Document Type:	Statistical Analysis Plan
Official Title:	A Randomized, Double-Blind, Placebo-Controlled Phase II Study to Investigate the Efficacy and Safety of Riociguat in Patients With Diffuse Cutaneous Systemic Sclerosis (dcSSc)
NCT Number:	NCT02283762
Document Date:	06 MAR 2019

A Randomized, Double-Blind, Placebo-Controlled Phase II Study to Investigate the Efficacy and Safety of Riociguat in Patients With Diffuse Cutaneous Systemic Sclerosis (dcSSc)

Riociguat in diffuse cutaneous systemic sclerosis (dcSSc)

Bayer study drug	BAY 63-2521 / Riociguat		
Study purpose:	To evaluate exploratory outcomes and safety in subjects with diffuse cutaneous systemic sclerosis (dcSSc) treated with riociguat in the long-term extension phase		
Clinical study phase:	II	**Date:**	06 MAR 2019
Study No.:	16277	**Version:**	Final 1.0
Author:	PPD		

This Statistical Analysis Plan is produced on a word-processing system and bears no signatures.

The approval of the Statistical Analysis Plan is documented in a separate Signature Document.

Table of Contents

1. Introduction 5

2. Study Objectives 5

3. Study Design 6
3.1 Design overview – Long term extension phase 6
3.1.1 Dose-Titration period (Visits 12 to 17; Week 52 to 62) 6
3.1.2 Open-label Extension period (Visit 18; Week 64 onward) 6
3.1.3 Termination visit and Safety follow-up visit 7
3.1.4 Definition of rescue therapy 7
3.2 End of study 7

4. General Statistical Considerations 8
4.1 General Principles 8
4.2 Handling of Dropouts 8
4.3 Handling of Missing Data 8
4.4 Interim Analyses and Data Monitoring 9
4.5 Data Rules 9
4.5.1 Definition of baseline and handling of repeated measurements 9
4.5.2 Definition of treatment-emergence 9
4.5.3 Definition of regions 9
4.6 Unblind Review 9

5. Analysis Sets 10
5.1 Assignment of analysis sets 10

6. Statistical Methodology 10
6.1 Population characteristics 10
6.1.1 Demographics and baseline characteristics 10
6.1.2 Medical history and concomitant medication 11
6.1.3 Study medication duration and exposure 11
6.2 Efficacy 12
6.2.1 mRSS 13
6.2.2 Pulmonary function testing 13
6.2.3 PROs / HRQoL 13
6.2.4 Worsening of end-organ diseases 14
6.2.5 Subgroup analysis 14
6.3 Pharmacokinetics/pharmacodynamics 15
6.4 Safety 15
6.4.1 Adverse events and mortality 15
6.4.2 Laboratory data 16
6.4.3 Other safety parameters 16
6.4.4 Subgroup analysis 17
6.4.5 Possible changes of observations in main phase 17

7. Document history and changes in the planned statistical analysis 18

8. References 19

9. Appendices 20

Abbreviations

ACR	American College of Rheumatology
ADS	Analysis dataset
AE	Adverse event
ATC	Anatomical therapeutic chemical
BDG	Bayer Drug Grouping
BMI	Body mass index
CM	Concomitant medication
CRF	Case report form
CRP	C-reactive protein
CSR	Clinical study report
CTEPH	Chronic thromboembolic pulmonary hypertension
dcSSc	Diffuse cutaneous systemic sclerosis
DL_{CO}	Diffusion capacity of the lung for carbon monoxide
DMC	Data Monitoring Committee
ECG	Electrocardiogram
EU	European Union
EULAR	European League Against Rheumatism
FAS	Full analysis set
FVC	Forced vital capacity
HAQ-DI	Health Assessment Questionnaire disability index
HRCT	High resolution computed tomography
HRQoL	Health-related quality of life
ILD	Interstitial lung disease
ITT	Intent-to-treat
IxRS	Telephone- or web-based response system
lcSSc	Limited cutaneous systemic sclerosis
LOQ	Limit of quantification
LTE	Long term extension
MedDRA	Medical Dictionary for Regulatory Activities
MH	Medical history
mRSS	modified Rodnan skin score
NSAID	Nonsteroidal anti-inflammatory drug
PAH	Pulmonary arterial hypertension
PROs	Patient-reported outcomes
PT	Preferred term
RAVE	Validated electronic system used for data collection
SAE	Serious adverse event
SAP	Statistical analysis plan
SAS	Statistical analysis system
SBP	Systolic blood pressure
SF-36	Short Form 36
SHAQ	Scleroderma Health Assessment Questionnaire
SOC	System organ class

