Document Type:	Statistical Analysis Plan Amendment 3	**Protocol/Study No.:**	RIN-PF-301
Finalization Date:	25 AUG 2025	**Total Number of Pages Including Appendices:**	55
Classification:	Confidential		

Title:

A Randomized, Double-blind, Placebo-controlled, Phase 3 Study of the Efficacy and Safety of Inhaled Treprostinil in Subjects with Idiopathic Pulmonary Fibrosis

Author: [redacted]

CONFIDENTIAL AND PROPRIETARY, UNITED THERAPEUTICS CORPORATION
All content contained herein is confidential and proprietary information of United Therapeutics Corporation and shall not be disclosed in whole or in part except as permitted by a signed contract with United Therapeutics Corporation.
© 2025 United Therapeutics Corporation

TABLE OF CONTENTS

TABLE OF CONTENTS 2
Table of In-Text Tables 3
Table of In-Text Figures 4
ABBREVIATIONS AND DEFINITIONS 5
1 PREFACE 6
2 STUDY ESTIMAND/OBJECTIVE AND ENDPOINTS 6
2.1 PRIMARY OBJECTIVE, ENDPOINT, AND ESTIMAND 6
2.2 SECONDARY OBJECTIVES, ENDPOINTS, AND ESTIMANDS 8
2.3 OTHER ENDPOINTS 9
3 STUDY DESIGN 10
4 RANDOMIZATION 11
5 STRATIFICATION 11
6 SEQUENCE OF PLANNED ANALYSES 11
7 SAMPLE SIZE CONSIDERATIONS 12
8 ANALYSIS POPULATIONS 12
9 GENERAL CONSIDERATIONS FOR DATA ANALYSES 13
9.1 COVARIATES 13
9.2 EXAMINATION OF SUBGROUPS 14
9.3 PREMATURE DISCONTINUATION AND MISSING DATA 14
9.4 MULTIPLE COMPARISONS AND MULTIPLICITY 15
9.5 DERIVED AND TRANSFORMED DATA 15
9.5.1 Spirometry 15
9.5.2 Lung Diffusion Capacity 17
9.5.3 Clinical Worsening 17
9.5.4 Acute Exacerbations of IPF 19
9.5.5 Overall Survival at Week 52 20
9.6 ASSESSMENT WINDOWS 20
9.6.1 Multiple Evaluations Within the Same Analysis Window 21
10 STUDY POPULATION 22
10.1 PATIENT ACCOUNTABILITY 22
10.2 PROTOCOL DEVIATIONS 23
10.3 OTHER DESCRIPTIONS OF STUDY POPULATION 23
10.3.1 Demographics 23
10.3.2 Medical History and IPF History 23
10.3.3 Concomitant Medications 24
11 EFFICACY ANALYSES 24
11.1 PRIMARY EFFICACY MEASURE 25
11.1.1 Primary Estimand 25
11.1.2 Primary Efficacy Analysis 26

11.1.3 Sensitivity and Subgroup Analyses....29
11.2 SECONDARY EFFICACY MEASURES....32
11.2.1 Time to Clinical Worsening....35
11.2.2 Time to Acute Exacerbation of IPF....37
11.2.3 Overall Survival at Week 52....38
11.2.4 Change in % Predicted FVC at Week 52....38
11.2.5 Change in K-BILD Total Score and Domain Scores at Week 52....39
11.2.6 Change from Baseline in DLCO at Week 52....40
11.3 EXPLORATORY EFFICACY MEASURES....40
11.3.1 Change from Baseline in Absolute FVC at Weeks 16, 28, and 40....41
11.3.2 Change from Baseline in NT-proBNP at Week 52....41
11.3.3 Change from Baseline in Resting Supplemental Oxygen at Week 52....41
11.3.4 Evaluation of Pharmacogenomics and Biomarkers....42
12 SAFETY ANALYSES....42
12.1 EXTENT OF EXPOSURE....42
12.2 ADVERSE EVENTS....42
12.3 CLINICAL LABORATORY PARAMETERS....43
12.3.1 Clinical Chemistry....44
12.3.2 Hematology....44
12.4 VITAL SIGNS....45
12.5 ELECTROCARDIOGRAMS....45
13 REFERENCES....46
14 APPENDICES....47
14.1 LIST OF TABLES....47
14.2 LIST OF LISTINGS....50
14.3 LIST OF FIGURES....51
15 SUMMARY OF CHANGES FROM PREVIOUS VERSIONS....52

Table of In-Text Tables

Table 2-1 Estimands for Secondary Endpoints....8
Table 9-1 Grading System for FEV1 and FVC (Graded Separately)....16
Table 9-2 Derivation of Time to First Clinical Worsening....18
Table 9-3 Derivation of Time to First Respiratory Hospitalization....19
Table 9-4 Derivation of Time to First Acute Exacerbation of IPF....19
Table 9-5 Derivation of Time to Death (Overall Survival)....20
Table 9-6 Assessment Windows for Scheduled Visits....21
Table 11-1 Derivation of Time to Death (when death is the first clinical worsening event)....34

