NOVARTIS

Biostatistics & Statistical Programming /
Novartis Institutes for BioMedical Research

VAY736

CVAY736X2207

A subject-, investigator-, and sponsor-blinded, randomized, placebo-controlled, multicenter study to investigate efficacy, safety, and tolerability of VAY736 in patients with idiopathic pulmonary fibrosis

Statistical Analysis Plan (SAP)

Author(s): Personal Protected Data

Document type: SAP Documentation – NIBR

Document status: Final - Amendment 1

Release date: 06 August 2020

Number of pages: 40

Commercially Confidential Information (CCI)

Table of contents

Commercially Confidential Information

Commercially Confidential Information

List of tables

Commercially Confidential Information

1 Introduction

1.1 Scope of document

The RAP documents contain detailed information to aid the production of Statistics & Programming input into the Clinical Study Report (CSR) for trial **"CVAY736X2207".**

The Statistical analysis plan (SAP) describes the implementation of the statistical analysis planned in the protocol. It includes statistical and analytical plans for the analyses described hereafter;

Commercially Confidential Information

1.2 Study reference documentation

This SAP has been developed using Clinical Trial Protocol version v07 CCI

1.3 Study objectives

1.3.1 Primary objectives

Primary Objective	Endpoints
• To assess the efficacy of VAY736 in patients with idiopathic pulmonary fibrosis.	• Change from baseline to end of treatment epoch (48 weeks of treatment) in Forced Vital Capacity (FVC).

1.3.2 Secondary objectives

Secondary Objectives	Endpoints
• To evaluate the impact on VAY736 on safety	• Adverse Events
• To assess the impact of VAY736 on survival	• a) All-cause mortality b) IPF-related mortality
• To assess the impact of VAY736 on "progression free survival (PFS)"	• a) Events, defined as: death (all-cause mortality) OR "progression" (relative reduction in FVC ≥ 10%) b) Events, defined as: death (IPF-related mortality) OR "progression" (relative reduction in FVC ≥ 10%)
• To assess the impact of VAY736 on "disease progression"	• a) Progression, defined as: relative reduction in FVC ≥10% b) Progression, defined as: relative reduction in Diffusing Capacity of the Lungs (DLCO) ≥ 15% c) Progression, defined as: absolute reduction in Six Minute Walk Distance (6MWD) ≥ 50 m
• To assess the impact of VAY736 on a "Composite Endpoint"	• a) Events, defined as: death (all-cause mortality), OR relative reduction in FVC≥10%, OR relative reduction in DLCO ≥15%, OR relative reduction in 6MWD ≥50m b) Events, defined as: death (IPF-related morality), OR relative reduction in FVC≥10%, OR relative reduction in DLCO ≥15%, OR relative reduction in 6MWD ≥50m
• To assess the impact of VAY736 on pulmonary physiology	• a) Change from baseline to the end of treatment epoch (48 weeks of treatment) in DLCO b) Change from baseline to the end of treatment epoch (48 weeks of treatment) in Total Lung Capacity (TLC).

• To assess the impact of VAY736 on exercise capacity	• a) Change from baseline to the end of treatment epoch (48 weeks of treatment) in 6-minute walk distance b) Change from baseline to the end of treatment epoch (48 weeks of treatment) in distance-saturation product c) Change from baseline to the end of treatment epoch (48 weeks of treatment) in post Six Minute Walk Test (6MWT) Borg Scale
• To assess the impact of VAY736 on gas exchange	• Change from baseline to the end of treatment epoch (48 weeks of treatment) in resting oxygen saturation (on room air)
• To assess the immunogenicity of VAY736	• Serum anti-VAY736 antibodies
• To assess the pharmacokinetics of VAY736 after multiple s.c. doses	• Serum VAY736 concentrations

1.3.3 Exploratory objectives

Commercially Confidential Information

Commercially Confidential Information

1.4 Study design and treatment

This is an exploratory (non-confirmatory) randomized, patient-, investigator-, sponsor blinded, and placebo controlled study of VAY736 in IPF patients. This study will investigate the safety and efficacy of CCI VAY736 administered s.c. every 4 weeks for 48 weeks. Approximately 58 subjects will be randomized in a 1:1 ratio on top of local standard of care (SOC), to receive VAY736 or placebo. Randomization will be stratified by SoC medication (3 levels: nintedanib, pirfenidone, or no treatment). The randomization scheme will be performed separately within each stratum to ensure balance of the treatment groups with respect to the SoC medication.

