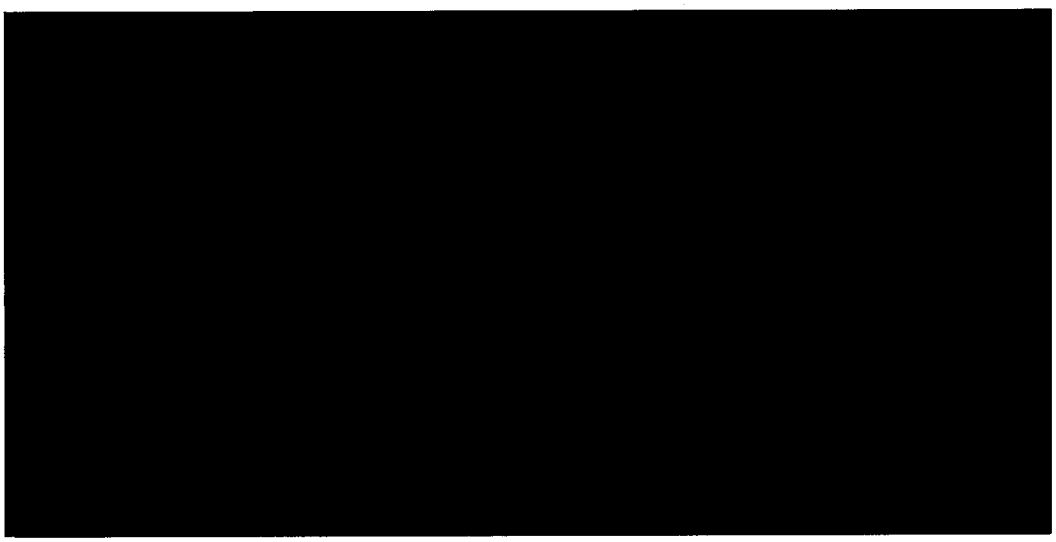
Appendix 1 SF-36v2 Scoring Algorithm[2]











Appendix 2 Criteria for Potentially Clinically Significant Values (PCSV)

Parameter	PCSV	Comments
Clinical Chemistry		
ALT*	>3 and ≤ 5 ULN and baseline ≤ 3 ULN*	Enzyme activity must be expressed in ULN, not in IU/L.
	>5 and ≤ 10 ULN and baseline ≤ 5 ULN	FDA DILI Guidance July 2009.
	>10 and ≤ 20 ULN and baseline ≤ 10 ULN	Each category is calculated independently.
	>20 ULN and baseline ≤ 20 ULN	* At least one level is required, multiple levels are optional for phase 2/3 studies. If it is desirable to get the distribution across the different PCSV levels, additional shift table on ≤3, >3 to ≤5, > 5 to ≤10, >10 to ≤20, and > 20 category for baseline vs. post baseline may be provided

Parameter	PCSV	Comments
AST*	>3 and ≤ 5 ULN and baseline ≤ 3 ULN* >5 and ≤ 10 ULN and baseline ≤ 5 ULN >10 and ≤ 20 ULN and baseline ≤ 10 ULN >20 ULN and baseline ≤ 20 ULN	Enzyme activity must be expressed in ULN, not in IU/L. FDA DILI Guidance July 2009. Each category is calculated independently. * At least one level is required, multiple levels are optional for phase 2/3 studies. If it is desirable to get the distribution across the different PCSV levels, additional shift table on ≤3, >3 to ≤5, > 5 to ≤10, >10 to ≤20, and > 20 category for baseline vs. post baseline may be provided
Alkaline Phosphatase	>1.5 ULN and baseline ≤ 1.5 ULN	Enzyme activity must be expressed in ULN, not in IU/L. FDA DILI Guidance July 2009.

Parameter	PCSV	Comments
Total Bilirubin*	>1.5 and ≤ 2 ULN and baseline ≤ 1.5 ULN* >2 ULN and baseline ≤ 2.0 ULN	Must be expressed in ULN, not in µmol/L or mg/L. Categories are cumulative. FDA DILI Guidance July 2009. * At least one level is required, multiple levels are optional for phase 2/3 studies. If it is desirable to get the distribution of significant level, additional shift table on ≤1.5, >1.5 to ≤2.0 and > 2.0 category for baseline vs. post baseline may be provided
Conjugated Bilirubin	(Direct Bilirubin >35% Total Bilirubin and Total Bilirubin >1.5 ULN) and (Direct Bilirubin <=35% Total Bilirubin or Total Bilirubin <=1.5 ULN) at baseline	Conjugated bilirubin dosed on a case-by-case basis.

