Statistical Analysis Plan

Protocol Number: MT-7117-G02

A Phase 2, Multicenter, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Study to Evaluate Efficacy, Safety, and Tolerability of MT-7117 in Subjects With Diffuse Cutaneous Systemic Sclerosis

Version Number: 1.0
Date: 8 April 2024
NCT number: NCT04440592

Statistical Analysis Plan

Protocol No. MT-7117-G02

A Phase 2, Multicenter, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Study to Evaluate Efficacy, Safety, and Tolerability of MT-7117 in Subjects with Diffuse Cutaneous Systemic Sclerosis

Prepared By:	Mitsubishi Tanabe Pharma Development America
Version:	V1.0
Date:	8APR2024

APPROVAL FORM

Statistical Analysis Plan

Protocol No.	MT-7117-G02
Protocol Title	A Phase 2, Multicenter, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Study to Evaluate Efficacy, Safety, and Tolerability of MT-7117 in Subjects with Diffuse Cutaneous Systemic Sclerosis
Version / Date	V1.0 / 8APR2024

Authors:

Statistics Author	
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
Clinical Pharmacokinetics Author	
[redacted]	[redacted]
[redacted]	[redacted]

Approved by:

Statistic Approver	
[redacted]	
Clinical Pharmacokinetic Approver	
[redacted]	

TABLE OF CONTENTS

ABBREVIATIONS

Abbreviations	Definitions
ACR	American College of Rheumatology
AE	adverse event
AESI	Adverse event of special interest
ALT	alanine transaminase
ALP	alkaline phosphatase
ANCOVA	analysis of covariance
AR	Autoregressive
AST	aspartate transaminase
ATC	anatomical therapeutic chemical
BDR	blinded data review
BMI	body mass index
CI	confidence interval
CRISS	Composite response index in diffuse systemic sclerosis
DP	decimal places
dcSSc	diffuse cutaneous systemic sclerosis
ECG	Electrocardiogram
EPP	Erythropoietic protoporphyria
[redacted]	[redacted]
FAS	full analysis set
FVC	Forced vital capacity
HAQ-DI	Health Assessment Questionnaire Disability Index
iDMC	Independent data monitoring committee
ILD	Interstitial lung disease
ITT	intent-to-treat
MedDRA	medical dictionary for regulatory activities
[redacted]	[redacted]
MMRM	mixed model repeated measures
PD	Pharmacodynamics
PK	Pharmacokinetics
PT	preferred term
%pFVC	Percent predicted forced vital capacity
[redacted]	[redacted]
[redacted]	[redacted]
QTcB	Corrected QT interval using Bazett's formula
QTcF	Corrected QT interval using Fredrecia's formula
SAP	statistical analysis plan
SAE	serious adverse event
SAF	safety population
SD	standard deviation
SOC	system organ class
SSc	Systemic sclerosis
TEAE	treatment emergent adverse event
TESAE	treatment emergent serious adverse events

ULN	upper limit of normal range
VAS	Visual Analog Scale
WHO	World Health Organization

1 INTRODUCTION

This statistical analysis plan (SAP) is based on the final protocol dated 05 Mar 2020 and protocol amendment version 5.0 dated 03 Nov 2021. The plan covers statistical analysis, tabulations, and listings of the study data to assess the efficacy, safety, and PK of MT-7117 compared to placebo.

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

The SAP is prepared by MTPA data science and reviewed by MTPA clinical study team and MTPC data science following GLB-BST-SOP002 ver.7. The statistical analyses and production of the outputs described in the SAP will be conducted and QCed by [redacted] using SAS version 9.4 or higher. The final analyses and outputs will be approved by MTPC/MTPA data science.

This SAP will be finalized prior to database lock.

PGx and PD analysis results will be reported separately from the CSR.

2 STUDY OBJECTIVE AND ENDPOINTS

2.1 Study Objectives

Primary:

- To evaluate the efficacy of MT-7117 treatment in subjects with diffuse cutaneous systemic sclerosis (dcSSc) using the American College of Rheumatology Composite Response Index in Diffuse Systemic Sclerosis (ACR CRISS) at Week 52.

Secondary:

- To evaluate the efficacy of MT-7117 treatment for up to 52 weeks using patient-reported outcomes (PROs) as measured by the Health Assessment Questionnaire Disability Index (HAQ-DI) and Patient Global Assessment.
- To evaluate the efficacy of MT-7117 treatment for up to 52 weeks on pulmonary function as measured by percent predicted forced vital capacity (%pFVC).
- To evaluate the efficacy of MT-7117 treatment for up to 52 weeks using the Physician Global Assessment.
- To evaluate the efficacy of MT-7117 treatment for up to 52 weeks using the modified Rodnan Skin Score (mRSS).
- To evaluate ACR CRISS at Weeks 16, 26, and 39.

- To evaluate ACR CRISS Score improvement proportion up to 52 weeks.

Exploratory:

- [redacted]

Safety:

- To evaluate the safety and tolerability of MT-7117 treatment for up to 56 weeks in subjects with dcSSc.

Pharmacokinetics:

- To determine the pharmacokinetic (PK) profile of MT-7117 in subjects with dcSSc.

2.2 Study Endpoints

2.2.1 Primary Endpoint

- The ACR CRISS composite score (0-1) at Week 52

2.2.2 Secondary Efficacy Endpoints

- Change in HAQ-DI from baseline at Weeks 16, 26, 39, and 52.
- Change in Patient Global Assessment from baseline at Weeks 16, 26, 39, and 52.
- Change in percent predicted forced vital capacity (%pFVC) from baseline at Weeks 16, 26, 39, and 52.
- Change in Physician Global Assessment from baseline at Weeks 16, 26, 39, and 52.
- Change in mRSS from baseline at Weeks 16, 26, 39, and 52.
- ACR CRISS Score at Weeks 16, 26, and 39.
- ACR CRISS Score improvement at Weeks 16, 26, 39 and 52: Proportion of subjects with ≥ 25% improvement in mRSS, HAQ-DI, Patient Global Assessment, Physician Global Assessment, or ≥ 5% improvement in FVC for at least 3 of the 5 ACR CRISS measures.
- ACR CRISS score responder (CRISS ≥ 0.6) at Weeks 16, 26, 39, and 52..

2.2.3 Exploratory Efficacy Endpoints

- [redacted]

2.2.4 Safety Endpoints

- Treatment-emergent adverse events ([TEAEs] including serious adverse events [SAEs], AEs leading to withdrawal, and adverse events of special interest [AESIs]).
- Physical examination.
- Vital signs (blood pressure, pulse rate, respiratory rate, and body temperature).
- Clinical laboratory examinations (hematology, coagulation, biochemistry, and urinalysis), including liver function markers (ALT, AST, gamma glutamyl transpeptidase [GGT], ALP, direct and total bilirubin).
- 12-lead electrocardiogram (ECG).
- Nevi (Melanocytic Lesions) appearance (assessed by a dermatologist or other qualified site staff). Any nevi (Melanocytic Lesions) undergoing change of clinical concern during active treatment will be biopsied for follow-up and evaluated by a central pathology laboratory.

2.2.5 Pharmacokinetics Endpoint

- Assessment of plasma concentrations of MT-7117 measured at scheduled visits.

3 STUDY DESIGN

3.1 Study Design

This is a Phase 2, multicenter, randomized, double-bind, placebo-controlled, parallel-group study to evaluate efficacy, safety, and tolerability of MT-7117 in subjects with dcSSc.

The duration of the study is approximately 60 weeks; screening period up to 4 weeks, double-blind treatment period of 52 weeks, and a safety follow-up period of 4 weeks after last dose.

Following completion of the all screening assessments, subjects will return to the clinical site for start of the double-blind treatment period (Visit 2, Day 1) and baseline assessments will be evaluated and confirmed the eligibility (e.g., safety assessments, ACR CRISS components, and nevi (Melanocytic Lesions) evaluation, if nevi (Melanocytic Lesions) are present.

Eligible subjects will be randomized in a 1:1 ratio to either MT-7117 at starting of [redacted] every day (QD) or matching placebo in a double-blind manner, stratified by autoantibody status of [redacted] performed at screening (positive or negative).

Study drug will be administered once daily orally in the morning with or without food.

During the double-blind treatment period, subjects will return to the clinical site and assessments will be collected as described in protocol.

The decision about dose reduction and tolerability should be made according to the pre-specified criteria and the Investigator's clinical judgment.

If a subject is experiencing intolerance to study drug, they should first be conservatively managed using standard of care and maintain the starting dose of study drug (active or placebo).

If the subject continues to present significant intolerable AEs at the following scheduled or unscheduled visit and the Investigator deems it necessary, the daily dose of study drug can be reduced from [redacted] QD to [redacted] QD while tolerability is managed with standard of care.

If the subject continues to present significant intolerable AEs at the following scheduled or unscheduled visit and the Investigator deems it necessary, the dose of study drug can be reduced from [redacted] QD to [redacted] QD.

If the subject continues to present significant intolerable AEs at the decreased dose of [redacted] QD and the Investigator deems it necessary, the subject may discontinue study treatment.

All dose reductions during the Treatment Period will be done in [redacted] decrements. Dose reductions will be conducted via the Interactive Web-based Response System (IWRS).

Subjects who do not tolerate a minimum dose of [redacted] day will be withdrawn from the study.

Subjects with worsening scleroderma such as skin thickening or other scenarios will be allowed to use rescue therapy starting at or after Week 26, and those with confirmation

of predicted FVC decline (after 2 separate confirmations within 4 weeks via an unscheduled visit) will be allowed to use rescue therapy starting at or after Week 16.

End of treatment (EOT) or early termination (ET) assessments will be performed for all subjects, who will undergo efficacy and safety evaluations. Efficacy evaluations will include change in the composite score of ACR CRISS Components, PGIC, PGIS, UCLA SCTC GIT 2.0 questionnaire, PROMIS-29 test questionnaire, and certain biomarkers.

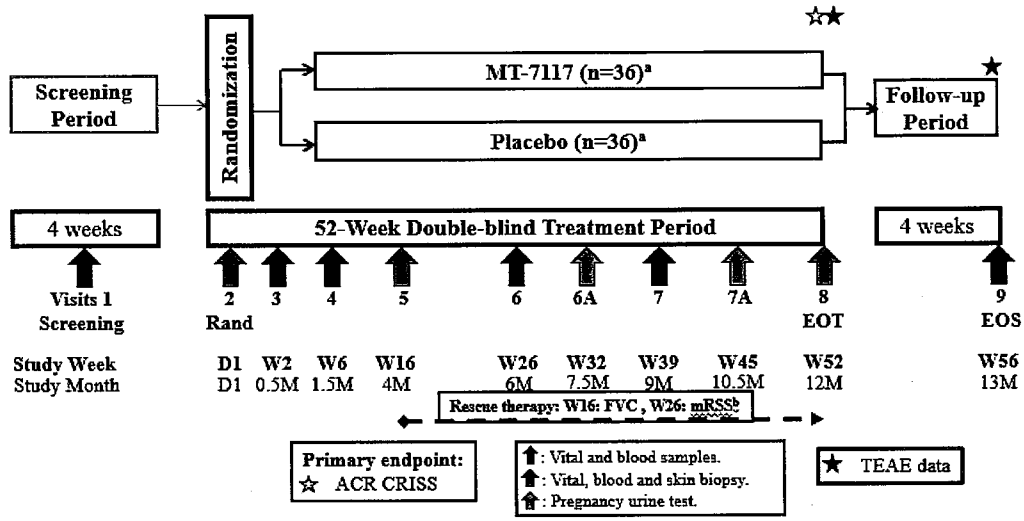
Subjects who discontinue study treatment before EOT (Week 52) should be encouraged to continue in the study and complete all the required study assessments through Week 52.