SSc	Systemic sclerosis
TEAE	Treatment-emergent adverse event
TLF	Tables, Listings and Figures
TID	*ter in die*, 3 times a day
VAS	Visual analog scale
WHO-DD	World Health Organization Drug Dictionary

1. Introduction

Systemic sclerosis (SSc) is a rare, orphan disease featuring chronic, fibrosing, autoimmune responses characterized by small vessel vasculopathy, autoantibody production, and fibroblast dysfunction leading to increased deposition of extracellular matrix. Systemic sclerosis is further divided into 2 subtypes defined by the extent of skin involvement: limited cutaneous systemic sclerosis (lcSSc) and diffuse cutaneous systemic sclerosis (dcSSc).

Both dcSSc and lcSSc are associated with internal organ involvement; however, patients with dcSSc are at greater risk for clinically significant major organ dysfunction. Diffuse cutaneous SSc is one of the most fatal rheumatic diseases, and is associated with substantial morbidity and many detrimental effects on health-related quality of life (HRQoL).

Currently, no therapy has been proven to reverse the vascular and fibrotic damage in patients with scleroderma. However, due to the high medical need, a number of drugs, such as methotrexate, mycophenolate mofetil, cyclophosphamide, azathioprine, and cyclosporine, are used off-label in an attempt to slow the progression of fibrosis. Current treatment options only target various SSc-related symptoms. No disease-modifying drug is available for SSc. In the European Union (EU), only bosentan is approved "to reduce the number of new digital ulcers in patients with SSc and ongoing digital ulcer disease" thus addressing also only one aspect of the disease.

Based on the positive results of riociguat in patients with PAH and chronic thromboembolic pulmonary hypertension (CTEPH) together with the compound's known anti-proliferative and antifibrotic effects as seen in vitro and in animal models, patients with SSc may benefit from treatment with riociguat. The current study will be the first study testing riociguat in this indication. The efficacy and safety of riociguat in patients with dcSSc will be evaluated.

This Statistical Analysis Plan is based on the following document(s):

Integrated Clinical Study Protocol, Amendment 7, version 6.0 dated 17 APR 2018

The SAP for the main treatment phase of the study was finalized prior to unblinding to describe the evaluation the efficacy and safety of 52-week treatment (plus 30-day safety follow-up, when applicable) with riociguat compared to the placebo (16277 Statistical Analysis Plan - Final Version 4_0). The SAP describing the interim benefit-risk evaluation for the long term extension (LTE) phase was finalized on 12 JUL 2018 (16277 Statistical Analysis Plan for Benefit-risk update version 1.0). This SAP describes the complete evaluation of the LTE phase. Results from the final LTE analyses will be presented in the Addendum to the Clinical Study Report (CSR).

2. Study Objectives

The overall objective of the main phase of study 16277 was to evaluate the efficacy and safety of 52 weeks of treatment with riociguat versus placebo in subjects with dcSSc. For details on the primary and secondary objectives of the study, see the SAP of the main treatment phase.

The main focus of the evaluation of the LTE phase described in this SAP was to allow a continued exploratory assessment of benefit risk profile of riociguat.

3. Study Design

3.1 Design overview – Long term extension phase

The study design consists of a main study treatment phase followed by a LTE phase. At Visit 12 (Week 52), patients randomized to the placebo arm in the main treatment phase were initiated active treatment with riociguat, as part of the LTE phase. During the first 10 weeks, a titration/sham-titration was performed to maintain the study blind, after which the extension phase becomes open-label. During the LTE phase, clinical outcomes were continued to be measured for exploratory long-term treatment effects. Adverse event (AE) information was continued to be collected throughout the LTE phase.

As defined in the protocol Section 4.1.3, the LTE phase was planned to continue up to 6 years after the last patient last visit (LPLV) in the main treatment phase. However, this study was terminated (LPLV planned in MAR 2019), following the Sponsor´s decision to stop pursuing the development of riociguat in dcSSc indication.