A screening epoch of 6 weeks will be used to assess eligibility.

Randomized subjects will participate in a 48 week treatment epoch

A scheme of the study design is provided in the figure below.

Figure 1-1 Study design

Commercially Confidential Information

2 First interpretable results (FIR)

First interpretable results (FIR) will be provided for this trial.

Commercially Confidential Information

The template shows the analysis / results to be presented in the FIR. Study outputs required to be created at the time of the FIR will be highlighted in TFL shells document and marked as "Key" in the Programming Deliverables Tracker (PDT) output list.

3 Interim analyses

Commercially Confidential Information

Commercially Confidential Information

4 Statistical methods

4.1 Analysis sets

For subjects for which the actual treatment received does not match the randomized treatment, the treatment actually received will be used for the analysis and/or some visits may be excluded.

The **Safety analysis set** will include all subjects that received any study drug.

The **PD analysis set** will include all subjects from safety analysis set with available PD data and no protocol deviations with relevant impact on PD data.

The **PK analysis set** will include all subjects from safety analysis set with at least one available valid (i.e. not flagged for exclusion) PK concentration measurement and with no protocol deviations that impact on PK data.

The **IG analysis set** will include all subjects from safety analysis set with at least one available valid (i.e. not flagged for exclusion) IG concentration measurement and with no protocol deviations that impact on IG data.

Protocol deviations will be identified during the data cleaning process. A comprehensive list of all Protocol Deviations is maintained within the Data Review Plan.

The analysis sets and protocol deviation are related as follows:

Table 4-1 Protocol deviation codes and analysis sets

Commercially Confidential Information

If updates to this table are needed, an amendment to the SAP needs to be implemented prior to DBL.

To understand impact of Covid-19 situation all protocol deviation linked to Covid-19 will be identified and listed separately.

4.2 General presentation of descriptive summaries and listings

The following summary statistics will be presented unless otherwise stated specifically. Qualitative/categorical data (e.g., gender, race) will be summarized by frequency counts and percentages. Percentages will be calculated using the number of patients in the relevant treatment arm or subgroup as the denominator.

Continuous data (e.g., age, body weight) will be summarized using appropriate descriptive statistics (i.e. mean, standard deviation (SD), median, minimum and maximum) by treatment arm.

Based on the nature of the data, few analyses require additional statistics such as geometric mean, geometric coefficient of variation (CV) and it will be specifically requested in respective section.

In general, listings will be presented

- with country/center/subject identifier and age/sex/race as mandatory columns unless otherwise specified
- by treatment, patient and chronological time

4.3 General definitions

Study treatment:
Study treatment will be noted as VAY736/Placebo throughout this document and CSR deliverables.

Study day:
The study day for safety/efficacy assessments (e.g. adverse event onset, laboratory abnormality occurrence, vital sign measurement, Electrocardiogram (ECG), PK sampling, FVC values, etc.,) will be calculated using the randomization date as the reference date. For assessments occurring:
on or after the randomization date, the study day will be calculated as (date of assessment/occurrence) – (date of randomization) + 1. Then study day 1 will be the day of randomization.
before the randomization date, the study day will be calculated as (date of assessment/occurrence) – (date of randomization).
For example, if a serious adverse event starts 1 day before the randomization date, the study day displayed on the listing will be negative, i.e. -1.

Baseline:
For all evaluations, the last available pre-dose assessment before or on the date of randomization is taken as 'baseline' value or 'baseline' assessment.

Unscheduled assessment:
There will be no analysis visit window defined, by consequence descriptive summary statistics by visit will not include unscheduled/repeat assessment.
For all analyses regarding abnormal assessments or analyses based on worst post–baseline value (e.g. laboratory, ECGs, vital signs), all post-baseline values will be included (scheduled, unscheduled, repeat). All unscheduled and repeat measurements will be included in listings.

Treatment analysis epoch:
Treatment epoch will be defined as any post-baseline assessment performed between the first dose of study drug date and the date of Week 49 visit if not missing or the date of last visit attended in treatment epoch (End of Study Treatment visit 199) in case of drop-out or the death date if subject died in treatment epoch.