If a subject decides to permanently withdraw consent from the study, every attempt should be made to have the subject complete the ET visit and complete the study assessments listed at Week 52.

In the event a subject does not return to the clinical site for the ET and/or safety follow-up visit, the Investigator must make every effort to contact the subject by telephone to for the collection and source documentation of safety information such as AEs, dosing compliance, and review of concomitant medications and date of last dose of study drug. Return of unused medication will be performed by courier where allowed.

A safety follow-up visit will occur at Week 56 (Visit 9) for subjects who complete the double-blind treatment period (Week 52) and 4 weeks after the last dose of study drug for subjects who ET from the study.

Further details can be found in the Study Schema below.



D: Day; M: Month; W: Week, EOT: End Of Treatment, EOS: End Of Study, TEAE: Treatment Emergent Adverse Event
FVC: forced vital capacity, mRSS: modified Rodnan Skin Score, ACR CRISS: American College of Rheumatology Composite Response Index in Diffuse Systemic Sclerosis

a) Starting daily dose will be MT-7117 [redacted] QD or matching placebo, with possibility to reduce the dose in a stepwise manner from [redacted] QD to [redacted] QD and possibly from [redacted] QD to [redacted] QD to manage subjects' tolerability to study drug.
b) At the discretion of the Investigator, rescue therapy may start as early as Week 16 with confirmation of predicted FVC decline, and as early as Week 26 with worsening scleroderma such as skin thickening or other scenarios.

Figure 1 Study Design Schema

3.2 Schedule of Study Procedures

Study assessments and corresponding event schedules are summarized in the time and events schedule. The schedule of assessment is shown in [redacted]

Table 1: Schedule of Activities

	Screening riod	Double-blind Treatment Period									Safety Follow-up Period
Visit Number	**1**	**2**	**3**	**4**	**5**	**6**	**6A Telephone Visit**	**7**	**7A Telephone Visit**	**8 (EOT/ET)**	**9[a] (EOS)**
Study Week	**-4**	**1**	**2**	**6**	**16**	**26**	**32**	**39**	**45**	**52**	**56**
Study Day ± Window	**–28 to -1**	**1 (Baseline)**	**15 (± 3 days)**	**43 (± 5 days)**	**113 (± 7 days)**	**183 (± 7 days)**	**225 (± 7 days)**	**274 (± 7 days)**	**316 (± 7 days)**	**365 (± 7 days)**	**393 (± 7 days)**
Informed consent	X										
Inclusion/exclusion criteria	X	X									
Prior medications	X										
Demographics	X										
Medical history	X										
Randomization		X									
Dispensing of study drug		X	X	X	X	X		X			
Drug accountability			X	X	X	X		X		X	
Body weight	X	X	X	X	X	X		X		X	
Height	X										

	Screening riod	Double-blind Treatment Period									Safety Follow-up Period
Visit Number	**1**	**2**	**3**	**4**	**5**	**6**	**6A Telephone Visit**	**7**	**7A Telephone Visit**	**8 (EOT/ET)**	**9[a] (EOS)**
Study Week	**-4**	**1**	**2**	**6**	**16**	**26**	**32**	**39**	**45**	**52**	**56**
Study Day ± Window	**–28 to -1**	**1 (Baseline)**	**15 (± 3 days)**	**43 (± 5 days)**	**113 (± 7 days)**	**183 (± 7 days)**	**225 (± 7 days)**	**274 (± 7 days)**	**316 (± 7 days)**	**365 (± 7 days)**	**393 (± 7 days)**
Physical examination[b]	X	X	X	X	X	X		X		X	X
Vital signs[c]	X	X	X	X	X	X		X		X	X
12-lead ECG	X					X				X	
SSc-specific autoantibody test[d]	X										
[redacted]	X									X	
[redacted]	X					X		X		X	
[redacted]		X				X		X		X	
[redacted]	X			X	X	X		X		X	
Pregnancy test[e]	X	X	X	X	X	X	X	X	X	X	X

	Screening riod	Double-blind Treatment Period									Safety Follow-up Period
Visit Number	**1**	**2**	**3**	**4**	**5**	**6**	**6A Telephone Visit**	**7**	**7A Telephone Visit**	**8 (EOT/ET)**	**9[a] (EOS)**
Study Week	**-4**	**1**	**2**	**6**	**16**	**26**	**32**	**39**	**45**	**52**	**56**
Study Day ± Window	**–28 to -1**	**1 (Baseline)**	**15 (± 3 days)**	**43 (± 5 days)**	**113 (± 7 days)**	**183 (± 7 days)**	**225 (± 7 days)**	**274 (± 7 days)**	**316 (± 7 days)**	**365 (± 7 days)**	**393 (± 7 days)**
Hematology/coagulation, biochemistry, and urinalysis[f]	X	X	X	X	X	X		X		X	X
ESR[g]	X										
PK sampling (blood)[h]		X	X	X	X	X		X		X	
Blood sampling for [redacted]		X									
[redacted]		X			X					X	
[redacted]		X			X	X				X	

	Screening Period	Double-blind Treatment Period									Safety Follow-up Period
Visit Number	**1**	**2**	**3**	**4**	**5**	**6**	**6A Telephone Visit**	**7**	**7A Telephone Visit**	**8 (EOT/ET)**	**9[a] (EOS)**
Study Week	**-4**	**1**	**2**	**6**	**16**	**26**	**32**	**39**	**45**	**52**	**56**
Study Day ± Window	**–28 to -1**	**1 (Baseline)**	**15 (± 3 days)**	**43 (± 5 days)**	**113 (± 7 days)**	**183 (± 7 days)**	**225 (± 7 days)**	**274 (± 7 days)**	**316 (± 7 days)**	**365 (± 7 days)**	**393 (± 7 days)**
[redacted]		X			X	X				X	
[redacted]		X			X	X				X	
[redacted]		X			X	X				X	
ACR CRISS					X	X		X		X	
[redacted]					X	X		X		X	
mRSS	X	X			X	X		X		X	
%pFVC	X	X			X	X		X		X	
SHAQ[m]		X			X	X		X		X	
PtGA		X			X	X		X		X	
PhGA		X			X	X		X		X	
[redacted]				X	X	X		X		X	
[redacted]		X		X	X	X		X		X	

	Screening riod	Double-blind Treatment Period								Safety Follow-up Period
Visit Number	**1**	**2**	**3**	**4**	**5**	**6**	**6A Telephone Visit**	**7**	**7A Telephone Visit**	**8 (EOT/ET)**	**9[a] (EOS)**
Study Week	**-4**	**1**	**2**	**6**	**16**	**26**	**32**	**39**	**45**	**52**	**56**
Study Day ± Window	**–28 to -1**	**1 (Baseline)**	**15 (± 3 days)**	**43 (± 5 days)**	**113 (± 7 days)**	**183 (± 7 days)**	**225 (± 7 days)**	**274 (± 7 days)**	**316 (± 7 days)**	**365 (± 7 days)**	**393 (± 7 days)**
[redacted]		X				X				X	
[redacted]		X				X				X	
[redacted]		X				X				X	
[redacted]		X	X	X	X	X		X		X	X
Nevi (Melanocytic Lesions) evaluation[o]	X		X	X		X				X	X
Subject Question for study drug[p]										X	
Concomitant medications[q]	X	X	X	X	X	X		X		X	X
Record Adverse events	X	X	X	X	X	X		X		X	X

Abbreviations: ACR CRISS = American College of Rheumatology Composite Response Index for Systemic Sclerosis; ECG = electrocardiogram; EOS = End of Study; EOT = End of treatment; [redacted] SHAQ= Scleroderma Health Assessment Questionnaire ; HIV = human immunodeficiency virus; mRSS = modified Rodnan Skin Score; [redacted] %pFVC = percent predicted forced vital capacity; [redacted]; PK = pharmacokinetic; [redacted]; MC1R= melanocortin-1 receptor; SNPs= single-nucleotide polymorphisms, PhGA=Physician Global Assessment; PtGA=Patient Global Assessment; ET=Early Termination; AE=Adverse Event; eCRF= electronic Case Report Form; HAQ-DI= Health Assessment Questionnaire Disability Index; VAS= Visual Analog Scale

a. All subjects (completers and those who ET) will return to the clinical site 4 weeks after their last dose for a safety follow-up visit. For subjects who ET and will not revisit the clinical site, the site will perform a scheduled phone call for the collection of safety assessments (e.g., AEs, concomitant medication, and date of last dose of study drug).
b. A complete physical examination will be performed at Visit 1 and an abbreviated physical examination will be performed at all other specified time points.
c. Vital signs will include blood pressure, pulse rate, respiratory rate, and body temperature (e.g., oral, axillary or tympanic body temperature).
d. [redacted].
e. For female subjects of child-bearing potential, a serum pregnancy test will be performed at Visit 1, a urine pregnancy test (on-site) will be performed at all other site visits, and a urine dipstick pregnancy test at home or other home testing arrangements will be performed at Telephone Visits (Visit 6A and 7A).
f. Refer to [redacted] for laboratory evaluations.
g. ESR will be analyzed locally at the site level and at Screening only.
h. PK blood samples for MT-7117 will be collected at Visit 2 predose, Visits 3, 4, 5, 6, 7, and Visit 8 at any time during each visit. Date and time of study drug dose and date and time of PK sample collection will be recorded.
i. Blood sampling will be collected for optional PGx analysis for those subjects who have provided informed consent.
j. [redacted]

l. [redacted]s.

m. SHAQ is composed of HAQ-DI and VAS.

n. [redacted]

o. Nevi (Melanocytic Lesions) evaluation will be performed by a local dermatologist or qualified site staff. Baseline nevi (Melanocytic Lesions) evaluation will be performed at any time during the screening period before randomization at Visit 2 (Day 1). The nevi (Melanocytic Lesions) evaluation at Visit 9 is to assess for the reversibility if any suspicious nevi (Melanocytic Lesions) changes that were observed during treatment as per the Investigator's (and/or dermatologist's or other qualified site staff) judgment. Any follow-up will be recorded in the eCRF. Nevi (Melanocytic Lesions) assessment will be described in a separate document.

p. Subjects will be asked whether they believe they received active or placebo treatment.

q. Concomitant medications at screening, baseline and used during study treatment will be reviewed, recorded and discussed with the Sponsor and medical Monitor (as needed).