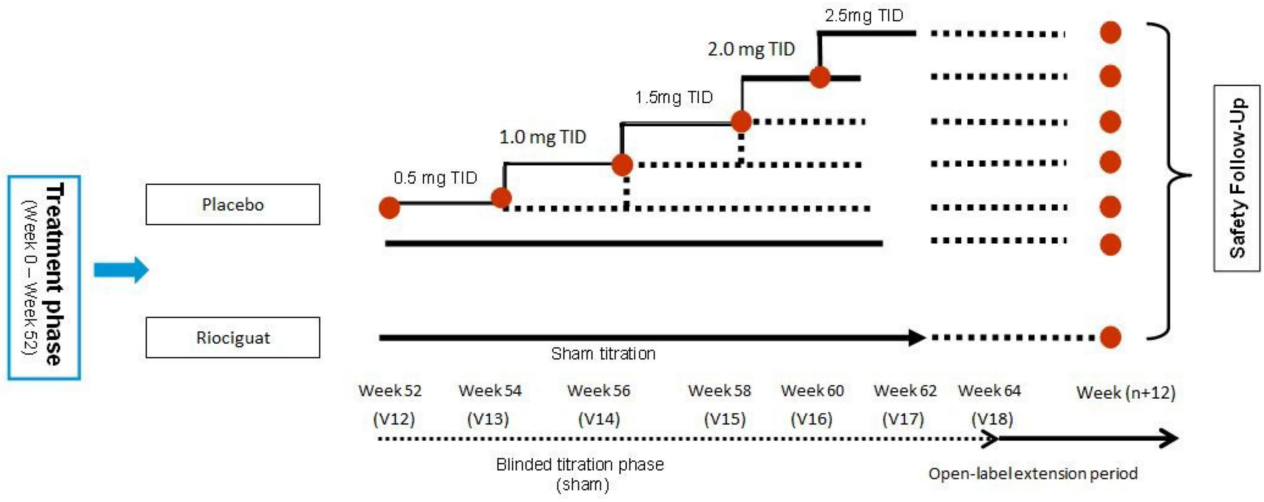
3.1.1 Dose-Titration period (Visits 12 to 17; Week 52 to 62)

At Visit 12 (Week 52), all patients were assigned in IxRS to treatment with riociguat. To maintain the study blinding, neither the patients prior treatment assignment nor dose were unblinded. During the first 10 weeks of LTE phase, patients previously on placebo were up-titrated on riociguat according to the dose titration algorithm described below (Section 3.1.2). Patients randomized to riociguat in the main treatment phase underwent sham titration to maintain the blinding during this period.

3.1.2 Open-label Extension period (Visit 18; Week 64 onward)

After completing the double-blind dose-titration period, from Week 64 onwards the study was open-label and the investigator and the patient were able to see the current treatment dose. All patients returned to the clinic for visits every 3 months ± 2 weeks. Ongoing titration of riociguat within the range of 0.5 mg to 2.5mg TID remained at the discretion of the investigator, and dose reductions for safety reasons were allowed.

Figure 3–1 - LTE phase design



Dose Interruptions

Although not intended, patients could interrupt their intake of study medication for various reasons (e.g., hospitalization in a remote hospital without study medication access, safety reasons or side effects). If treatment was interrupted, the following rules were applied:

LTE phase:

- ≤3 consecutive days without treatment (9 missing doses) in the dose-titration period: restart at the same dose
- >3 days but ≤14 consecutive days without treatment in the dose-titration period: treatment can be restarted at the discretion of the investigator at 0.5 mg TID lower than the last dose
- >14 consecutive days without treatment in the dose-titration period: discontinue the patient from study medication.
- Interruptions in the open-label extension period:
    - >13 and ≤28 consecutive days: treatment can be restarted at the discretion of the investigator at 0.5 mg TID lower than the last dose.
    - >28 consecutive days: discontinue the patient from study medication.

3.1.3 Termination visit and Safety follow-up visit

A termination visit should have been performed for patients who discontinued, per clinical study protocol, from study medication for any reason except death or lost to follow-up, and should have occurred as soon as possible after the patient received his/her last dose of study drug. In general, at the Termination Visit the same safety and efficacy relevant measurements and procedures should have been performed as at Visit 12. If the Termination Visit was performed after Visit 12 (Week 52) the patient's and physician's global assessment, patient interference with skin assessment, and tender and swollen joint count assessment must not have been performed.