Commercially Confidential Information

5 Statistical methods for Pharmacokinetic (PK) parameters

5.1 Variables

Commercially Confidential Information

No derivation of PK parameters is planned for this study but a modeling of the data may be performed possibly pooling these data with data coming from previous studies. This optional analysis will be described and reported separately in a modeling plan and a modeling report as appropriate.

5.2 Descriptive analyses

Commercially Confidential Information

Summary statistics will include mean (arithmetic and geometric), Standard Deviation (SD), Coefficient of Variation (CV) (arithmetic and geometric), median, minimum and maximum. At each visit, geometric means and CVs will not be reported in case of concentrations below LLOQ or zero values.

Commercially Confidential Information

Arithmetic CV is defined as: SD*100/Arithmetic mean.

Geometric CV is defined as: 100*sqrt(exp(sdlogs**2)-1), where sdlogs represents the SD of log-e transformed data.

5.3 Statistical model, assumptions and hypotheses

Modeling of the data may be performed as appropriate. As the PK data from the current study may be pooled with data from previous studies, the PK analysis will be described and reported separately in a modeling plan and modeling report.

5.3.1 Model checking procedures

Not applicable.

5.3.2 Graphical presentation of results

Individual concentration-time profiles and mean profiles with SD bars will be presented.

6 Statistical methods for Pharmacodynamic (PD) parameters

6.1 Primary objective

The primary objective of the study is to assess the efficacy of VAY736 on top of current standard-of-care therapy (pirfenidone, nintedanib, or no treatment) using FVC (L) at the end of treatment epoch (48 weeks of treatment) in patients with idiopathic pulmonary fibrosis.

6.1.1 Variables

The primary endpoint is the change from baseline in FVC (L).

6.1.2 Descriptive analyses

Summary statistics of change from baseline in FVC by treatment and visit will be presented by epoch. In addition to the overall summary a summary by baseline standard-of-care medication (stratification factor: pirfenidone, nintedanib, or no treatment) will be produced.

Individual data will be listed by treatment, epoch and visit.

Commercially Confidential Information

6.1.3 Statistical model, assumptions and hypotheses

The following hypothesis will be tested for VAY736+SoC versus Placebo+SoC:

H0: There is no difference or worsening in change from baseline in FVC (L) from baseline to 48 weeks of treatment between VAY736+SoC and placebo+SoC

H1: There is an improvement in change from baseline in FVC (L) from baseline to 48 weeks of treatment between VAY736+SoC and placebo+SoC

Test will be conducted at a one-sided 10% level and 2-sided 80% CI of the estimates will be displayed. Sense of the one-sided test for each criteria is given in table of appendix 1 (section 10.1).

The primary variable, change from baseline in FVC (L) will be analyzed using a Mixed Effects Model for Repeated Measures (MMRM) in the PD analysis set and considering all assessments collected during the treatment epoch. The model will include treatment and visit as fixed effects, standard of care treatment (Nintedanib, Pirfenidone, or no treatment) as factor and baseline value as a covariate, treatment-by-visit and baseline-by-visit as interaction terms. The within subject correlation will be modeled using a repeated statement in SAS Proc Mixed with unstructured covariance matrix. In case of non convergence the model, an alternative AR(1) covariance may be implemented. The posterior will be derived as a t-distribution using a PROC MIXED as detailed in section 3. The main comparison will be a contrast between treatment groups at 48 weeks of treatment. This approach assumes data is Missing At Random (MAR).

6.1.3.1 Model checking procedures

The analyses of residuals will be performed to assess model assumptions.

6.1.3.2 Sensitivity analysis

Missing not at random (MNAR) analysis

Given the severity of disease and the likelihood of some missing data, the analysis will also include approaches to assess treatment under the assumption of missing not at random (MNAR). Missing data will be imputed using methods specific to non-monotone and monotone missing patterns. In a first step, non-monotone missing data will be imputed 1000 times using MCMC method. Then monotone missing data will be imputed by pattern-mixture model (MNAR) performed on the 1000 samples (O'Kelly & Ratitch 2014). A conservative analysis where missing values in the active group are imputed based on the behavior observed in the placebo group (unconditional reference approach) will be used. The objective is to assess if a potential effect observed in the analysis based on MAR assumption is still present when a more conservative procedure is used. The detail of the sensitivity analyses is described in the appendix 2 (section 10.2).