Table of In-Text Figures

Figure 3-1 RIN-PF-301 Study Design....10
Figure 11-1 Other Efficacy Endpoints....34

ABBREVIATIONS AND DEFINITIONS

AE	Adverse event
ATC	Anatomical Therapeutic Chemical
BMI	Body mass index
CEC	Clinical Endpoint Committee
CI	Confidence interval
DLCO	Diffusion capacity of lungs for carbon monoxide
DMC	Data Monitoring Committee
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
FEV_1	Forced expiratory volume in 1 second
FVC	Forced vital capacity
IPF	Idiopathic pulmonary fibrosis
ITT	Intent-to-Treat
K-BILD	King's Brief Interstitial Lung Disease Questionnaire
K-M	Kaplan-Meier
MMRM	Mixed-effect model for repeated measures
NT-proBNP	N-terminal pro-brain natriuretic peptide
OSA	Obstructive sleep apnea
PT	Preferred Term
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SOC	System Organ Class

1 PREFACE

This Statistical Analysis Plan (SAP) provides further details of the planned analyses for the RIN-PF-301 study as presented in the study protocol. This plan, based on the original RIN-PF-301 study protocol dated 14 October 2020 and subsequent study protocol amendments (the latest version study protocol amendment 4 dated 26 July 2023), provides further details of the planned analyses stated in the study protocol as well as any additional planned analyses. Additional post hoc or unplanned analyses that are not defined in this SAP may be performed. Such analyses will be documented in the Clinical Study Report.

2 STUDY ESTIMAND/OBJECTIVE AND ENDPOINTS

2.1 PRIMARY OBJECTIVE, ENDPOINT, AND ESTIMAND

RIN-PF-301 will compare inhaled treprostinil with placebo in subjects with idiopathic pulmonary fibrosis (IPF), with randomization stratified by IPF background therapy (nintedanib or pirfenidone vs. no background therapy). The primary objective of RIN-PF-301 is to evaluate the superiority of inhaled treprostinil against placebo for the change in absolute forced vital capacity (FVC) from baseline to Week 52. The primary endpoint is the change in absolute FVC in subjects with IPF from baseline to Week 52. The treatment effect of primary interest is for randomized subjects who received at least 1 study dose, regardless of the extent of treatment adherence or early study treatment discontinuation but excluding the effects of elective lung transplantation and initiation of commercial inhaled treprostinil products. All available data, including post treatment discontinuation data for subjects who prematurely discontinued treatment, will be used for the primary estimand.

Following the International Council for Harmonisation (ICH) E9 (R1) guidance, the estimand for the primary efficacy endpoint contains the following attributes:

- Population: Subjects ≥40 years of age with IPF
- Endpoint: Change in absolute FVC from baseline at Week 52
- Treatment condition: Inhaled Treprostinil or placebo on top of IPF background therapy or no IPF background therapy
- Intercurrent events (3 types of intercurrent events):
  - Death: a composite strategy will be implemented, in which the occurrence of death is incorporated into the definition of the endpoint.

  - Early study treatment discontinuation, initiation or change of background therapy during the study, initiation of rescue medication, and initiation of prostacyclin therapy (other than commercial inhaled treprostinil products): a treatment policy strategy will be implemented, the values for the variable of interest after intercurrent events will be included.
  - Elective lung transplantation and initiation of commercial inhaled treprostinil products: a hypothetical strategy will be implemented.
- Population-level summary: Difference in mean changes in absolute FVC at Week 52 between inhaled treprostinil and placebo.

Subjects who discontinue study treatment prematurely are encouraged to remain in the study and continue their clinical visits until Week 52. The reasons for treatment discontinuation will be collected and treatment discontinuation will be considered an intercurrent event. The treatment policy will be the primary strategy for handling the intercurrent events of early treatment discontinuation, initiation or change of IPF background therapy during the study, initiation of rescue medication, and initiation of prostacyclin therapy (excluding commercial inhaled treprostinil products). Initiation of rescue medication includes chronic concomitant medication (azathioprine, cyclosporine, mycophenolate mofetil, tacrolimus, oral corticosteroids >20 mg/day, or the combination of oral corticosteroids + azathioprine + N-acetylcysteine; cyclophosphamide; rituximab, endothelin receptor antagonists, phosphodiesterase type 5 inhibitors, or soluble guanylate cyclase stimulators) initiated during the study when deemed as safe and clinically necessary by the investigators. Initiation of prostacyclin therapy includes epoprostenol, treprostinil (except for commercial inhaled treprostinil products), iloprost, or beraprost and IP receptor agonists (selexipag). With the treatment policy strategy, the values for the variable of interest after intercurrent events will be collected and included in the primary analysis.