A safety follow-up visit (30 [+5] days after the last dose of study medication) were to be performed for all the patients.

3.1.4 Definition of rescue therapy

During the LTE phase (ie, after completion of all efficacy- and safety-related procedures at Visit 12), except for the contraindicated nitrates or NO donors and PDE5 inhibitors, the addition of any other concomitant medication was at the discretion of the investigator.

Use of rescue therapy agents (4 agents: methotrexate, mycophenolate mofetil, cyclophosphamide or azathioprine are counted as defined in the main SAP) will be tabulated by therapy from Week 52 to the end of the study date and overall by timepoint.

3.2 End of study

For each participating EU country, the end of the study according to the EU Clinical Trial Directive will be reached when the last visit of the last subject for all centers in the respective country has occurred.

The end of the study as a whole will be reached as soon as the end of the study according to the above definition has been reached in all participating countries (EU and non-EU).

4. General Statistical Considerations

4.1 General Principles

The statistical evaluation will be performed by using the software package SAS release 9.2 or higher (SAS Institute Inc., Cary, NC, USA).

The statistical analysis will conform to Bayer Global Standard Tables version 3.0, 16 JAN 2017 with any riociguat project-specific options for these tables, and also any additional riociguat project-specific tables (Riociguat Standard Tables version 2.0, 19 JUN 2013). Subject data listings will conform to Bayer Global Standard Listings version 3.1, 16 JAN 2017. Additional data summaries not contained within these standards will be study specific.

All variables will be analyzed by descriptive statistical methods. The number of data available and missing data, mean, standard deviation, minimum, median, quartiles (if data are clearly non-normal) and maximum will be calculated for continuous data. Frequency tables will be generated for categorical data.

All the analyses will be performed in patients valid for the full analysis set (FAS), who took part in the LTE. This is the long-term safety analysis set.

4.2 Handling of Dropouts

The frequency of enrolled subjects not completing LTE phase and associated reasons will be summarized.

A patient who discontinues study participation prematurely for any reason is defined as a "dropout", if the patient has already been entered to LTE phase.

Any patient, if entered to long-tem extension phase, removed from the trial should undergo the assessments at the termination visit.

4.3 Handling of Missing Data

All missing or partial data will be presented in the subject data listing as they are recorded on the Case Report Form (CRF).

Kaplan-Meier plot for "Time to end of study treatment" will be provided.

General rules

When appropriate, the following rules will be implemented so as not to exclude subjects from descriptive analyses due to missing or incomplete data:

- **Safety Variables**
  When only partial dates are available, the following rules will be used for the derivation:

  If either the day or month of the start date of the adverse event is missing, then a worst case assumption is made for the treatment-emergent flag. For example, if study medication starts on 15 JAN 2016 and the AE start date is recorded as JAN 2016, then

this is considered treatment-emergent, as it is possible the adverse event started while the patient is on study medication.

4.4 Interim Analyses and Data Monitoring

No formal interim analyses for LTE phase will be done. There will be no Data Monitoring Committees either.

4.5 Data Rules

Analysis datasets (ADS) containing all derived variables needed for the statistical evaluation will be generated. The structure of the ADS and the contained variables will be described in a separate specification document.

Efficacy analysis datasets will be created that include key data, such as demography, flag for use of rescue medication, baseline efficacy. Relative days and flags for treatment emergent events are included in the databases.

The rules for data handling are described in detail in the Project Data Handling Rules, current version 1.2, dated 09 JAN 2014, and any updated versions becoming available during the course of this study.

4.5.1 Definition of baseline and handling of repeated measurements

Baseline for the main phase is defined as the last set of non-missing measurements taken prior to the first intake of study medication, called ‘baseline for main treatment phase’.

In case of multiple measurements per post baseline visit, the last non-missing value per visit will be taken for analysis.

For LTE phase, Week 52 will serve as baseline measurement for LTE (‘baseline for LTE phase’) to be compared with measurements during the LTE phase.

4.5.2 Definition of treatment-emergence

Values will be considered treatment-emergent if they start within 2 calendar days after the last day of study drug administration, i.e. the treatment-emergent window will be 2 days.