Analysis of potential Covid-19 impact

As primary analysis may exclude some data due to Covid-19, a sensitivity analysis will also include all PD data regardless of protocol deviations linked to Covid-19.

Covid-19 may interfere with study outcomes due to the following:

- Missing visits with missing FVC assessment and/or missing treatment doses due to restriction in travel or closed site. These missing data are considered as missing completely at random therefore introducing no systematic bias in treatment effect.
- Missing visits due to adverse event linked to Covid-19
- Infection by Covid-19.
- Alteration in recruitment after the start of pandemic as some patients might want to avoid participation in the study due to the Covid-19 risk.

The date of when Covid-19 occurred in different countries is tabulated below. Any patient data collected after that date may be confounded by complications due to the pandemic. For example, if a patient has any on-treatment observation collected after the start date for the specific region, that patient's estimated treatment effect is considered potentially affected by the pandemic.

Table 6-1 COVID-19 date

Region/Country	Start Date	End Date
China	01-Jan-2020	End date has not yet been defined
South Korea	20-Feb-2020	End date has not yet been defined
Japan	21-Feb-2020	End date has not yet been defined
Italy	23-Feb-2020	End date has not yet been defined
Rest of the World	01-Mar-2020	End dates have not yet been defined

If sufficient data is present the interaction of covid-19 by treatment will be assessed by means of descriptive statistics:

- Demographics of patients recruited pre- and during COVID period by treatment and for all
- Number of patients with missed visits pre- and during COVID period overall and by treatment and for all
- Adverse events pre- and during COVID period by treatment and for all
- Boxplots of FVC change from baseline pre- and during COVID period by treatment and for all

6.1.3.3 Graphical presentation of results

Plot of model estimates of mean change from baseline and their standard error will be presented over the treatment epoch.

Plot of observed mean change from baseline and their standard error will be presented over the whole study period i.e. double-blind period CCI

6.2 Secondary objectives

6.2.1 Variables

6.2.1.1 Survival (all-cause mortality and IPF-related mortality)

Only events restricted to treatment epoch will be analyzed.

Survival all-cause mortality (S1) is defined as time from randomization to death due to any cause.

Survival IPF related mortality (S2) is defined as time from randomization to death due to IPF related cause.

IPF-related death will be assessed by an Adjudication Committee.

Survival time will be derived in days as follows:

(date of death or censoring date – date of randomization) + 1

For S1, if a patient is not known to have died (all-cause mortality) at the end of the treatment epoch, survival time will be censored at the latest date in treatment epoch where the patient was known to be alive.

For S2, if a patient is not known to have died due to IPF related cause at the end of the treatment epoch, survival time will be censored at the latest date in treatment epoch where the patient was known to be alive or died due to non-IPF related cause.

6.2.1.2 Progression-free Survival

Only events restricted to treatment epoch will be analyzed.

Progression free survival (PFS) is defined as follows:

- PFS1: time from randomization to progression (relative reduction in FVC $\geq$ 10%) or death due to all causes
- PFS2: time from randomization to progression (relative reduction in FVC $\geq$ 10%) or death due to IPF related causes

For PFS1, if a subject has not progressed nor died at the end of the treatment epoch, PFS will be censored at the date of last non-missing assessment of FVC

For PFS2, if a subject has not progressed nor died (IPF related mortality) at the end of the treatment epoch, PFS will be censored at the maximum date of (the last non-missing assessment of FVC during treatment epoch and the non-IPF related death).

6.2.1.3 Disease Progression

Only events restricted to treatment epoch will be analyzed.

Disease progression is defined as the time from randomization to progression.

The following progression events will be considered:

- Progression, defined as: relative reduction in FVC $\geq$ 10%;

- Progression, defined as: relative reduction in DLCO ≥ 15%;
- Progression, defined as: absolute reduction in 6MWD ≥ 50 m.

For the subjects who have not progressed at the end of the treatment epoch, disease progression will be censored at the date of the last FVC, DLCO or 6MWD assessment in treatment epoch as applicable.

6.2.1.4 Composite endpoint

Only events restricted to treatment epoch will be analyzed.