[redacted].

3.2.1 Screening Phase and Rescreening

Screening assessments will be performed up to 28 days prior to Day 1 of the double-blind treatment period. All screening evaluations must be completed and reviewed to confirm that potential subjects meet all eligibility criteria.

At the end of the screening period, if a subject has not met all eligibility criteria at the end of the screening period, the subject will be registered as a screen failure. Screen failures may be eligible for rescreening 1-time following consultation with the Sponsor and Medical Monitor.

Note: If the rescreening occurs more than 30 days after the first screening and signing of the original ICF, all screening procedures, including ICF, must be repeated. If within 30 days, repeat assessments should be discussed with the Medical Monitor.

If a subject does not meet eligibility criteria for laboratory and pulmonary function tests (PFTs), the subject may undergo repeat laboratory and PFTs up to 2 additional times during the 4-week screening period. Repeat of laboratory or PFTs is not considered rescreening.

3.2.2 Double-blind Treatment Period

Subjects who successfully complete the screening period will return to the clinical site on Day 1 of the double-blind treatment period to reconfirm his/her eligibility by reviewing the inclusion and exclusion criteria and evaluate other baseline criteria prior to randomization. To be randomized into the double-blind treatment period, the subject must meet the required criteria at screening and at baseline. Eligible subjects will be randomized, and study drug administration will begin on Day 1 (Visit 2) after confirmation of all baseline criteria.

Study visits will occur at the study site per the Schedule of Assessments (Table 1).

3.2.3 Dose Reduction Visit

If a subject experience any intolerable AEs, the dose can be reduced. Subjects with intolerable AEs, will have to come back to the clinical site for an unscheduled visit where the following procedures will be performed:

- Physical examination including vital signs.
- Adverse events and concomitant medications, if any, will be assessed since from the previous visit.
- Local urine pregnancy test (if applicable).
- Safety blood and urine samples will be collected and submitted to the central laboratory for analysis.

- Study drug will be collected, IWRS will allocate new study drug, and treatment compliance will be reviewed.

If a dose reduction is deemed necessary by the Principal Investigator, the IWRS will be used for assignment of the new [redacted] packs to be provided to the subject for dose reduction and study drug dispensation. The date of dose reduction visit and the reason will be recorded in the eCRF.

3.2.4 Dose Reduction Criteria

Subjects will be informed regarding potential side effects. If the side effects experienced by subjects are not manageable through nonpharmacologic and pharmacologic treatment (using standard of care), dose reduction (from [redacted] QD to [redacted] QD, or from [redacted] QD to [redacted] QD) should be considered to manage the incidence and/or intensity of the subject-reported AEs (including but not limited to vomiting, diarrhea, headache, hyperpigmentation).

For safety reasons, the Investigator may also decide to implement a dose reduction for a subject. The following are criteria for Investigator initiated dose reduction:

- AEs which are classified as serious and/or moderate/severe (in severity), which did not improve with clinical management according to standard of care per Investigator's judgement, AND
- AEs assessed as having reasonable possibility of relationship to MT-7117 per Investigator's judgement.

Examples of severe AEs include AE which limit self-care activities of daily living (independently eat, dress, walk/transfer from 1 position to another, bathe, toilet, maintaining bowel/bladder continence).

Note: Dose interruption can be considered after the discussion with the Sponsor Medical Monitor in the case, such as

1) in subjects with suspicious lesions or change in nevi (Melanocytic Lesions) that had their dose temporarily interrupted can resume study treatment at the same dose they were interrupted at after the evaluations are complete and nevi (Melanocytic Lesions) and/or other lesions are determined to be benign; or

2) in a subject with an initial positive urine pregnancy test, awaiting a confirmatory test. If the confirmatory test is definitively negative, dosing may resume at the same dose they were interrupted; or

3) in a subject with any laboratory or clinical abnormality such as meeting the criteria of AESI, awaiting a confirmatory test and/or a judgement whether it is clinically significant. Resumption of dosing at the same dose they were interrupted will be considered based on the discussion with the Sponsor Medical Monitor and Investigator and should be closely followed up.

At any time during the study, subjects may be withdrawn from study treatment if continuing study treatment would be detrimental to the subject's safety in the opinion of the Investigator.

3.3 Sample Size and Power Considerations

The planned sample size of 72 is expected to provide power for the comparison between MT-7117 treatment group and placebo group for ACR CRISS score at Week 52. The calculation of sample size assumes a 2-sided alpha level of 0.05 and uses EAST (Version 6.5) for Wilcoxon Mann Whitney test. The randomized 72 subjects (36 subjects for MT-7117 and 36 subjects for placebo) with a 1:1 allocation ratio would provide 80% power to detect treatment difference of 0.35 (35% improvement on MT-7117 treatment group compared to placebo group) in ACR CRISS score at Week 52 with an associated standard deviation (SD) of 0.5.

4 PLANNED ANALYSIS

4.1 Interim analysis and Data Monitoring Committee

No formal interim analysis of efficacy/safety data will be conducted prior to collection of all data necessary for the statistical analysis.

An external independent Data Monitoring Committee (iDMC) will be established to perform regular review of the safety data to ensure the ongoing safety and risk-benefit of participating subjects until the last subject completes the double-blind treatment period. The frequency of data review will be described in the iDMC Charter. The safety monitoring analyses required for the iDMC's review will be performed by Cytel Inc. which is a Contract Research Organization. These interim analyses are described in the iDMC Statistical Analysis Plan (SAP).

4.2 Final Analysis

Final Analysis will take place after the study complete and data base is locked. The analyses defined in this document will be performed.

5 ANALYSIS POPULATIONS

5.1 Randomized (RAND) Population

The Randomized (RAND) population includes all randomized subjects.

5.2 Intent-to-Treatment (ITT) Population

The Intent to treat (ITT) population includes all randomized subjects who receive at least 1 dose of study drug. The ITT population will be used for all efficacy analyses. The subject randomized treatment will be used for efficacy analyses.

5.3 Per-protocol (PP) population

The per-protocol population includes all ITT subjects who do not have to be excluded due to relevant protocol deviations (to be discussed in the blinded data review meeting) and who complete Week 52 (the end of double blinded treatment period). The PP will be used to perform supportive analysis for the only primary endpoint (CRISS).

5.4 Safety (SAF) Population

The safety analysis population is defined as all randomized subjects who received at least 1 dose of study drug. The safety population will be used for all safety analyses. The subject actual treatment received will be used for safety analyses. For subjects who took more than one treatment, the highest dose level will be used for safety analyses.

5.5 Pharmacokinetic (PK) population

The PK population includes all randomized subjects who receive at least 1 dose of study drug and who have at least 1 postdose value of plasma concentration to be included in the PK analysis without important protocol deviations which may affect the PK of study drug. PK analysis will only be performed on samples with active study drug.

6 STATISTICAL CONSIDERATIONS

6.1 Descriptive Statistics

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 (e.g. on some occasions, percentages may be calculated out of the total number of subjects with available data at a particular visit and/or time point).
Unless otherwise specified, all data will be summarized by analysis visits and treatment group.

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

Data will be handled as follows except for the results of blinded data review meeting (BDRM).

7.1 Analysis Variable Definitions

7.1.1 Study Subjects

7.1.1.1 Protocol Deviation

Protocol deviations will be documented, reviewed and determined in the BDRM. Protocol deviations potentially influencing the evaluation of the primary endpoint will be defined as major deviations. The major protocol deviations will be selected in this meeting.

At least the following will be included as major protocol deviations:

- I/E criteria violation with significant impact on the primary endpoint
- Took prohibited medication during treatment

The BDRM meeting minutes will serve as the main document for protocol deviations.

7.1.1.2 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) SSc Disease Duration

SSc Disease Duration (days) = the date of informed consent – the date of diagnosis (the first non-Raynaud's phenomenon manifestation) + 1

If the date of diagnosis is missing, the first day of the month will be used.

If the the date and month of diagnosis are missing, then the first day of the first month (January) will be used.

7.1.1.3 Medical History

Medical history will be coded according to the MedDRA version 25.0 or higher.

7.1.1.4 Prior or Concomitant Medication

Medications will be coded according to the WHO Drug Dictionary (WHO-DD) B3 MAR 2019 version or higher.

(1) Prior Medication

Any prior medication, including prescription and over-the-counter medications, taken within 30 days before screening and all SSc-related medications used since diagnosis will be recorded on the eCRF and source documents. The following information will be collected: name of medication, dose, duration, and reason for use.

(2) Concomitant Medication

Concomitant medication(s) is defined as any medication, other than study drug, which is taken during the treatment period, including prescription, over-the-counter medications, herbals, dietary supplements, and recreational drugs. All concomitant medication(s) taken while the subject is participating in the study will be recorded.

Concomitant medication(s) will be provided only if deemed necessary by the Investigator or the subject's personal physician.

Analysis rules to determine prior medications and concomitant medications

Medications with a stop date before the first date of study drug dosing will be defined as prior medications. Medications with start date or stop date on or after the first date of study drug dosing or ongoing up to week 56 will be defined as concomitant medications.

If the medication start date is incomplete, then it will be imputed as follows for the purpose of determining concomitant use:

- If the start date is completely missing, the start date will be equal to the first dose date.

However, if the stop date is not missing and is before the first dose date, then the stop date will be used instead of the start date. It will not be considered as concomitant medications.

- If the start date is missing, the first day of the month will be used.
- If the start date and month are missing, then the first day of the first month (January) will be used.

If the medication stop date is partial, then it will be imputed as follows for the purpose of determining concomitant use:

- If the stop date is completely missing and the medication is not ongoing, the stop date will be equal to the last dose date or date of completion/withdrawal, whichever is the latest.
- If the stop date is missing, the last day of the month will be used.
- If the stop date and month are missing, then the last day of the last month (December) will be used.

7.1.1.5 Treatment Duration and Compliance

(1) Treatment Duration

Study medication exposure in days will be calculated for each subject using the following:

Treatment duration (days)
=Date of last study drug up-to week 52 – date of first study drug + 1

If the date of the first study drug or the date of the last study drug is missing, the first day of the month will be used.
If the date and month of the first study drug or the those of the last study drug are missing, then the first day of the first month (January) will be used.
Interruptions, compliance, and dose changes are not taken into account for duration of exposure.

(2) Treatment Compliance

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

$$Treatment\ compliance(\%) = \frac{Treatment\ duration\ (days) - Days\ of\ missed\ doses + Days\ of\ overdose}{Treatment\ duration\ (days)} \times 100\%$$

Days of missed doses : Counted days only if the patient does not take at least 1 of the daily dose

of 3 investigational product tablets

7.1.1.6 Interruption duration

Study drug interruption duration will be calculated for each subject who experienced study drug interruption using the following:

Interruption duration (days) = End Date of Study Drug Interruption – Start Date of Study Drug Interruption + 1

If the Start Date of Study Drug Interruption or the End Date of Study Drug Interruption is missing, the first day of the month will be used.
If the date and month of the Start Date of Study Drug Interruption or the those of the End Date of Study Drug Interruption are missing, then the first day of the first month (January) will be used.
Drug Interruption : if all 3 investigational product tablets per day are not taken.