4.5.3 Definition of regions

The following regions are used for further subgroup analyses:

- Europe and Australia/New Zealand: Belgium, Netherlands, Switzerland, Germany, Czech Republic, Hungary, Turkey, Spain, Italy, France, UK, Australia, and New Zealand
- North America: US and Canada
- East Asia: Japan.

4.6 Unblind Review

The results of the unblinded validity review meeting will be documented in the Validity Review Report.

5. Analysis Sets

5.1 Assignment of analysis sets

Long term safety analysis set will be used in LTE phase. It will consist of all patients randomized and treated with study medication (FAS), who have continued treatment with study medication in the LTE phase of the study.

6. Statistical Methodology

Although dose titration is in use, the descriptive analyses will be shown to the following treatment groups:

- Patients randomized to riociguat in the main phase who stayed on riociguat in the LTE phase (called Riociguat-Riociguat in outputs)
- Patients randomized to placebo in the main phase who switched to riociguat in the LTE phase (called Placebo-Riociguat in outputs)
- Total of all patients.

Any comparison between the two treatment groups defined above should be done very carefully, since all patients received riociguat in the LTE phase (ie. initial randomization is not preserved). The decision whether or not a patient takes part in the LTE phase may depend on the treatment which was given in the main phase and thus the groups may not be comparable. All results are purely exploratory.

Number of decimal places for summary statistics will be the following:

Figure 6–1 - Decimal places for summary statistics

Statistic	Number of digits
Minimum, maximum	Same as original data
Mean, median	1 more than in original data
SD	2 more than in original data
Frequencies (%)	1 decimal place

6.1 Population characteristics

6.1.1 Demographics and baseline characteristics

Demographic variables and baseline characteristics will be summarized by treatment group and overall for long term safety analysis population.

The following demographic data will be recorded:

- Date of birth (month and year) (age)
- Sex
- Ethnicity

- Smoking status, including number of cigarettes per day
- Alcohol consumption

Summary statistics (number of data available and missing data, mean, standard deviation, minimum, median and maximum) will be presented for continuous variables. Frequency tables will be presented for categorical variables. Smoking history and status will be summarized. Age will be summarized as a continuous variable and as a categorical variable categorized into two groups (< 65 years, ≥ 65 years).

6.1.2 Medical history and concomitant medication

Medical history (MH) and concomitant medication (CM) will be summarized by treatment group and overall.

Medical history findings will be coded by Medical Dictionary for Regulatory Activities (MedDRA) codes. System organ class (SOC) and preferred term (PT) will be used in tabulations.

Medical history findings (i.e., previous diagnoses, diseases or surgeries) meeting all criteria listed below will be collected:

- Pertaining to the study indication
- Started before signing of the informed consent
- Considered relevant to the study
- Medical history related to concomitant therapy

In the following differentiation between medical history and AEs, the term “condition” may include abnormal e.g., physical examination findings, symptoms, diseases, laboratory, ECG.

- Conditions that started before signing of informed consent and for which no symptoms or treatment are present until signing of informed consent are recorded as medical history (e.g., seasonal allergy without acute complaints).
- Conditions that started before signing of informed consent and for which symptoms or treatment are present after signing of informed consent, at unchanged intensity, are recorded as medical history (e.g., allergic pollinosis).
- Conditions that started or deteriorated after signing of informed consent will be documented as AEs.

Prior and concomitant medications will be coded by Anatomical Therapeutic Chemical (ATC) classification system according to the World Health Organization Drug Dictionary (WHO-DD). The ATC class is taken from the WHO-DD code (first character) and the ATC subclass is taken from the WHO-DD code (first 3 characters).

Bayer Drug Groupings (BDG) will be used to select concomitant medications of special interest. The selected concomitant medications will be summarized by parent BDG, BDG and substance name. See the main SAP section 6.1.2 for reference.

6.1.3 Study medication duration and exposure

Study medication duration and exposure will be summarized by treatment group and overall.

The duration of study medication (in days) is derived by calculating the study medication durations by the following formula: period (last dose date – first dose date + 1). Descriptive statistics of treatment duration as well as frequency counts by treatment duration categories will be presented. Dose titration by visit, dose titration sequence and reasons for up- and down-titration by visit and dose will also be summarized using frequency counts from titration phase of LTE period.