Time to composite events will be considered, where events of interest are defined as below:

- Death (all-cause mortality), OR relative reduction in FVC ≥ 10%, OR relative reduction in DLCO ≥ 15%, OR relative reduction in 6MWD ≥ 50 m;
- Death (IPF-related mortality), OR relative reduction in FVC ≥ 10%, OR relative reduction in DLCO ≥ 15%, OR relative reduction in 6MWD ≥ 50 m.

For the subjects who do not experience the events at the end of the treatment epoch, the time to event will be censored at the latest date of the FVC, DLCO or 6MWD assessments in the treatment epoch.

6.2.1.5 Pulmonary Physiology

Change from baseline in DLCO and in TLC during the treatment epoch

6.2.1.6 Exercise capacity

Change from baseline in 6-minute walking distance (6MWD), in distance-saturation product (DSP) and in post 6-minute walking test (6MWT) Borg dyspnea score during the treatment epoch.

Note: DSP = 6MWD * end of test oxygen saturation.

6.2.1.7 Gas Exchange

Change from baseline in resting oxygen saturation during treatment epoch

6.2.2 Descriptive analyses

A list and a summary of the events (Survival, Progression-free Survival, Disease Progression and Composite Endpoint) within and outside the treatment epoch (as applicable) will be

provided. In the analyses on composite events a summary of the single events (the one occurring first) will be provided too.

Summary statistics of change from baseline by treatment and visit will be presented for DLCO and TLC, 6-minute walking distance, in distance-saturation product and in post 6MWT, resting oxygen saturation

Commercially Confidential Information

6.2.3 Statistical model, assumptions and hypotheses

6.2.3.1 Time to Events analyses CCI

The Kaplan-Meier estimates of the proportion of patients with the event of interest in each treatment group, along with 80 % two-sided confidence intervals using Greenwood's formula, will be provided. The two treatment groups will be compared using a one-sided log rank test at the 10% significance level. The hazard ratio and its associated 80% two-sided confidence interval will be estimated based on a Cox regression model using treatment group as main factor.

6.2.3.2 Pulmonary Physiology, Exercise capacity, Gas Exchange CCI

The same MMRM model described for the primary endpoint will be used in these analyses using SAS Proc Mixed.

Commercially Confidential Information An ANCOVA model will be fitted with treatment group as main factor, standard of care therapy (stratification factor) as a cofactor and baseline as covariate.

6.2.3.3 Model checking procedures

For the Time to Events Analyses, the assumption of proportional hazard will be visually assessed reviewing the Kaplan-Meier curves.

For the other analyses, the analyses of residuals will be performed to assess model assumptions.

6.2.3.4 Graphical presentation of results

6.2.3.4.1 Time to Events Analyses

Kaplan-Meier curves by treatment will be plotted.

6.2.3.4.2 Other analyses

Plots of mean values (with error bars representing SD) by treatment and visit will be released.

6.3 Exploratory objectives

Commercially Confidential Information

Commercially Confidential Information

6.3.2 Descriptive analyses

Commercially Confidential Information

7 Statistical methods for safety and tolerability data

7.1 Variables

Adverse events, vital signs (blood pressure, pulse rate, body temperature), ECG intervals, laboratory measurements, as well as subject demographics, baseline characteristics, and treatment information.

7.2 Descriptive analyses

Subject demographics and other baseline characteristics

All data for background and demographic variables will be listed by treatment group and subject. Summary statistics will be provided by treatment group.

Treatment

Data for study drug administration (rescue medication) and concomitant therapies will be listed by treatment group and subject.

Summary statistics will be provided for treatment compliance by treatment group.

The use of the standard of care medication (pirfenidone, nintedanib, or no treatment) will be summarized by treatment group.

Vital signs

All vital signs data (including pulse rate, systolic and diastolic blood pressure in sitting position, temperature and oxygen saturation) will be listed by treatment, subject, and visit and abnormalities as described below will be flagged.