7.1.1.7 Duration of SSc-Interstitial lung disease (SSc-ILD)

Duration of SSc-ILD will be calculated for each subject using the following:
Duration of SSc-ILD (days) = The date of informed consent - The date of diagnosis of SSc-ILD + 1

7.1.2 Efficacy assessments

7.1.2.1 Primary Efficacy Endpoint

The primary efficacy endpoint is the ACR CRISS score which is a composite endpoint assessed in a 2-step process that calculates the probability of improvement for each subject ranging from 0.0 (no improvement) to 1.0 (marked improvement).

Step 1: Subjects who develop new or worsening of cardiopulmonary and/or renal involvement due to systemic sclerosis are considered as not improved (irrespective of improvement in other core items) and assigned a probability of improving equal to 0.0. Specifically, if a subject develops any of the following

- New scleroderma renal crisis (See protocol for definition)
- Decline in %pFVC ≥ 15% (relative), confirmed by another %pFVC within a month, high resolution computer tomography (HRCT) to confirm ILD (if

previous high-resolution computer tomography of chest did not show ILD) and %pFVC below 80% predicted*

- New onset of left ventricular failure (defined as left ventricular ejection fraction ≤ 45%) requiring treatment*
- New onset of pulmonary arterial hypertension (PAH) on right heart catheterization requiring treatment.* PAH is defined as mean pulmonary artery pressure ≥ 20 mm Hg at rest and an end-expiratory pulmonary artery wedge pressure ≤ 15 mm Hg and a pulmonary vascular resistance > 3 Wood units

*= Attributable to systemic sclerosis.

The step 1 is up to medical judgement and the data in the CRF will be used directly for deriving the endpoint. If the data in the CRF is "NOT done" or "missing" then the ACR CRISS score will be missing.

Step 2: For the remaining subjects, Step 2 involves computing the predicted probability of improving for each subject using the following equation (equation to derive predicted probabilities from a logistic regression model):

$$\frac{exp[-5.54 - 0.81 * \Delta_{MRSS} + 0.21 * \Delta_{FVC\%} - 0.40 * \Delta_{Pt-glob} - 0.44 * \Delta_{MD-glob} - 3.41 * \Delta_{HAQ-DI}]}{1 + exp[-5.54 - 0.81 * \Delta_{MRSS} + 0.21 * \Delta_{FVC\%} - 0.40 * \Delta_{Pt-glob} - 0.44 * \Delta_{MD-glob} - 3.41 * \Delta_{HAQ-DI}]}$$

where ΔMRSS indicates the change in MRSS from baseline to each visit, ΔFVC denotes the change in FVC% predicted from baseline to each visit , $\Delta_{Pt\text{-}glob}$ indicates the change in subject global assessment, $\Delta_{MD\text{-}glob}$ denotes the change in physician global assessment, and ΔHAQ-DI is the change in HAQ-DI. All changes are absolute change ($Time_2$–$Time_{baseline}$).

At the Investigator's discretion, if a subject's medical conditions are worsening such as %pFVC decline, an HRCT scan may be performed as a part of standard of care based on a local guidance to confirm worsening.

7.1.2.2 Secondary Efficacy Endpoints

- ACR CRISS Responder
  An ACR CRISS score of 0.60 or greater indicates that a subject improved on treatment and a score of less than 0.6 suggests that a subject has not improved.
- HAQ-DI
  The HAQ-DI is a self-administered instrument that measures physical disability in 8 different domains of function: dressing/grooming, arising, eating, walking, hygiene,

reach, grip, and common activities. Subjects are asked about 2 or 3 activities associated with each category and choose whether they are able to perform the activity 'without any difficulty', 'with some difficulty', 'with much difficulty', or 'unable to do'. An additional section asks subjects to indicate if they used any aids or devices or needed assistance to perform daily activities.

A negative change from baseline indicates improvement. The Minimal Clinically Important Difference (MCID) for improvement ranges from 0.14-0.22 (Khanna et al. 2006; Pope 2011)

HAQ-DI is calculated based on the following points.

- There are four possible response for each component:
  Without ANY difficulty = 0
  With SOME difficulty = 1
  With MUCH difficulty = 2
  UNABLE to do = 3
- A category score is determined from the highest score of the components in that category (except when aids and devices are taken into account; see below). For example, if there are three components in a category and the responses were 1, 2 and 0 to the components in the category, then the score for the category would be 2.
- If a category consists of only two questions
- and has one missing response, then the non-missing response will be used as the value for the category. If a category consists of three questions and has one missing response, then the category score will be the higher of the two responses. If a category consists of three questions and has two missing responses, then the category will be considered missing.
- To calculate the HAQ-DI, the patient must have a category score for at least six of the eight categories. The HAQ-DI is the sum of the category scores, divided by the number of categories that have a score.
- The HAQ-DI takes into account the patient's use of aids or devices in the scoring for a category. For each of the eight categories, there is an aids or devices companion variable(s) that is used to record any assistance the patient uses. Where aids or devices are indicated by a patient for a category or help from another person is required for a category, if the maximum score for the category is <2, the category score is increased to 2 to reflect the use of an aid or device, or help. If the maximum value is ≥ 2, the score is not modified. In the event that a category score is missing

but a corresponding aid or device is listed, then the score for that category will reflect the use of the aid or device (i.e., it will be scored as 2).

- Where "other" is ticked for use of aids or devices, the use of the aid or device will not be assigned to a category and will therefore not be reflected in the category scoring.
- If there are fewer than 6 categories with responses, an HAQ-DI cannot be calculated.
- The assignment of aids and devices to categories is done as follows:

Aid or device	Associated with category
Cane, walker, crutches, wheelchair	Walking
Devices used for dressing	Dressing and grooming
Built up or special utensils	Eating
Special or built up chair	Arising
Raised toilet seat	Hygiene
Bath tub seat	Hygiene
Jar opener (for jars previously opened)	Grip
Bath tub bar	Hygiene
Long-handled appliances for reach	Reach
Long-handled appliances for bathroom	Hygiene

- No aids or devices are assigned to the category 'Activities'.
- Help from another person for errands and chores is assigned to the category 'Activities'.

- Patient Global Assessment

  The Patient Global Assessment is used to assess the subject's rating of their overall disease activity. Subjects rate their perceived health on an 11-point scale from 0 (excellent) to 10 (extremely poor).

  A negative change from baseline indicates improvement. The MCID has been reported by Sekhon and Pope 2010 to be -6.70 mm on Global VAS scale 100 mm. Therefore, -0.67 point will be defined as MCID in this assessment.

- Observed FVC and %pFVC

  Pulmonary function tests (PFTs) are a group of tests that measure how well the lungs are performing. The test includes how well the subject is able to breathe and how effective the lungs are able to bring oxygen to the rest of the body. PFT evaluators need to be trained and appropriately qualified to perform both Spirometry and DLco tests.

FVC measurements will be conducted in the clinical site at around the same time of day where possible with the subject in a sitting upright position. Subjects should undertake at least 3 attempts to generate acceptable and reproducible FVC data per ATS/ERS guidelines. The highest acceptable value will be selected following quality review by a centralized over read specialist in agreement with the site Investigator. Single-breath DLco will be performed at selected sites/countries in accordance with recently published ATS/ERS (the American Thoracic Society/European Respiratory Society) guidelines using equipment and testing techniques that meet ATS/ERS requirement. At least 2 acceptable tests that meet repeatability criteria will be performed. The best value (mean from acceptable values) will be selected by a centralized over read specialist in agreement with the site Investigator.

The acceptability of the FVC and DLco data will be determined by a centralized over read specialist. The centralized over read specialist will state whether a session has been accepted (passed QA) or rejected (failed QA). As the lungs are involved in most of dcSSc patients, there is some difficulties for dcSSc patient to repeat PFTs until they have acceptable values to meet the ATS/ERS criteria. Even if there was no accepted (passed QA) value at each visit, we accept that value as long as there is no issue with the flow loop legibility, since, for some dcSSc subjects, this is their best performance. Records of such manoeuvres should be retained since they may contain useful information. The QA passed values will be used at each visit unless there is no acceptable values at a visit.

PFTs data will include observed FVC, %predcted FVC, observed DLco, and %predicted DLco such as below.

1) Observed FVC is the best FVC and the highest FVC value from accepted values.

2) FEV is the forced expiratory volume in one seconde, which is the volume delivered in the first second of the best FVC manoeuvre..

3) %pFVC is the best FVC as a percentage of predicted FVC (Best FVC/pFVC x 100%).

4) Observed DLco is the best DLco uncorrected for Hb diffusing capacity of the lung for carbon monoxide in mLCO/min/mmHg.

5) %pDLco is the uncorrected best DLco as a percentage of corrected predicted DLco for hemoglobin (Best DLco/pDLco x 100%).

- Physician Global Assessment

  The Physician Global Assessment is used to assess the physician's rating of overall health of the subject. Physicians rate the perceived health of the subject on an 11 point

scale from 0 (excellent) to 10 (extremely poor). Higher scores indicate worse disease in terms of severity, damage, or overall disease, but there is no standardization for the scale (Pope 2011). A negative change from baseline indicates improvement. Expert consensus has suggested a range of 8-13 units on Global VAS scale 100 mm.for the MCID (Gazi et al. 2007). Therefore, -0.8 to -1.3 points will be defined as MCID in this assessment.

- Modified Rodnan Skin Score (MRSS)
  The mRSS will be assessed by a qualified and trained Efficacy Assessor to limit the variability. The same Efficacy Assessor will measure skin thickness on 17 different body areas on a scale from 0 (normal) to 3 (severe). Global average scoring method (The assessor scores individual sub-areas, as needed for differences within large areas and takes average of sub-areas to attain the score for that overall area) will be used as a scoring technique in this study. Worsening of skin fibrosis as measured by the mRSS score was associated with a disease progression and overall mortality in subjects with dcSSc, indicating that the mRSS is a good surrogate marker for disease progression.
  The mRSS is a measure of skin thickness across 17 different body sites and has a range of 0 to 51 in whole units. Total score is derived from the sum of the 17 different body areas. If at least one domain score is missing, the total score will be regarded as missing. The MCID has been reported by Dinesh Khanna 2019 to be 3 to 5 units for dcSSc patients.