The formula for compliance is:

(Total dispensed – total returned/ (Days between visits x dose[# of pills]) X 100

= What was taken / What should've been taken

x 100 = % Compliance

6.2 Efficacy

In the LTE phase, the following endpoints will be described:

- mRSS
- Pulmonary function testing:
  - FVC% predicted
  - FVC (l)
  - DLco% predicted
  - DLco (mmol/min/kPa)
- PROs / HRQoL:
  - Scleroderma Health Assessment Questionnaire (SHAQ)
  - Short Form 36 (SF-36)
  - Patient-Reported Outcomes Measurement Information System (PROMIS)-29
  - University of California, Los Angeles, Scleroderma Clinical Trial Consortium Gastrointestinal Scale (UCLA SCTC GIT) 2.0.
- Worsening of end-organ disease (cardiac, renal, pulmonary, gastrointestinal & digital ischemia)
- Digital ulcer net burden
- Proportion of patients developing new digital ulcers
- Evaluate digital ulcers for patients with and without DUs at baseline by line plots
- All cause mortality
- Composite endpoint will be assessed from Week 52 to the end of the study (in addition, each individual endpoint will be described separately):
  - mRSS progression (defined as an increase in mRSS by > 4 units or >=20%)
  - worsening of FVC (defined as an absolute change of FVC% predicted ≤ -10)

- new organ involvement (as defined in CRISS Step 1 main SAP).

Main study phase and LTE phase data will be shown in summary tables. Change from baseline and change from Week 52 (which is the baseline for the LTE analysis) will be shown. Progression/regression/improvement/worsening rates will be assessed from Week 52 until the end of the study.

6.2.1 mRSS

- Summary statistics for mRSS
- mRSS progression rate (defined as increase in mRSS by > 5 units and ≥ 25%)
- mRSS regression rate (defined as decrease in mRSS by > 5 units and ≥ 25%)
- mRSS progression rate (defined as increase in mRSS by > 4 units or ≥ 20%)
- Percentage of subjects with ≥ 20%, 40%, or 60% improvement in mRSS
- Percentage of subjects with ≥ 20%, 40%, or 60% worsening in mRSS

6.2.2 Pulmonary function testing

- Summary statistics for pulmonary function testing (FVC (forced vital capacity)% predicted, FVC (l), DLco % predicted, DLco (mmol/min/kPa)
- Number of subjects who experience worsening (absolute and relative decline) in FVC% predicted by 15% or more
- Number of subjects who experience worsening (absolute and relative decline) in FVC% predicted by 10% or more

6.2.3 PROs / HRQoL

- SF-36 (Variables: Bodily Pain, General Health, Mental Health, Physical Functioning, Role Emotional, Role Physical, Social Functioning, Vitality, Mental Component Score, Physical Component Score, Mental Health Enhanced score and Health Utility Index) [1.] Norm-based scores in addition to scores 0-100 will be reported.
- SHAQ (VAS variables: pain, intestinal problems, breathing problems, Raynaud's symptoms, finger ulcers and overall severity). All the VAS scores will be converted to 0-3 scale by multiplying results in mm by factor 0.03 and rounding by 0.1 decimals [2.]
- Gastrointestinal involvement as assessed by University of California, Los Angeles, Scleroderma Clinical Trial Consortium Gastrointestinal Scale (UCLA SCTC GIT) 2.0 instrument
- Patient-Reported Outcomes Measurement Information System (PROMIS)-29 (subset of sites belonging to English-speaking countries) [3.]

- Number and proportion of patients with HAQ-DI improvement (ΔHAQ-DI≤-0.21) or non-improvement (ΔHAQ-DI>-0. 21)
- Number and proportion of patients with HAQ-DI improvement (ΔHAQ-DI≤-0.25) or non-improvement (ΔHAQ-DI>-0.25)

6.2.4 Worsening of end-organ diseases

Possible worsened adverse events (cardiac, renal, pulmonary, gastrointestinal & digital ischemia) in the LTE phase identified by certain SOC and PT will be reviewed by an (internal) medical expert applying the definitions of "worsening of end-organ disease" as defined in the Adjudication Committee charter. After evaluation of medical expert, investiagtors will be asked to confirm the finding. Worsened events, which are confirmed by investigators, will be marked in Rave for reporting. In contrast, in the main phase, adjudication of clinical events was done by an (external) adjudication committee. Thus, only events from the LTE phase will be tabulated. Careful interpretation is necessary when comparing the results from the main phase and the LTE phase.