Table 7-1 Notable vital signs and body weight

Vital Sign Variable	Notable Criteria

Pulse (beats/min)	>120 bpm or < 50 bpm
Systolic BP (mmHg)	≥ 160 mmHg or ≤ 90 mmHg
Diastolic BP (mmHg)	≥ 100 mmHg or ≤ 50 mmHg
Temperature (°C)	>38.3 °C/ 101°F
Body weight (kg)	More than 7% decrease from baseline weight

ECG evaluations

All ECG data (heart rate, QRS duration, RR, PR, QT and QTcF intervals) will be listed by treatment, subject and visit, and abnormalities as described below will be flagged. Summary statistics of value will be provided by treatment and visit. CCI

Table 7-2 ECG abnormalities

ECG abnormalities Variable	Notable Criteria
QT (ms)	Increase ≥ 30 ms
	Increase ≥ 60 ms
	New > 450 ms
	New > 480 ms
	New > 500 ms
QTcF (ms)	Increase ≥ 30 ms
	Increase ≥ 60 ms
	New > 450 ms
	New > 480 ms
	New > 500 ms
PR (ms)	increase > 25% or new > 200 ms
	New > 200 and ≤ 220 ms
	New > 220 ms
QRS (ms)	increase >25% or new > 110 ms
	New > 110 and ≤ 120 ms
	New > 120 ms
RR (ms) / HR (bpm)	RR increase > 25% or HR > 100 bpm
	RR decrease > 25% or HR < 60 bpm

Clinical laboratory evaluations

All laboratory data will be listed by treatment, subject, and visit and if normal ranges are available abnormalities will be flagged. Summary statistics will be provided by treatment and visit. Newly occurring liver enzyme abnormalities described in Table 7-3 will be summarized in treatment epoch. Newly occurring abnormality is defined when the abnormality occurs post baseline during the treatment epoch for patient without this abnormality at baseline .

Table 7-3 Liver enzyme abnormalities

Definition
ALT > 5x ULN

Definition
ALT > 10x ULN ALT > 20x ULN
ALT or AST > 3x ULN ALT or AST > 5x ULN ALT or AST > 8x ULN
ALT or AST > 3x ULN & TBL > 1.5x ULN ALT or AST > 3x ULN & TBL > 2x ULN
ALP > 1.5x ULN ALP > 2x ULN
AST = Aspartate aminotransferase; also known as SGOT, ALT = Alanine aminotransferase; also known as SGPT, ALP = Alkaline phosphatase, TBL = Total bilirubin,

Listing and summary tables will be provided for the final analysis CCI

Adverse events

All information obtained on adverse events will be displayed by treatment and subject.

Commercially Confidential Information

The number and percentage of subjects with adverse events occurring after the first dosing will be tabulated by body system and preferred term with a breakdown by treatment. A subject with multiple adverse events within a body system is only counted once towards the total of this body system and treatment. The same summary will be done for serious event.

The number and percentage of subjects with adverse events occurring after the first dosing by maximum severity of adverse events will be tabulated by body system and preferred term with a breakdown by treatment.

The number and percentage of subjects with adverse events occurring after the first dosing classified as related to study drug will be tabulated by body system and preferred term with a breakdown by treatment. The same summary will be done for serious event.

An overall summary of AEs will present incidence and occurrence of all AEs, AEs of mild intensity, AEs of moderate intensity, AEs of severe intensity, study drug-related AEs, serious AEs and AEs leading to discontinuation of study treatment.

If for a same patient, several consecutive AEs (irrespective of study treatment causality, seriousness and severity) occurred with the same SOC and PT:

- a single occurrence will be counted if there is ≤ 1 day gap between the end date of the preceding AE and the start date of the consecutive AE
- more than one occurrence will be counted if there is > 1 day gap between the end date of the preceding AE and the start date of the consecutive AE.

The number of deaths resulting from SAEs suspected to be related to study treatment and SAEs irrespective of study treatment relationship will be provided by SOC and PT.

Listing and summary tables, by epoch, will be provided for the final analysis

Commercially Confidential Information

7.3 Graphical presentation

Boxplots to visualize trends in longitudinal safety data (vitals, ECG, lab parameter) will be created.

Commercially Confidential Information

Commercially Confidential Information

Commercially Confidential Information

9 Reference list

Carpenter JR, Roger JH, Kenward MG (2013) Analysis of longitudinal trials with protocol deviation: a framework for relevant, accessible assumptions and inference via multiple imputation. *Journal of Biopharmaceutical Statistics*, 23(6):1352–1371.