Parameter	PCSV	Comments
ALT/AST and Total Bilirubin	(ALT >3 ULN and TBILI>2 ULN) and baseline (ALT <=3 ULN **or** TBILI <=2 ULN) (AST >3 ULN and TBILI>2 ULN) and baseline (AST <=3 ULN or TBILI <=2 ULN) (ALT >3 ULN and TBILI>1.5 ULN) and baseline (ALT <=3 ULN or TBILI <=1.5 ULN) (AST >3 ULN and TBILI>1.5 ULN) and baseline (AST <=3 ULN or TBILI <=1.5 ULN)	FDA DILI Guidance July 2009.
ALT/AST and Total Bilirubin and ALP	(ALT >3 ULN and TBILI>2 ULN and ALP < 2 ULN) and baseline (ALT <=3 ULN **or** TBILI <=2 ULN or ALP >=2 ULN) (AST>3 ULN and TBILI>2 ULN and ALP < 2 ULN) and baseline (AST <=3 ULN or TBILI <=2 ULN or ALP >=2 ULN)	FDA DILI Guidance July 2009.

Parameter	PCSV	Comments
CPK*	>3 and ≤ 10 ULN and baseline ≤ 3ULN* >10 ULN and baseline ≤ 10ULN	FDA Feb 2005. Am J Cardiol April 2006. Categories are cumulative. * At least one level is required, multiple levels are optional for phase 2/3 studies. If it is desirable to get the distribution of significant level, additional shift table on ≤3, >3 to ≤10, and > 10 category for baseline vs. post baseline may be provided
Creatinine	≥150 µmol/L (Adults) and baseline < 150 µmol/L >=30% change from baseline and <100% change from baseline ≥100% change from baseline	Benichou C., 1994. 3 independent criteria
Uric Acid Hyperuricemia Hypouricemia	>408 µmol/L and <=408 µmol/L at baseline <120 µmol/L and >= 120 µmol/L at baseline	Harrison- Principles of internal Medicine 17th Ed., 2008. Two independent criteria
Blood Urea Nitrogen	≥17 mmol/L and <17 mmol/L at baseline	Two independent criteria

Parameter	PCSV	Comments
Chloride		Two independent criteria
Hypochloremia	<80 mmol/L and baseline ≥ 80 mmol/L	
Hyperchloremia	>115 mmol/L and baseline ≤ 115 mmol/L	
Sodium		Two independent criteria
Hyponatremia	≤129 mmol/L and baseline > 129 mmol/L	
Hypernatremia	≥160 mmol/L and baseline <160 mmol/L	
Potassium		FDA Feb 2005.
Hypokalemia	<3 mmol/L and baseline ≥ 3 mmol/L	Two independent criteria
Hyperkalemia	≥5.5 mmol/L and baseline <5.5 mmol/L	
Total Cholesterol	≥7.74 mmol/L and < 7.74 mmol/L at baseline	Threshold for therapeutic intervention.
Triglycerides	≥4.6 mmol/L and < 4.6 mmol/L at baseline	Threshold for therapeutic intervention.

Parameter	PCSV	Comments
Glucose		
Hypoglycaemia	(≤3.9 mmol/L and <LLN) and (>3.9 mmol/L or >=LLN) at baseline	ADA May 2005.
Hyperglycaemia	≥11.1 mmol/L (unfasted); ≥7 mmol/L (fasted) and < 11.1 mmol/L (unfasted); <7 mmol/L (fasted) at baseline	ADA Jan 2008.
HbA1c	>8% and <= 8% at baseline	
Albumin	≤25 g/L and >25 g/L at baseline	
CRP	>2 ULN or >10 mg/L (if ULN not provided) and <=2 ULN or <=10 mg/L (if ULN not provided) at baseline	FDA Sept 2005.
Hematology		
WBC	<3.0 Giga/L and >=3.0 Giga/L at baseline (Non-Black); <2.0 Giga/L and >=2.0 Giga/L at baseline (Black) ≥16.0 Giga/L and < 16 Giga/L at baseline	Increase in WBC: not relevant. To be interpreted only if no differential count available.
Lymphocytes	>4.0 Giga/L and <= 4.0 Giga/L at baseline	