6.2.5 Subgroup analysis

Descriptive analyses of the primary efficacy outcome measure (change in mRSS from baseline to the end of the study and change in mRSS from Week 52 to end of study) will be performed for the following subgroups:

- region (North America, Europe and Australia/New Zealand, East Asia)
- gender (males/females)
- age (age < 65 years/age ≥ 65 years)
- mRSS at baseline (10 - 16 units/17 - 22 units)
- disease duration at baseline (0 - 6 months, 7 - 12 months, 13 - 18 months)
- antibody at baseline:
    - SCL-70, RNA polymerase III; both positive, both negative, either one positive
    - Anti-centromere B; negative (< 10 U/mL) / positive (>= 10 U/mL)
    - Anti-centromere B positive + SCL-70 and RNA polymerase III negative
    - Anti-centromere B positive + either SCL-70 or RNA polymerase III positive
- ILD (interstitial lung disease) per medical history (defined with preferred terms: interstitial lung disease and pulmonaryfibrosis) at baseline (yes/no)
- FVC%, predicted at baseline (<50, 50-75, > 75)
- FVC%, predicted at week 52 (<50, 50-75, > 75)
- hsCRP elevated at baseline (≤ 3.0mg/L, > 3.0 mg/L; and ≤ 10.0 mg/L, > 10.0 mg/L)
- use of corticosteroids at baseline (yes/no)

- tendon friction at baseline (yes/no)
- mRSS at week 52 (<10 units/ 10 - 16 units/ 17 - 22 units/ >22 units)

6.3 Pharmacokinetics/pharmacodynamics

BAY 63-2521 and M1 trough concentrations will be summarized per visit, according to previous dose for trough concentrations. During LTE phase PK was assessed only through visits from V12 to V17. The analyses will be focused on descriptive statistics. The following statistics will be calculated for each of the sampling points: arithmetic mean, standard deviation and coefficient of variation (CV), geometric mean, geometric standard deviation (re-transformed standard deviation of the logarithms), and CV, minimum, median, maximum value and the number of measurements. Boxplot of concentration of BAY 63-2521 by visit will be created.

Means at any time will only be calculated if at least 2/3 of the individual data were measured and were above the limit of quantification (LOQ). For the calculation of the mean value a data point below LOQ will be substituted by one half of the limit.

6.4 Safety

The safety analysis will be performed in the population of patients valid for the long-term safety analysis set. All tabulations will be descriptive only. Data from LTE phase until end of study will be shown. Absolute values and changes from Week 52 (which is the baseline for the LTE phase analysis) will be shown.

Adverse events starting in the main phase and continuing into the LTE phase will be included.

6.4.1 Adverse events and mortality

The incidence of AEs (includes incidence of TEAEs and incidence of post-treatment AEs) will be summarized using MedDRA (version 21.1) preferred terms grouped by primary system organ class. The version number of MedDRA relevant for study evaluation will be stored in the study database.

The incidence of treatment-emergent AEs (TEAEs) occurring during the LTE phase will be tabulated by treatment groups defined for LTE phase. More specifically, AEs are considered to be treatment-emergent if they have started or worsened after first application of study drug up to 2 days after end of treatment with study drug.

An overall summary of the number and percentage of patients with TEAEs will be presented by treatment groups for events from Week 52 to the end of the study. This summary will include the number and the percentage of patients with drug-related TEAEs, treatment emergent serious adverse events (TESAEs), drug-related TESAEs, maximum intensity, AEs leading to permanent discontinuation and AEs with outcome of death.

Incidences of subjects with TEAE**s** occurring during LTE phase will be summarized by treatment arm and MedDRA terms using frequency tables for the following AE types:

- TEAEs
- Drug-related TEAEs
- TESAEs
- Drug-related TESAEs
- TEAEs of special interests
- Drug-related TEAEs of special interest

Tables for maximum intensity of TEAEs, TESAEs and drug-related TEAEs will be provided.

TEAEs which occurred during the main phase and were ongoing (i.e. not reported as "recovered/resolved" or "recovered/resolved with sequelae") at the entry in LTE phase will be summarized.