Ratitch, B. and O'Kelly, M. (2011), Implementation of Pattern-Mixture Models Using Standard SAS/STAT Procedures, *in Proceedings of PharmaSUG 2011 (Pharmaceutical Industry SAS Users Group)*, SP04, Nashville.

European Medicines Agency (2010), Guideline on Missing Data in Confirmatory Clinical Trials.

O'Kelly, M., and Ratitch, B. (2014), Clinical Trials with Missing Data: A Guide for Practitioners, *West Sussex: Wiley.[https://doi.org/10.1002/9781118762516.fmatter]*.

Rubin, D. B. (1976), "Inference and Missing Data," Biometrika, 63, 581–592.

Rubin, D. B. (1987), Multiple Imputation for Nonresponse in Surveys, New York: John Wiley & Sons.

Patel AS, Siegert RJ, Brignall K, Gordon P, Steer S, Desai SR, Maher TM, Renzoni EA, Wells AU, Higginson IJ, Birring SS. The development and validation of the King's Brief Interstitial Lung Disease (K-BILD) health status questionnaire. Thorax 2012; 67 (9), 804-810 [R12-4171]

St George's Respiratory Questionnaire Manual, *University of London 2009*

A Composite Physiologic Index Derived from Disease Extent Observed by Computed Tomography, *ATS Journals Vol 167. pp 962–969, 2003.*

10 Appendix

10.1 Appendix 1: Definition of benefit for individual endpoints

Measurement	Endpoint	Direction of benefit*
Spirometry	FVC	Increase
Lung Volumes	TLC	Increase
Diffusion lung capacity	DLCO	Increase
	Commercially Confidential Information	
6MWT	6MWD	Increase
	Borg dyspnea score	Decrease
	Commercially Confidential Information	

* In the direction of

10.2 Appendix 2: Implementation of the sensitivity analyses

This appendix is adapted from Ratitch, B. and O'Kelly, M. (2011).

SAS PROC MI permits to perform control-based imputation. In this appendix, we will assume that data are stored in a SAS dataset, DIN, with the following variables:

- SUBJID – subject identification number
- SOC_TYPE - stratification factor (with 0 representing no background therapy, 1 nintedanib, 2 pirfenidone)
- SOC_T1 - stratification factor (with 1 representing nintedanib and 0 other cases)
- SOC_T2 - stratification factor (with 1 representing pirfenidone and 0 other cases)
- DRUGNUM - treatment arm (with 0 representing control arm, and 1 representing the Active group)
- BASE, W5, W9-W49, corresponding to the criteria at BASE(baseline), and post-baseline visits W5 (week 5), W9 (week 9) up to W49 (week 49) by 4 weeks
- LASTVIS – last study visit (takes values 0 through 49)

A- General MAR-based multiple imputation

If a dataset has some subjects with non-monotone missingness, the most frequently used imputation method in PROC MI is based on the Markov Chain Monte Carlo methodology and assumes a multivariate normal distribution over all variables included in the imputation model. For datasets with monotone missingness, a number of different methods are available for continuous as well as categorical variables, including some parametric and non-parametric methods. Fortunately, standard software facilitates using the MCMC method to impute data only partially – just enough to obtain a dataset with monotone missingness. Then the remaining missing items can be imputed using other available methods if deemed more suitable. This two-step procedure is the one we use, and is illustrated below using the example dataset described in Section 3 and above in this appendix.

In PROC MI the MCMC method is used to partially impute non-monotone data with the `impute=monotone` option to obtain a dataset with monotone missingness.

```
proc mi data=DIN out=DIN_MONO nimpute=1000 seed=123181;
   var DRUGNUM SOC_T1 SOC_T2 BASE W5 W9-W49;
   mcmc chain=multiple impute=monotone nbiter=200 niter=100;
run;
```

Note that the SOC type factor is dichotomized in 2 continuous variables as in this first call only numeric variable can be used.

Values of the MCMC parameters NBITER (the number of MCMC updates allowed for the sampler to converge to the posterior distribution, and NITER (the number of iterations between each saved draw from the posterior) may need to be increased to allow the MCMC estimates stabilize.

Assumptions underlying this partial imputation step would be of the MAR type, i.e., that subjects with intermediate missing data follow the same model as other subjects in their respective treatment arm that continue the study. This is a reasonable assumption for this partial imputation process because subjects tend to miss intermediate visits due to scheduling conflicts or other reasons generally unrelated to their medical condition under study.