7.1.2.3 Exploratory Efficacy Endpoints

- [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

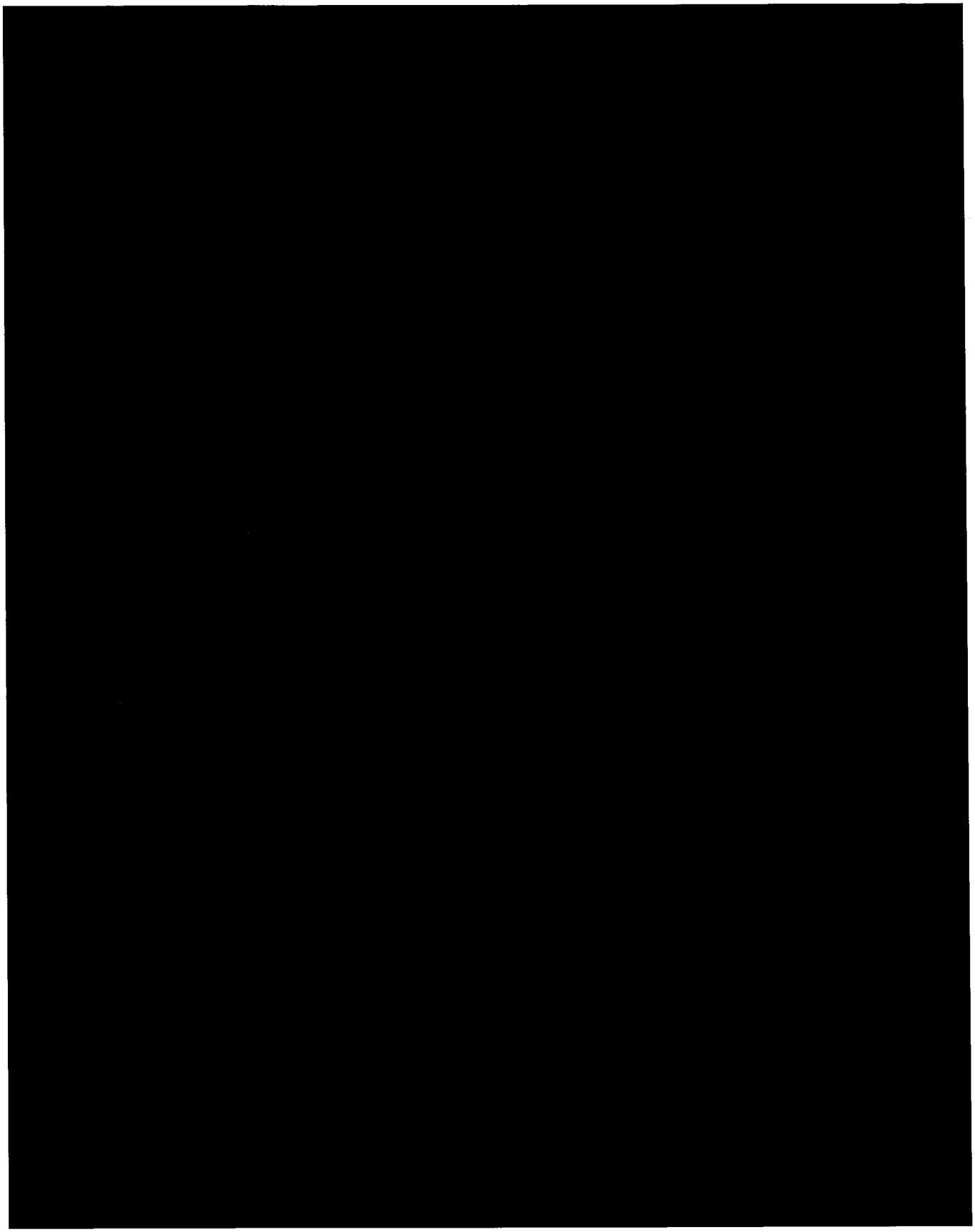
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]

[redacted]

- [redacted]

[redacted]

[redacted]



- [redacted]

[redacted]

[redacted]-Responder".

*Step1 of the new Revised CRISS is same as that of Revised CRISS.

7.1.3 Safety Assessments

7.1.3.1 Adverse Events

All AEs and SAEs will be recorded in the source documents. All AEs and SAEs that occur from the time written informed consent is obtained until the end of the Safety Follow-up Period (4 weeks after the last treatment visit, for safety monitoring) will be recorded in the eCRF.

An AE is any untoward medical occurrence in a subject or clinical investigation subject administered a pharmaceutical product that does not necessarily have a causal relationship with this treatment.

Adverse events will be coded according to the MedDRA version 25.0 or higher.

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

Protocol definition

AEs will be classified as 'treatment-emergent' if they arise following the first administration of study drug in the double blind treatment period (after randomization) or if a predose AE increases in severity following dosing in the double blind treatment period (after randomization).

Analysis rule to determine TEAE using onset date of AEs

An AE is classified as treatment emergent if it newly occurred on or after the first dose of study drug. According to this data handling, for the severity part in the above TEAE definition, the AE with upgraded severity and with the new onset date during the treatment period will be

automatically included as TEAE.

If the AE start date is incomplete, it will be imputed as follows for the purpose of determining TEAE and AE duration:

- If the start date is completely missing, the start date will be equal to the date of the first dose date of study treatment. However, if the stop date is not missing and is before the date of the first dose of study treatment, then the stop date will be used instead of the start date and the AE will not be considered as TEAE.
- If the start date is missing, but the month and year are not missing and are equal to the month and year of the first study dose, then start date will be equal to the date of the first dose date of study treatment. If the month of the start date is not equal to the month of the first study dose, then start date will be equal to the first date of the month of the start date. these events will be considered as TEAE.
- If the start date and month are missing, then the first day of the first month (January) will be used.

If an AE stop date is incomplete, it will be imputed as follows for the purpose of determining TEAE and AE duration:

- If the AE stop date is completely missing, then the stop date will be equal to the subject's last observed date.
- If the stop date is missing, but the month and year are not missing and are equal to the month and year of the last observed date, then stop date will be equal to last observed date. If the month of the stop date is earlier than the last observed month, then stop date will be equal to last day of the month of the stop date.
- If the stop date and month are missing, then the last day of the last month (December) will be used.
- If the AE stop date is prior to the date of the first dose date of study treatment, the AE will not be considered as TEAE.

(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) Serious Adverse Events

A serious Adverse Event (SAE) is defined as any untoward medical occurrence that at any dose:

- Results in death;
- Is life-threatening;

- Requires hospitalization or prolongation of existing hospitalization;
- Results in persistent or significant disability or incapacity;
- Is a congenital anomaly or birth defect;
- Is an important medical event;

(4) Trial Emergent Adverse Events

An AE is considered "Trial Emergent Adverse Events" if it occurred before the first dose of study drug at Day 1.

(5) Duration of Adverse Events

Duration of the AE and time to the AE occurrence from start of study treatment will be calculated and presented in days.

- Duration of Adverse Events (days) = AE stop date – AE start date + 1,
- Time to AE onset (days) = AE start date –The first administration date of study drug + 1

(6) AESI

One AESI that will be considered during this study includes hepatic AESIs, defined as

- Clinically significant liver dysfunction, as follows:
  - ALT or AST > 8 × ULN
  - ALT or AST > 5 × ULN for more than 2 weeks
  - Elevated total bilirubin > 2 × ULN and ALT or AST > 3 x ULN, or
  - Symptoms consistent with liver dysfunction (e.g., fatigue, nausea, vomiting, abdominal pain or tenderness, fever, rash, eosinophilia > 5%) with concomitant ALT or AST values > 3 × ULN.
- Hepatic AEs or hepatic laboratory abnormalities that lead to study drug interruption or discontinuation (see protocol).
- Any other laboratory or clinical abnormality that Sponsor Medical Monitor and/or Investigator considered as significant.

7.1.3.2 Laboratory Tests

If a laboratory test value is indeterminate due to a problem with the test sample, the value will be calculated according to the following rules:

For data below or equal to the lower limit of quantification, the lower limit of quantification will be used for tabulation. For values reported as below the lower limit of quantification, Data exceeding the lower limit of quantification will be imputed by subtracting 1 from the last digit

of the lower limit of quantification for tabulation. For data above or equal to the upper limit of quantification, the upper limit of quantification will be used for tabulation. Data exceeding the upper limit of quantification will be imputed by adding 1 to the last digit of the upper limit of quantification for tabulation.

If the data is ' >1100 ' , the data is handled as '1101'.
If the data is ' <1100 ' , the data is handled as '1099'.

(1) Criteria for pre-defined limit
Potential Drug-induced Liver Function Criteria at any post-baseline timepoint:

- ALT or AST > 8 x ULN
- ALT or AST > 5 x ULN
- ALT or AST > 3 x ULN
- Total Bilirubin > 2 x ULN
- ALT or AST > 3 x ULN and total bilirubin > 2 x ULN

Abnormal Liver Function Criteiria at any post-baseline timepoint:

- ALT (>2*ULN, >3*ULN, >5*ULN, >8*ULN)
- AST (>2*ULN, >3*ULN, >5*ULN, >8*ULN)
- GGT (>2.5*ULN, >5*ULN, >8*ULN)
- ALP (>1.5*ULN, >2.5*ULN, >5*ULN)
- Direct and total bilirubin (>1.5*ULN, >2*ULN, >3*ULN).
- ALT or AST >3xULN and ALP <2xULN and T-BIL >2xULN (Combination Abnormality (Hy's law))

7.1.3.3 12-Lead ECG

(1) Criteria for pre-defined limit

baseline corrected QT interval (QTc) < 450 msec and > 500 msec at EOT

baseline QTc < 450 msec and 500 msec ≥ QTc > 480 msec at EOT

baseline QTc < 450 msec and 480 msec ≥ QTc > 450 msec at EOT

Increase from baseline in QTc > 30 msec

Increase from baseline at EOT in QTc > 60 msec

These criteria will be applied to both corrected QT interval using Bazett's formula (QTcB) and corrected QT interval using Frederica's formula (QTcF).

7.1.4 Data Handling of PK data and melanin density data.

The PK data and MD data handling will be confirmed during blinded data review (BDR). PK data and MD data that are considered "invalid" will be flagged in the listing. Due to the nature of PK data and MD data, some issues may only be discovered after PK data and MD data are unblinded. Should new issues be identified post unblinding, and new data handling rules would have to be applied, a separate PK data and MD data handling document will be produced to provide detailed rationale and decision making. If there is clear evidence that PK sample handling errors, MD handling errors, or other factors are 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 and MD data handling document.