Incidences of TEAEs and TESAEs per 100 person-years will be tabulated. The rate per 100 person-years is calculated as

Rate per 100 person-years = number of events / (total drug exposure in years / 100),

where 365.25 days are taken as one year.

The incidence of all post-treatment AEs (i.e., AEs occurring more than 2 days after end of treatment with study drug) will be tabulated separately.

AEs with outcome of death will be tabulated from Week 52 to the end of the study.

SAEs, deaths, AE leading to discontinuation and AEs of special interest (as defined in the protocol Section 7.5.1.6) will be listed. The date, relative day (to study medication) will be included.

Further summaries of AEs by intensity and outcome, may be provided, consistent with Bayer Global Medical Standards.

Incidences of AEs of special interest will be tabulated. The protocol Section 7.5.1.6 specifies symptomatic hypotension and serious hemoptysis as AEs of special interest.

6.4.2 Laboratory data

The safety evaluation of laboratory data will include:

- Incidence rates of treatment-emergent laboratory values outside of normal range.
- Incidence rates of pre-specified laboratory data abnormalities.
- Descriptive analysis of continuous laboratory parameters, and their changes from baseline (i.e. Week 52) by visit.
- Categorical analysis of transitions from low, normal, high from baseline (i.e. Week 52) to post baseline.

6.4.3 Other safety parameters

Descriptive analysis of weight and vital signs, pulse oximetry and their changes from Week 52 (which is the baseline for the LTE analysis) until end of the study will be performed. Number of subjects with low/normal/high vital signs, will be tabulated by treatment group and

visit. Assessments for orthostatic test will be listed. The number of subjects with abnormal ECG findings, as well as the number of subjects with treatment-emergent ECG abnormalities will be tabulated by treatment group from Week 52 to the end of the study.

6.4.4 Subgroup analysis

A summary table (number and percentage of patients) of the following groups will be created by treatment group:

- Patients with ILD per MH (defined with PTs: interstitial lung disease and pulmonary fibrosis) by treatment group (already listed as subgroup analysis in section 6.2.5)
- Patients with ILD confirmed by HRCT (confirmed, if interstitial fibrosis and /or ground glass found on HRCT) devided to sub-categories:
  - All patients with ILD per HRCT
  - All patients with ILD per HRCT done pre-therapy (HRCT prior randomization)
  - All patients with ILD per HRCT done post-baseline (HRCT after randomization)
- Patients with ILD per MH and confirmed by HRCT (if a patient with ILD in MH has HRCT done pre-randomization with interstitial fibrosis and/or ground glass found)
- Patients with worsening or new onset of ILD per AE (defined with PTs interstitial lung disease and pulmonary fibrosis in AE data)
- Patients with worsening or new onset of ILD (defined with PTs interstitial lung disease and pulmonary fibrosis in AE data) confirmed by HRCT (Worsening is confirmed, when interstitial fibrosis and/or ground glass finding is defined as worsened on HRCT. New onset is confirmed, when interstitial fibrosis and/or ground glass found on HCRT. HRCT confirms AE in question, if it has been done 28 days before – 28 days after AE start date and fulfills the requirements defined before.)

Descriptive analyses by ILD (yes/no) per MH will be perfomed as follows:

- TEAEs by treatment group
- TESAEs by treatment group
- Lung function tests (FVC%, FVC, DLco% predicted, DLco (mmol/min/kPa),) by treatment group
- Discontinuations of study medication by treatment group

Obligatory HRCTs were performed at baseline or during the study to be able to confirm possible ILD finding in MH or AE.

6.4.5 Possible changes of observations in main phase

If there are corrected observations in AE, MH or CM domains after analysis of main treatment phase, those will be listed per domain by subject.

7. Document history and changes in the planned statistical analysis

- Approval of the SAP version 4.0 dated 22 Dec 2017.
- Approval of the SAP for Benefit-risk update version 1.0 dated 12 Jul 2018.

8. References

1. SF-36v2® Advanced Scoring Guidelines:

   How to submit data, what you will receive back, and scoring timeline Final. Revised: VL, KM, MW, AY, 11/13/2014

2. IMACS Form 04a: Instructions for the Health Assessment Questionnaire
3. PROMIS ADULT PROFILE INSTRUMENTS

9. Appendices

NA