If the proportion of non-monotone missing data is small compared to the number of subjects that discontinue from the study permanently, imputing these few non-monotone values using the partial MCMC method under MAR assumptions would have a small impact on the overall analyses. Note that we will use 1000 imputations. This number is much larger than the default in most software packages. This high number of imputations decrease at the best the uncertainty of the imputation and using as many as 1000 or so imputations does not present any computational difficulties with the current technology.

The output data set DIN_MONO from the above call to PROC MI can now be used as input to a subsequent call of PROC MI to impute the remaining monotone missing values. Since dataset DIN_MONO already contains 1000 copies of the data with partial imputations, the subsequent call to PROC MI will use a `BY _Imputation_` statement and request one imputed dataset (`nimpute=1` option) within each BY group. This second call will be based on a Not Missing At Random (NMAR) approach as described in B part.

B- NMAR data imputation

Reference-based pattern imputation (unconditional reference)

Assuming that non-monotone missing data have already been imputed using the MCMC method as described above we will then break the imputation process into a sequence of MIs, where each MI is intended to impute missing values at one time-point only, as follows.

i. In order to impute missing data at time-point 1 (W5) and based on the definition of the link function, first prepare an input dataset, making sure that it will contain only the intended donor pattern (P1: control subjects with non-missing values) and recipient patterns (P2: control arms with missing data and P4 active group with missing data). Separate the input dataset DIN_MONO into two datasets: DIN_MONO_IMP1, containing all control subjects and those subjects from the active group that have values at W4 missing (patterns P1, P2 and P4); and DIN_MONO_REST1, containing the subjects from the active group with non-missing values (pattern P4.

```
data DIN_MONO_IMP1  DIN_MONO_REST1;
   set DIN_MONO;
  *** subjects from the VAY736 non-missing values at W5 ***;
   if DRUGNUM=1 and LASTVIS>=5 then output DIN_MONO_REST1;
  ***All other included VAY736 with missing W5 ***;
   else output DIN_MONO_IMP1;
run;
```

i. Call PROC MI to impute missing data at time-point 1 using dataset DIN_MONO_IMP1 as input.

```
proc mi data= DIN_MONO_IMP1 out= DIN_REG_IMP1 nimpute=1
seed=234;
   by _Imputation_;
   var SOC_TYPE BASE W5;
   monotone regression (W5 = SOC_TYPE BASE);
run;
```

Note that treatment arm is not included as an effect in the model (it is not included in the VAR statement). Since subjects from the active group with non-missing values at time-point t are not included in the input dataset, they will not contribute to the estimation of an imputation model for time-point t. The imputation model will be estimated using control subjects only, while this call to PROC MI will impute missing data at time-point t for all subjects who need imputation at that time-point. This way, subjects from active group will be imputed based on the control subjects' model.

ii. Assemble a dataset containing all subjects while including the just-imputed visit t data so that they can serve as predictors for the imputation of the next visit.

```
data DIN_IMP1;
   set DIN_REST1 DIN_REG_IMP1;
run;
```

iii. Repeat steps (i)-(iii) for all other time-points sequentially using a reconstructed dataset (e.g., DIN_IMP1) as a starting point in step (i).

C- Modelling on M complete datasets

Once all missing data has been imputed then an ANCOVA of the change from baseline of FVC (CHANGE) at week 49 visit will be applied by _Imputation_ as per model below:

```
proc mixed data=DIN_FULLIMP;
     by _Imputation;
     where VISIT='Week 49'
     class DRGNUM SOC_TYPE;
     model CHANGE = DRGNUM SOC_TYPE BASE / s ddfm=kr;

     lsmeans DRGNUM / diff=control('0') cl;
     ods output Diffs=DIFF_MI LSMeans=LSM_MI;
quit;
```

D- Pooling analysis

Once the result by Imputation are computed then the parameter estimates will be combined as per Rubin (1976, 1987) method using PROC MIANALYZE.

```
proc mianalyze parms=DIFF_MI;
     modeleffects DRGNUM;
     stderr STDERR;
     ods output parameterestimates=MiDIFF;
run;
```

Commercially Confidential Information

Commercially Confidential Information

Commercially Confidential Information

Signatures

Personal Protected Data