7.2 Analysis Visit Definitions

Table 3 Analysis Visit Window for PtGA, PhGA, mRSS, FVC, DLco, HAQ-DI

Visit Number	Analysis Visit	Nominal day	Analysis Visit Window
1	Screening	NA	NA to -1
2	Baseline	First dose day	NA
3	Week 2	Day 15	Day 2 to 29
4	Week 6	Day 43	Day 30 to 77
5	Week 16	Day 113	Day 78 to 147
6	Week 26	Day 183	Day 148 to 228
7	Week 39	Day 274	Day 229 to 319
8	Week 52 (EOT)	Day 365	Day 320 to 379
10	Week 56 (EOS/FU)	Day 393	Day 380 to 400

Table 3 Analysis Visit Window for other than PtGA, PhGA, mRSS, FVC, DLco, HAQ-DI and pregnancy test

Visit Number	Analysis Visit	Nominal day	Analysis Visit Window
1	Screening	NA	NA to -1

2	Baseline	First dose day	NA
3	Week 2	Day 15	Day 2 to 29
4	Week 6	Day 43	Day 30 to 77
5	Week 16	Day 113	Day 78 to 147
6	Week 26	Day 183	Day 148 to 228
7	Week 39	Day 274	Day 229 to 319
8	Week 52 (EOT)	Day 365	Day 320 to 379
9	Early termination (ET)	End of treatment	End of treatment + 7
10	Week 56 (EOS/FU)	Day 393	Day 380 to 400

Table 4 Analysis Visit Window for pregnancy test

Visit Number	Analysis Visit	Nominal day	Analysis Visit Window
1	Screening	NA	NA to -1
2	Baseline	First dose day	NA
3	Week 2	Day 15	Day 2 to 29
4	Week 6	Day 43	Day 30 to 77
5	Week 16	Day 113	Day 78 to 147
6	Week 26	Day 183	Day 148 to 204
	Week 32	Day 225	Day 205 to 249
7	Week 39	Day 274	Day 250 to 294
	Week 45	Day 316	Day 295 to 340
8	Week 52 (EOT)	Day 365	Day 341 to 379
9	Early termination (ET)	End of treatment	End of treatment + 7
10	Week 56 (EOS/FU)	Day 393	Day 380 to 400

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

Unless otherwise specified, the baseline value will be the last observed value of the parameter of interest prior to the first intake of study drug (this includes unscheduled visits). For FVC , %pFVC and %pDLco etc in the pulmonary function test, the baseline value will be the last observed value prior to the first intake of study drug. The data of "accepted" or "failed" generated by the centralized over read specialist will be utilized for each pulmonary function test parameters. The data accepted by the centralized over read specialist will be used if it

exits. However, when there is no accepted data available at a visit, the failed data that will be approved in BDRM will also be used.

Unless otherwise specified, no data imputation will be performed using data from outside the allowable range.

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.

In the event that in a post baseline visit window there are only values from unscheduled visit, these values will be used to derive the value for this visit.

8 STATISTICAL METHODOLOGY

All available data will be listed. Listings will be presented in treatment, subject, visit (where applicable), and date (where applicable) order.

8.1 Study Subjects

8.1.1 Subject Disposition

The following will be provided:

Number of screened subjects: defined as those who met the inclusion criteria regarding the target indication and signed the ICF

Number (%) of screen falures and reasons of screen failure

Number of randomized subjects: defined as those who received a randomization number

Number (%) of subjects who took the first drug

Number (%) of subjects in the ITT population

Number (%) of subjects who completed Week 52

Number (%) of subjects who discontinued the Randomized, Double-Blind Treatment Period, and the reasons for discontinuation of Randomized, Double-Blind Treatment Period

Number (%) of subjects who performed the follow-up visit

Number (%) of subjects who did not perform the follow-up visit and the reason for not performing the follow-up visit

Subject who used rescue therapy

Subject who reduced the dose

8.1.2 Analysis Populations

Number (%) of subjects of each analysis population will be summarized in the RAND population.

8.1.3 Protocol Deviations

Major protocol deviations will be summarized on the ITT population. The number and percentage of subjects with any major protocol deviation and within each deviation category will be summarized by treatment group in the ITT population.

8.1.4 Demographic and Other Baseline Characteristics

Demographic and other baseline characteristics will be summarized on the ITT population. The following demographic and other baseline characteristics will be used. Demographic and other baseline characteristics will be summarized by Placebo, MT-7117 over all, MT-7117 [redacted] only, [redacted] regimen, [redacted] regimen, Total.

	category	descriptive
Sex	Male, Female	
Age(years)	18 – 65 >65	Yes
Height(cm)		Yes
Weight(kg)	<45 >=45	Yes
BMI(kg/m2)		Yes
Race	White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Pacific Islander, Other	
Country	USA, Canada, UK, Italy, Spain, Germany, Poland and Belguim (from IWRS data)	
Region	North America, Europe	
Ethnicity	Hispanic or Latino, Non-Hispanic or Latino, Other	
Disease Duration of SSc		Yes
SSc-ILD	Yes, No, Uncertain	
SSc-ILD based on the HRCT assessment	Yes, No, Uncertain	

Duration of SSc-ILD		Yes
mRSS		Yes
%pFVC		Yes
HAQ-DI		Yes
PtGA		Yes
PhGA		Yes
PGIS		Yes
[redacted]		Yes
[redacted]	Normal(<6.0), Abnormal(>=6.0)	
ESR		Yes
ESR	Normal(<28), Abnormal(>=28)	
[redacted]		Yes
[redacted]		Yes
Platlets		Yes
Platlets	Normal(<330), Abnormal(>=330)	
[redacted]		Yes
[redacted]	[redacted]	
[redacted]	[redacted]	
Number of MC1R variant (MC1R variant)	e.g. 0, 1, 2, or 3	Yes

Immunosupressant and corticosteroid	Any immunosuppressant, MMF, MTX, AZA, LEF, Corticosteroid	

8.1.5 Medical History

Medical history will be summarized on the SAF population. In the above tables, SOC is sorted by International order; then within SOC, PT is sorted by descending counts under MT-7117 over all column, then descending counts under Placebo column, then alphabetic order for PTs with the same count.

8.1.6 Prior or Concomitant Medications

The prior and concomitant medications will be summarized in table separately by ATC level 2, preferred name and treatment group and presented in data listing for the SAF population.

Regarding of the concomitant medication, the dosage of the following medications which is ongoing at the time of the first dose of study drug or is started after the first dose of study drug will be descriptively summarized. Also, the duration of the medication use will be descriptively summarized. Data describing units other than "mg" or "g" will be included only in the tabulation of time period. These medication codes is attached in the Appendix 13.1. For "METHYLPREDNISOLONE" and "METHYLPREDNISOLONE SODIUM SUCCINATE", the dose multiplied by 0.8 will be used for tabulation to convert these drugs to equivalent doses of corticosteroid.

Mycophenolate (up to 3g/day)

Mycophenolic acid (up to 2.14 g/day)

Methotrexate (up to 25 mg/Week)

Leflunomide (up to 20 mg/day)

Azathioprine (up to 3 mg/kg/day)

Corticosteroid (up to $\leq$ 40 mg/day of prednisone)

8.1.7 Prohibited Medications

The prohibited medications coded by the protocol Appendix 2 will be summarized and listed on the SAF population. The data will be coded using World Health Organization Drug Dictionary (WHO-DD). The summaries will be by ATC level 2 and preferred name and treatment.

8.1.8 Rescue Medications

The Rescue Medications will be summarized and listed on the SAF population. The data will be coded using World Health Organization Drug Dictionary (WHO-DD). The summaries will be by ATC level 2 and preferred name and period and treatment. The Rescue Medications and the reason (e.g., Worsening mRSS of a minimum of 5 points and at least 25% increase relative to baseline) to use it will be summarized by Placebo, MT-7117 over all, MT-7117 [redacted] only, [redacted] to [redacted] regimen, [redacted] to [redacted] regimen.

[Analysis period using the start date of the rescue medication]

Week 16 =< Period =< Week 26

Week 26 < Period =< Week 39

Week 39 < Period =< Week 52

8.1.9 Treatment Duration and Exposure and Treatment Compliance

Treatment duration and Exposure and Treatment compliance will be summarized on the ITT population.

Treatment duration (months) will be calculated by dividing by 30.4375 for treatment duration (days). Then, treatment duration months and Exposure will be summarized as a categorical variable using the following categories:

- 0 to <= 3 months
- > 3 months to <= 6 months
- > 6 months to <= 12 months
- >12 months
- [redacted]
- [redacted]
- [redacted]

The number and percentage of treatment compliance will be calculated using the range 80% <= compliance <= 120% as a subject being treatment compliant. The number and percentage of treatment compliance will be summarized by treatment and and following period.

- Baseline to 15 Study Day
- 16 Study Day to 43 Study Day
- 44 Study Day to 113 Study Day
- 114 Study Day to 183 Study Day
- 184 Study Day to 274 Study Day

- 275 Study Day to 365 Study Day

8.1.10 Interruption/permanent discontinuation of study drug and subject question for study medication

The number and percentage of study drug interruption and study drug permanent discontinuation will be calculated with it's reason by treatment on the ITT population. Interruption duration of study drug will be summarized by treatment on the ITT population. The number and percentage of subject question for study medication will be calculated by treatment on the ITT population.

8.2 Efficacy Assessments

All efficacy endpoints will be summarized using descriptive statistics or using the number with percentage and presented in data listing.

8.2.1 Primary Efficacy Endpoint

Primary Analysis

The primary treatment comparisons of interest are the ACR CRISS score for MT-7117 treatment group compared with placebo group at Week 52. MT-7117 treatment group includes all subjects treated with MT-7117 regardless down titrated or not. The treatment MT-7117 dose level at EOT will be compared with placebo as secondary analysis.

The primary endpoint will be summarized using descriptive statistics with interquartile range. The distribution of ACR CRISS score may not be normal distribution. For ACR CRISS score at Weeks 16, 26, 39, and 52 the comparison between MT-7117 treatment group and placebo group will be performed using the non-parametric analysis method. The point estimates and 2 sided 95% CIs and associated P-values for the difference between the treatment groups will be obtained using the Hodges-Lehman estimator corresponding to Wilcoxon's rank sum test. This non-parametric analysis will be performed with multiple imputation method, assuming missing at random. A supportive analysis will be performed at the same way for the primary endpoints using the per-protocol set (PPS). As a secondary analysis for the primary endpoint, the same analysis method will be applied to the intent to treat (ITT) population but the patients' data after taking rescue therapy will not be included. This approach will be performed to investigate a sensitivity for the primary analysis.

MI Step 1: The ACR CRISS score at week 52 will be multiply imputed using Fully Conditional Specification Method methodology under MAR assumption. The imputation for the missing ACR CRISS total score at week 52 will be conducted using all visit data of all subjects with available ACR CRISS total score up to week 52. The 50 different imputed datasets will be created using the MI procedure in SAS with 2023 for seed value. The following factors will be included in the imputation model: treatment group (TRT01P), randomization strata (the baseline autoantibody status of anti-RNA Polymerase III (positive or negative)).

MI Step 2: The imputed datasets generated with the approach described above do contain only non-missing values. Non-parametric analysis using The Hodges-Lehman estimator will be performed for ACR CRISS score on each of the 50 generated imputed datasets. The ACR CRISS score at each visit (Weeks 16, 26, 39, and 52) from two treatments groups will enter the corresponding analysis. The Location shift estimate and two-sided 95% confidence intervals will be obtained using the Hodges-Lehman estimator corresponding to Wilcoxon's rank sum test. The paired comparison between each MT-7117 dose level with placebo will also be performed using this non-parametric analysis method. The following SAS code will be used:

MI Step 3 : 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, ..., W52 according to the following code. In addition to the estimates, corresponding 95% confidence intervals and p-values will be calculated.