Parameter	PCSV	Comments
Neutrophils	<1.5 Giga/L and >=1.5 Giga/L at baseline (Non-Black); <1.0 Giga/L and >=1.0 Giga/L at baseline (Black)	International Consensus meeting on drug-induced blood cytopenias, 1991. FDA criteria.
Monocytes	>0.7 Giga/L <= 0.7 Giga/L at baseline	
Basophils	>0.1 Giga/L <= 0.1 Giga/L at baseline	
Eosinophils	(>0.5 Giga/L and >ULN) and (<=0.5 Giga/L or <= ULN at baseline)	Harrison- Principles of internal Medicine 17th Ed., 2008.
Hemoglobin	≤115 g/L and > 115 g/L at baseline for male; ≤95 g/L and> 95 g/L at baseline for Female.	Three criteria are independent.
	≥185 g/L and <185 g/L at baseline for Male; ≥165 g/L and < 165 g/L at baseline for Female Decrease from Baseline ≥20 g/L	Criteria based upon decrease from baseline are more relevant than based on absolute value. Other categories for decrease from baseline can be used (≥30 g/L, ≥40 g/L, ≥50 g/L).

Parameter	PCSV	Comments
Hematocrit	≤0.37 v/v and > 0.37 v/v at baseline for Male ; ≤0.32 v/v and > 0.32 v/v at baseline forFemale ≥0.55 v/v and < 0.55 v/v at baseline for Male ; ≥0.5 v/v and < 0.5 v/v at baseline for Female	Two Criteria are independent
RBC	Female <3 Tera/L and baseline ≥3 Tera/L ≥6 Tera/L and baseline < 6 Tera/L Male <4 Tera/L and baseline ≥4 Tera/L ≥7 Tera/L and baseline < 7 Tera/L	Unless specifically required for particular drug development, the analysis is redundant with that of Hb. Otherwise, consider FDA criteria.
Platelets	<100 Giga/L and >=100 Giga/L at baseline ≥700 Giga/L and < 700 Giga/L at baseline	International Consensus meeting on drug-induced blood cytopenias, 1991. Two independent criteria

Parameter	PCSV	Comments
Urinalysis		
pH	≤4.6 and > 4.6 at baseline ≥8 and < 8 at baseline	Two independent criteria
Vital signs		
PR	≤50 bpm and decrease from baseline ≥20 bpm ≥120 bpm and increase from baseline≥20 bpm	To be applied for all positions (including missing) except STANDING.
SBP	≤95 mmHg and decrease from baseline ≥20mmHg ≥160 mmHg and increase from baseline ≥20 mmHg	To be applied for all positions (including missing) except STANDING.
DBP	≤45 mmHg and decrease from baseline ≥10 mmHg ≥110 mmHg and increase from baseline ≥10 mmHg	To be applied for all positions (including missing) except STANDING.

Parameter	PCSV	Comments
Orthostatic Hypotension as per protocol, Orthostatic Hypotension confirmed by repeated assessment as per protocol	Su SBP < 160 mmHg - SBP St – Su ≤ - 20mmHg DBP St – Su ≤ - 10mmHg Su SBP ≥ 160 mmHg – SBP St – Su ≤ - 30 mmHg DBP St – Su ≤ - 15 mmHg Orthostatic PR ≥ 30 beats/min	
Weight	≥5% increase from baseline ≥5% decrease from baseline	FDA Feb 2007.

Parameter	PCSV	Comments
ECG : Ref.: CPMP 1997 guideline. ICH E14 2005		
HR	≤50 bpm and decrease from baseline ≥20 bpm ≥120 bpm and increase from baseline ≥20 bpm	
PR	≥220 ms and increase from baseline ≥20 ms	
QRS	≥120 ms & < 120 ms at baseline	
QTc	Absolute values (ms) >450 ms and baseline <=450 ms >480 ms and baseline <=480 ms >500 ms and <=500 ms at baseline Increase from baseline 30-60 ms Increase from baseline >60 ms	To be applied to QTcF formula.

Appendix 3 Reference Value in Each Laboratory Parameters

Parameter	Reference Value
RBC count (10^12/L)	Male Low < 4.3, High > 5.7 Female Low < 3.8, High > 5
Hemoglobin (g/L)	Male Low < 135, High > 175 Female Low < 115, High > 150
Hematocrit (L/L)	Male Low < 0.397, High > 0.524 Female Low < 0.348, High > 0.45
MCHC (g/dL)	Low < 30.2, High > 35.1
WBC count (10^9/L)	Low < 3.3, High > 9
Platelet count (10^9/L)	Low < 140, High > 340
AST (GOT) (U/L)	Low < 10, High > 40
ALT (GPT) (U/L)	Low < 5, High > 45
ALP (U/L)	Low < 100, High > 325
LDH (U/L)	Low < 120, High > 240
Gamma-GTP (U/L)	Male High > 80 Female High > 30
Total protein (g/L)	Low < 67, High > 83
Albumin (g/L)	Low < 38, High > 52
Total cholesterol (mmol/L)	Low < 3.108, High > 5.672
Total bilirubin (umol/L)	Low < 3.421, High > 20.525
BUN (mmol/L)	Low < 2.856, High > 7.14
Serum creatinine (umol/L)	Male Low < 53.925, High > 91.938 Female Low < 41.549, High > 69.838

Parameter	Reference Value
Na (mmol/L)	Low < 137, High > 147
K (mmol/L)	Low < 3.5, High > 5
Cl (mmol/L)	Low < 98, High > 108
Ca (mmol/L)	Low < 2.1, High > 2.6
HCO3- (mmol/L)	Low < 19.2, High > 27.1
Uric acid (umol/L)	Male Low < 226.043, High > 416.395 Female Low < 148.713, High > 416.395
Triglycerides (mmol/L)	Low < 0.339, High > 1.684
BAP (ug/L)	Male Low < 3.7, High > 20.9 Female Premenopause(30 - 44 years) Low < 2.9, High > 14.5 Postmenopausal(45 - 79 years) Low < 3.8, High > 22.6
Inorganic phosphorus (mmol/L)	Low < 0.808, High > 1.454
CPK (U/L)	Male Low < 40, High > 150 Female Low < 60, High > 270
CRP (mg/L)	High > 3

[redacted]

APPROVAL FORM

Statistical Analysis Plan(For Database at Week 68)

Protocol No.	MT-5547-J01
Protocol Title	A Phase 2/3 (Placebo-Controlled, Double-Blind, Comparative) Study on MT-5547 in Patients with Osteoarthritis Accompanied by Moderate to Severe Pain
Version / Date	1.0 / 19-April-2021

Authors:

Statistics Author	
Print Name:	[redacted]
Position:	Biostatistician
Clinical Pharmacokinetics Author	
Print Name:	[redacted]
Position:	Pharmacokinetics

Approved by:

Statistic Approver	
Print Name:	[redacted]
Position:	Responsible Biostatistician
Signature:	Refer to email
Approval date:	Refer to email
Clinical Pharmacokinetic Approver	
Print Name:	[redacted]
Position:	Responsible Pharmacokinetics
Signature:	Refer to email
Approval date:	Refer to email

[redacted]

[redacted]

[redacted]

[redacted]

APPROVAL FORM

Statistical Analysis Plan(For Database at Week 68)

Protocol No.	MT-5547-J01
Protocol Title	A Phase 2/3 (Placebo-Controlled, Double-Blind, Comparative) Study on MT-5547 in Patients with Osteoarthritis Accompanied by Moderate to Severe Pain
Version / Date	1.0 / 19-April-2021

Authors:

Statistics Author	
Print Name:	[redacted]
Position:	Biostatistician
Clinical Pharmacokinetics Author	
Print Name:	[redacted]
Position:	Pharmacokinetics

Approved by:

Statistic Approver	
Print Name:	[redacted]
Position:	Responsible Biostatistician
Signature:	Refer to email
Approval date:	Refer to email
Clinical Pharmacokinetic Approver	
Print Name:	[redacted]
Position:	Responsible Pharmacokinetics
Signature:	Refer to email
Approval date:	Refer to email

[redacted]

[redacted]