The following example SAS code will be used:
*Combine LS mean difference estimates;

[redacted]
[redacted]TDERR;
RUN;

Prespecified Subgroup Analysis

The treatment effect on the primary endpoint of CRISS score up to Week 52, the seconday endpoint of change from baseline to Week 52 in mRSS and new Revised CRISS up to Week 52 across different subgroups will be explored based on the ITT population for the following subgroups.

Baseline CRP (<6.0, >=6.0, mg/dL)

Baseline ESR (<28, >=28, mm/hr)

Baseline Platelets (<330, >=330, x10^9/L)

Baseline SSc disease duration (<=3 years, >3 and <=5)

Baseline mRSS score (<25, >=25)

[redacted]

[redacted]

Concomitant medication (MMF, no-MMF)

Concomitant medication (Immunosupressant(IM)*, non-IM)

Baseline disease duration and concomitant medication (<2 years+MMF, <2 years+no-MMF, >=2 years+MMF, >=2 years+no-MMF)

Region (North America, Europe)

MC1R variant (number of MC1R variants, e.g. 0, 1, 2, or 3 and each MC1R variant (MC1R variant: n≥3))

[redacted]

[redacted]

Baseline %pFVC (>50% and <=80% , >80%)

*Immunosupressant(IM): methotrexate, leflunomide, azathioprine, mycophenolate and mycophenolic acid

The impact of MT-7117 dose reduction to efficacy

Efficacy analysis of MT-7117 vs. Placebo will be basically performed based on the ITT population. In this study dose reduction due to significant intolerable AEs is allowed. If efficacy of MT-7117 [redacted] only, [redacted] to [redacted] regimen, [redacted] to [redacted] regimen is investigated, the randomization principle cannot be maintained so that there is comparibility of these regimens. However, given this limitation, in order to investigate the potential treatment effect, the treatment effect on the primary endpoint of CRISS score up to Week 52 and the seconday endpoint of change from baseline up to Week 52 in mRSS, HAQ-DI, Patient Global Assessment, %pFVC, Physician Global Assessment and Revised CRISS up to Week 52 across the following treatment regimens will be explored based on the ITT population. Then, demographic and disease characteristics specified in Section 8.1.4 will be similary evaluated.

Treatment comparison is the following;

MT-7117 [redacted] only,

MT-7117 [redacted]

MT-7117 [redacted]

vs. Placebo

8.2.2 Secondary Efficacy Endpoints

Continuous endpoint

Change from baseline at 16, 26, 39 and 52 in the followings;

- HAQ-DI
- Patient Global Assessment
- %pFVC
- Physician Global Assessment
- mRSS

The above endpoint will be analyzed using mixed-effect model for repeated measures (MMRM). The model will include fixed categorical terms for treatment, randomization strata (autoantibody status of anti-RNA Polymerase III at screening (positive or negative)), visit, and treatment by visit interaction together with continuous covariate terms for baseline value of the endpoint and baseline value by visit interaction. An unstructured correlation structure will be used to model the within-subject variance covariance errors. Should convergence of the model fail (due to the small numbers of subjects in this study), other variance covariance matrices such as autoregressive [AR(1)] correlation matrix will be used if appropriate. The Kenward-Roger approximation will be used to estimate the denominator degrees of freedom. From the model

described above, adjusted (least squares [LS]) means and standard errors will be produced by treatment and visit. Difference in adjusted means at each visit (MT-7117 vs. placebo) with standard errors, 95% CIs and associated P values will also be produced. All available data from all subjects will be used in the analysis without any imputation.

ACR CRISS score responder (CRISS $\geq$ 0.6)

Proportion of subjects with ACR CRISS score responder (CRISS $\geq$ 0.6) at Weeks 16, 26, 39, and 52 will be analyzed using logistic regression model employing SAS GLIMMIX procedure. The model will include fixed categorical terms for the treatment group (TRTPN), the randomization factor (screening autoantibodyanti-RNA Polymerase III, positive or negative - STRTNC), visit, and treatment group (TRTPN) by visit interaction together with continuous covariate terms for baseline mRSS (BLMRSS). The treatment odds ratio at each visit will be estimated using a contrast (using ilink PDIFF options in LSMEANS). An unstructured correlation structure (UN) will be used to model the within-subject variance covariance errors. Should convergence of the model fail (due to the small numbers of subjects in this study), other variance covariance matrices such as autoregressive [AR(1)] or compound symmetry (CS) correlation matrix will be used if appropriate. The following SAS code will be used.

[redacted]

8.2.3 Exploratory Efficacy Endpoints

Continous endpoint

- [redacted]

[redacted]

[redacted]

8.3 Safety Assessments

Safety assessments will be made on the SAF population. All data will be listed.

8.3.1 Adverse Events

The overall TEAEs are summarized by Placebo, MT-7117 over all, MT-7117 [redacted] only, [redacted] to [redacted] regimen, [redacted] to [redacted] regimenfor subjects with

- at least one TEAE,
- at least one treatment emergent adverse reaction,
- at least one serious TEAE,
- at least one serious treatment emergent adverse reaction,
- at least one TEAE leading to drug withdrawn,
- at least one treatment emergent adverse reaction leading to drug withdrawn,
- at least one hepatic TEAE,
- at least one hepatic TEAE of Special Interest (AESI)
- at least one hyperpigmentation TEAE,
- at least one Adverse Events of Special Interest (AESI)
- at least one TEAE leading to death.

The frequency and incidence of TEAEs will be summarized by System Organ Class (SOC) and Preferred Term (PT) by Placebo, MT-7117 over all, MT-7117 [redacted] only, [redacted] to [redacted] regimen, [redacted] to [redacted] regimen. For this table, SOC is sorted by International order; then within SOC, PT is sorted by descending counts under MT-7117 Total group, then descending counts under Placebo group, then alphabetic order for PTs with the same count.

The AE summaries will be presented by treatment group for the following:

- TEAEs by SOC and PT
- TEAEs by SOC, PT and severity
- TEAEs by SOC, PT and relationship to study drug

- Treatment emergent adverse reactions by SOC and PT
- Treatment emergent adverse reactions by SOC, PT and severity
- Serious TEAEs by SOC and PT
- Serious TEAEs by SOC and PT and relationship to study drug
- Serious treatment emergent adverse reactions by SOC and PT
- TEAEs leading to drug withdrawn by SOC and PT
- Ttreatment emergent adverse reaction leading to drug withdrawn by SOC and PT
- Hepatic TEAEs by SOC and PT
- Hepatic TEAE of Special Interest (AESI) by SOC and PT
- Hyperpigmentation TEAEs by SOC and PT
- TEAEs of special interest by SOC and PT
- TEAEs by SOC and PT for TEAEs with frequency >= 5% in total MT-7117 group

For each of the summaries will be done 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 intensity category (severe > moderate > mild) and/or maximum drug relationship category (reasonable possibility, no reasonable possibility). If intensity or relationship is found to be missing the most severe occurrence will be imputed for that particular summary.

All AEs for each subject, including multiple occurrences of the same event, will be listed.

Deaths that occur during the study will be listed in a data listing. The data listings for serious TEAE and TEAE leading to drug withdrawn and trial emergent adverse event will be generated as well.

The frequency and incidence of TEAEs will be summarized by System Organ Class (SOC) and Preferred Term (PT) by Placebo, MT-7117 over all, Treatment duration (0 to <=3, >3 to <= 6, >6 to < 12, =12 (months)). For this table, SOC is sorted by International order; then within SOC, PT is sorted by descending counts under MT-7117 Total group, then descending counts under Placebo group, then alphabetic order for PTs with the same count.

The AE summaries will be presented by treatment group for the following:

- TEAEs by SOC and PT
- Serious TEAEs by SOC and PT

8.3.2 Laboratory Tests including Pregnancy Test

Absolute values and changes from baseline will be summarized for the following laboratory tests parameters by treatment group.

Laboratory Test	Parameters
Hematology	Haemoglobin, Haematocrit, Red Blood Cell, Platelets, Mean corpuscular hemoglobin, Mean corpuscular hemoglobin concentration, Mean corpuscular volume, White blood cell count and differential, Erythrocyte sedimentation rate (ESR)[c]
Biochemistry	Alkaline phosphatase (ALP), Aspartate aminotransferase (AST), Alanine aminotransferase (ALT), γ-glutamyl transpeptidase (GGT), Potassium, Sodium, Chloride, Inorganic phosphate, Glucose, Bilirubin (direct and total), C-reactive protein (CRP), Cholesterol, Triglycerides, High density , ipoprotein-cholesterol (HDL), Low density lipoprotein-cholesterol (LDL), Protein (total), Albumin, Creatine phosphokinase (CPK), Creatinine, Ferritin, Calcium Blood urea nitrogen (BUN), hCG[b]
Coagulation	Prothrombin time (PT), International normalized ratio (INR), Activated partial thromboplastin time(aPTT)
Urinalysis	pH, Specific Gravity (Qualitative value): protein, glucose, ketones, urobilinogen, blood Microscopic examination[a]
Serology	Screening only: [redacted] Baseline (day 1 of Week 1 prior to randomization) and Weeks 16, 26, and 52: [redacted]

[a] Performed only if required, based on urinalysis results

[b] Female subjects of child-bearing potential only; serum pregnancy test will be performed at Visit 1 and a urine pregnancy test at all other site visits and a urine dipstick pregnancy test at home or other home testing arrangements will be performed at Telephone Visits (Visit 6A and 7A).

[c] ESR will be analyzed locally at the site level and at Screening only.

All laboratory data will be listed with clinically relevant values flagged (Low = Lower than lower limit of normal range, High = Higher than ULN range or Normal = Normal for or Hematology, Biochemistry, Urinalysis and Coagulation, N=Normal and A=Abnormal for Urinalysis). And then shift table of clinically relevant categories will be presented at post-baseline timepoint. Laboratory Urinalysis Absolute will be summarized by flag (-, +/-, 1+, 2+, 3+, 4+) and analysis visit.

The number and percentage of subjects with potential drug-induced liver function criteria and abnormal liver function criteiria at post-baseline timepoint will be summarized.

The figure of mean (or median) and standard error value of ALT, AST, total bilirubin, and ALP by visit will be plotted.

The pregnancy test result (Negative, Positive) will be summarized at each visit.

If laboratory values in central measurement and/or in local site are included, only values in central measurement will be summarized. In addition, all laboratory values including value measured in local site will be also listed.

8.3.3 Vital Signs

Absolute values and change from baseline will be descriptively summarized and shift table of clinically relevant categories (Normal, Not Clinically Significant, Clinically Significant) from baseline to EOT will be presented for the following parameters . All tabulations will be performed by treatment group.

- Systolic Blood Pressure (mmHg)
- Diastolic Blood Pressure (mmHg)
- Pulse Rate
- Respiratory Rate
- Body Temperature (°C)
- Weight (kg)

8.3.4 12-Lead ECGs

Absolute values and changes from baseline will be summarized and shift table of clinically relevant categories (Normal, Not Clinically Significant, Clinically Significant) from baseline to EOT will be presented for the following parameters.

- Heart Rate (bpm)
- RR (msec)

- QRS (msec)
- QT (msec)
- QTcF (msec)
- QTcB (msec)

The percentage of subjects with 12-lead ECG values outside pre-defined limit will be summarized. Number and percentage of subjects meeting the criteria listed 7.1.3.3 will be presented in tables. All tabulations will be performed by treatment group.

8.3.5 Physical Examinations

The number and percentage of the subjects for clinically relevant categories (Normal, Not Clinically Significant, Clinically Significant) will be summarized by treatment group.

- Abdominal
- Respiratory
- Cardiovascular
- General Appearance
- Head
- Eyes
- Ears/nose/throat
- Lymph nodes
- Musculoskeletal
- Neck
- Neurological
- Dermatological
- Other

8.3.6 Nevi appearance

Nevi (Melanocytic Lesions) appearance will be summarized with the number and percentage of the subjects' suspicious nevi (Melanocytic Lesions) found:

- Does the subject have any suspicious nevi found during assessment? (Y,N)
- Has study been temporarily treatment discontinued? (Y,N)
- Has the subject been referred for further evaluation? (Y,N)
- Was the biopsy performed and sent to the central lab? (Y,N)
- Were there any clinically significant findings? (Y,N)

All tabulations will be performed by treatment group.

8.4 Pharmacokinetics Evaluation

Pharmacokinetics evaluation will be made on the PK population. Plasma MT-7117 concentrations will be listed for each subject. Plasma concentrations of MT-7117 will be analyzed at randomization visit (Day 1 pre-dose), Week 2 (any time), Week 6 (any time), Week 16 (any time), Week 26(any time), Week 39(any time), and Week 52 (any time). PK sample collection times, most recent dosing times before PK sample collection, as well as derived actual sampling time relative to the most recent dose will be provided in a listing. The actual sampling time relative to the most recent dose will be calculated in hours and rounded to 2 DP and the mean values of actual sampling time will be calculated for each visit. Plots of individual concentration vs actual sampling time will be presented overlaid with treatment in different symbols for each treatment and overlaid with visit. For the calculation of the summary statistics (mean, SD, minimum, median, and maximum) of palsma concentrations at each sampling point, concentration values reported as below the limit of quantification (BLQ) will be set to 0.

Population PK analysis will be performed using the plasma concentration of MT-7117 obtained in this study in combination with data obtained from other clinical studies. Population PK analysis results will be reported separately as necessary.

8.5 Pharmacogenetics Evaluation

8.6 Exploratory endpoint

[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)
		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 $p<0.001$

(2) Plasma concentration

Statistic	Specification

Individual value, Mean, SD, Minimum, Maximum, Median	The same significant digits as they are reported

9.2 Treatment Groups to Report

Unless otherwise specified, the following treatment groups will be reported.

Treatment Group	For TFLs
MT-7117 [redacted] with dose reduction to [redacted]	MT-7117
Placebo	Placebo

For only CRISS score and mRSS, the following treatment groups will be reported as the secondary analysis.

Treatment Group	For TFLs
MT-7117 [redacted]	MT-7117 [redacted]
MT-7117 [redacted]	MT-7117 [redacted]
MT-7117 [redacted]	MT-7117 [redacted]
Placebo	Placebo

9.3 Analysis Visits to Report

Efficacy:

Analysis Visit	Variables										
	ACR CRISS	Revised CRISS	mRSS / %pFVC	SHAQ	PTGA/ PHGA	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Screening			X							X	
Baseline			X		X		X	X	X		X
Week 2									X		
Week 6						X	X		X		
Week 16	X	X	X	X	X	X	X		X		

Week 26	X	X	X	X	X	X	X	X	X	X	X
Week 39	X	X	X	X	X	X	X		X	X	X
Week 52	X	X	X	X	X	X	X	X	X	X	X
Week 56 (EOS/FU)									X		

Safety:

Analysis Visit	Variables				
	Lab	Vital Signs	ECGs	Nevi	Physical exam
Screening	X	X	X	X	X
Baseline	X	X			X
Week 2	X	X		X	X
Week 6	X	X		X	X
Week 16	X	X			X
Week 26	X	X	X	X	X
Week 39	X	X			X
Week 52	X	X	X	X	X
Week 56 (EOS/FU)	X	X		X	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.

10 CHANGE FROM THE PROTOCOL

- "The total number (%) of subjects in each analysis populations" in Subject Disposition was changed to "Number (%) of subjects in the ITT population".
- Year of birth was deleted from "Demographic and Other Baseline Characteristics".
- The "baseline mRSS by visit interaction" was deleted from generalized linear mixed effect model of "ACR CRISS score responder (CRISS $\geq$ 0.6)".

- The "Revised CRISS" and "new Revised CRISS" will be analyzed using generalized linear mixed effect model.
- "ACR CRISS Score improvement at Weeks 16, 26, 39 and 52: Proportion of subjects with ≥ 25% improvement in mRSS, HAQ DI, Patient Global Assessment, Physician Global Assessment, or ≥ 5% improvement in FVC for at least 3 of the 5 ACR CRISS measures." was deleted from 8.2.2 Secondary Efficacy Endpoints, because "the revised CRISS at each timepoint with >= the predefined 25% improvement in mRSS, HAQ-DI, PtGA and PhGA, or >= 5% improvement in %pFVC for at least 2 of the 5 core set measures and the revised CRISS at each timepoint with >= the predefined 25% worsening in mRSS, HAQ-DI, PtGA and PhGA, or >= 5% worsening in %pFVC for no more than 1 of the 5 core set measures" was analyzed in "new Revised CRISS" of 8.2.3 Exploratory Efficacy Endpoints.
- Following abnormal liver function criteiria was changed.
  ALT was changed from ">=2*ULN, >=3*ULN" to ">2*ULN, >3*ULN, >5*ULN, >8*ULN".
  AST was changed from ">=2*ULN, >=3*ULN" to ">2*ULN, >3*ULN, >5*ULN, >8*ULN".
  GGT was changed from ">=2*ULN, >=3*ULN" to ">2.5*ULN, >5*ULN, >8*ULN".
  ALP was changed from ">=2*ULN, >=3*ULN" to ">1.5*ULN, >2.5*ULN, >5*ULN".
  Direct and total bilirubin was changed from ">=2*ULN, >=3*ULN" to "1.5*ULN, >2*ULN, >3*ULN".
- Standard errors of ALT, AST, total bilirubin, and ALP by visit were changed to standard deviations.

11 SOFTWARE

All statistical analyses will be performed using SAS version 9.4 or higher.

12 REFERENCES

- [redacted]
- THE HEALTH ASSESSMENT QUESTIONNAIRE (HAQ) DISABILITY INDEX (DI) OF THE CLINICAL HEALTH ASSESSMENT QUESTIONNAIRE (VERSION 96.4) (https://www.niehs.nih.gov/research/resources/assets/docs/haq_instructions_508.pdf)

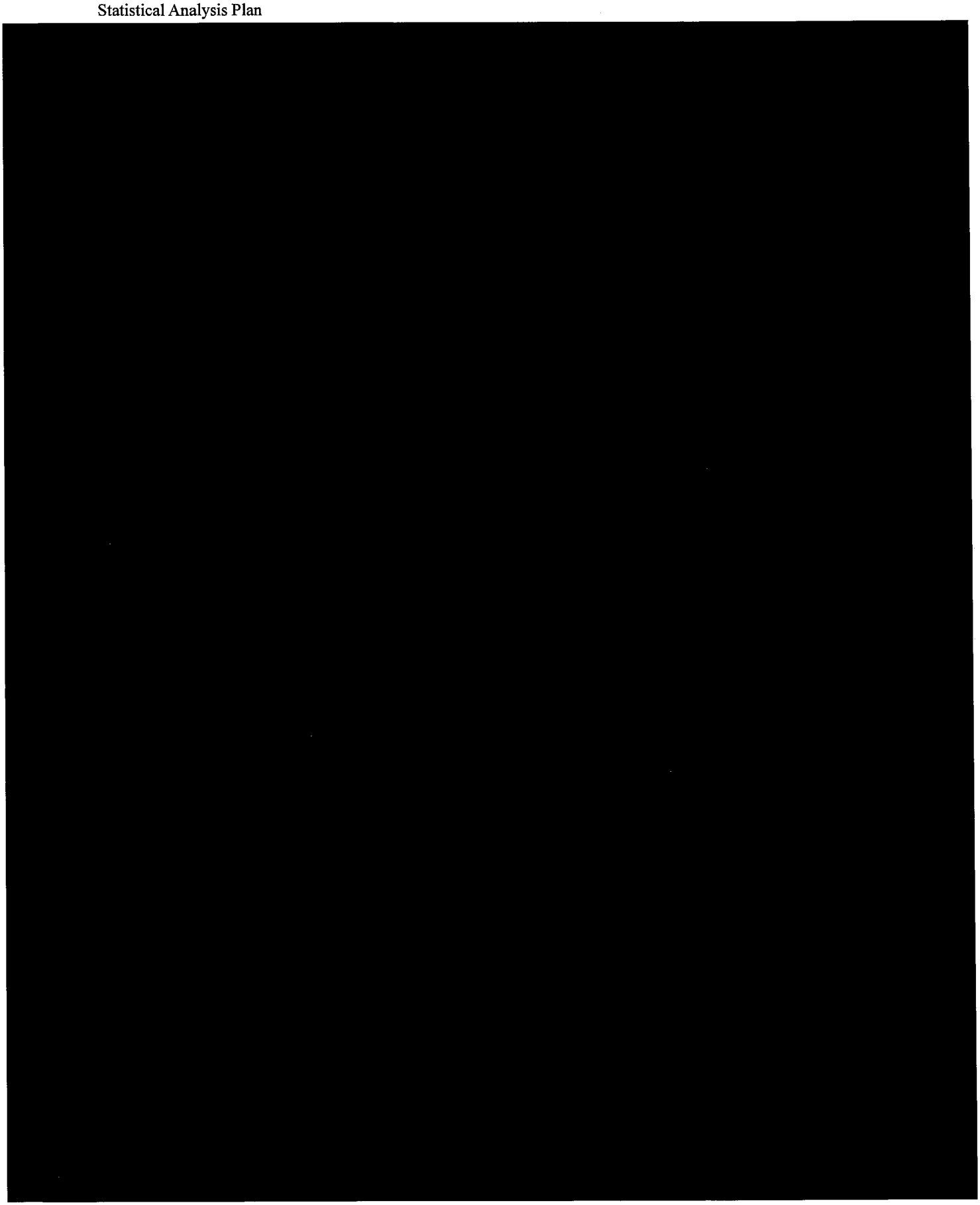
13 APPENDIX

13.1 Appendix 13.1 List of WHO-DD Drug record number, Sequence Number1,

Sequence Number2, Drug name

This file is attached to this dtatistical analysis plan.

13.2 Appendix 13.2



[redacted]