Composite variable strategy will be used for the intercurrent event of death. For subjects who died before Week 52, their missing FVC change from baseline values at scheduled visits on or after the death date will be assigned the 2.5th percentile of all observed values (FVC change from baseline) across both treatment groups at each visit.

Hypothetical strategy will be used for the intercurrent events of elective lung transplant and initiation of commercial inhaled treprostinil products. Values observed after the initiation of commercial inhaled treprostinil products during the study will not be included. For this study, it

is known that some patients are already in the lung transplant waiting list. The occurrence of a lung transplant is often driven by availability of donor lungs, rather than a specific event or disease progression which triggers the transplant. Thus, lung transplantation will be considered an elective procedure. Subjects who undergo a lung transplant prior to Week 52 will be discontinued from the study and the value of the variable after the lung transplant will not be measured or recorded.

After applying different strategies for intercurrent events as described above, any still missing or unmeasurable values at visits before (or at) Week 52 will be imputed appropriately in the primary analysis.

2.2 SECONDARY OBJECTIVES, ENDPOINTS, AND ESTIMANDS

The secondary endpoints of this study evaluate the effect of inhaled treprostinil on the following parameters in subjects with IPF:

- Time to first clinical worsening (including time to death, respiratory hospitalization, or ≥10% relative decline in % predicted FVC)
- Time to first acute exacerbation of IPF
- Overall survival at Week 52
- Change from baseline in % predicted FVC at Week 52
- Change from baseline in King's Brief Interstitial Lung Disease Questionnaire (K-BILD) score at Week 52
- Change from baseline in diffusion capacity of lungs for carbon monoxide (DLCO) at Week 52

The key attributes of estimands for the secondary endpoints are outlined in Table 2-1.

Table 2-1 Estimands for Secondary Endpoints

Endpoint	Population-level Summary	Intercurrent Event Strategies
Time to first clinical worsening event	Hazard ratio (inhaled treprostinil vs. placebo)	Treatment policy (for early study treatment discontinuation, initiation or change of background therapy during the study, initiation of rescue medication, and initiation of prostacyclin therapy – other than commercial inhaled treprostinil products) and hypothetical strategy (for elective lung transplantation

Endpoint	Population-level Summary	Intercurrent Event Strategies
		and initiation of commercial inhaled treprostinil products)
Time to first acute exacerbation of IPF	Hazard ratio (inhaled treprostinil vs. placebo)	Same as for time to first clinical worsening event
Overall survival at Week 52	Hazard ratio (inhaled treprostinil vs. placebo)	Same as for time to first clinical worsening event
Change from baseline in % predicted FVC at Week 52	Difference in mean changes in % predicted FVC at Week 52	Same as for primary endpoint of change in absolute FVC from baseline to Week 52
Change from baseline in K-BILD score at Week 52	Difference in mean changes in K-BILD score at Week 52	Same as for primary endpoint of change in absolute FVC from baseline to Week 52
Change from baseline in DLCO at Week 52	Difference in mean changes in DLCO at Week 52	Same as for primary endpoint of change in absolute FVC from baseline to Week 52

Abbreviations: DLCO, diffusion capacity of lungs for carbon monoxide; FVC, forced vital capacity; IPF, idiopathic pulmonary fibrosis; K-BLD, King's Brief Interstitial Lung Disease Questionnaire

The time to event endpoints, as described in Table 2-1, will be censored at the end of study date, elective lung transplantation date, or initiation of commercial inhaled treprostinil products date, whichever is earliest. Intercurrent event strategies for the continuous secondary endpoint (ie, change from baseline in % predicted FVC at Week 52, change from baseline in K-BILD score at Week 52, and change from baseline in DLCO at Week 52) will follow the same intercurrent event strategies as the primary endpoint. That is, to apply treatment policy for early study treatment discontinuation, initiation or change of IPF background therapy during the study, initiation of rescue medication and initiation of prostacyclin therapy (other than commercial inhaled treprostinil products); composite strategy for death; and hypothetical strategy for elective lung transplantation and initiation of commercial inhaled treprostinil products. For subjects who died before Week 52, their missing change from baseline values for % predicted FVC, K-BILD, and DLCO at scheduled visits on or after the death date will be assigned the 2.5th percentile of all observed values (change from baseline) across both treatment groups at each respective visit.

2.3 OTHER ENDPOINTS

The exploratory endpoints of this study evaluate the effect of inhaled treprostinil on the following parameters in subjects with IPF:

- Change from baseline in absolute FVC at Weeks 16, 28, and 40

- Change from baseline in N-terminal pro-brain natriuretic peptide (NT-proBNP) at Week 52
- Change from baseline in resting supplemental oxygen use at Week 52
- Evaluation of pharmacogenomics and biomarkers for subjects who consent to the sample collection

Safety endpoints will be assessed by reviewing the following parameters:

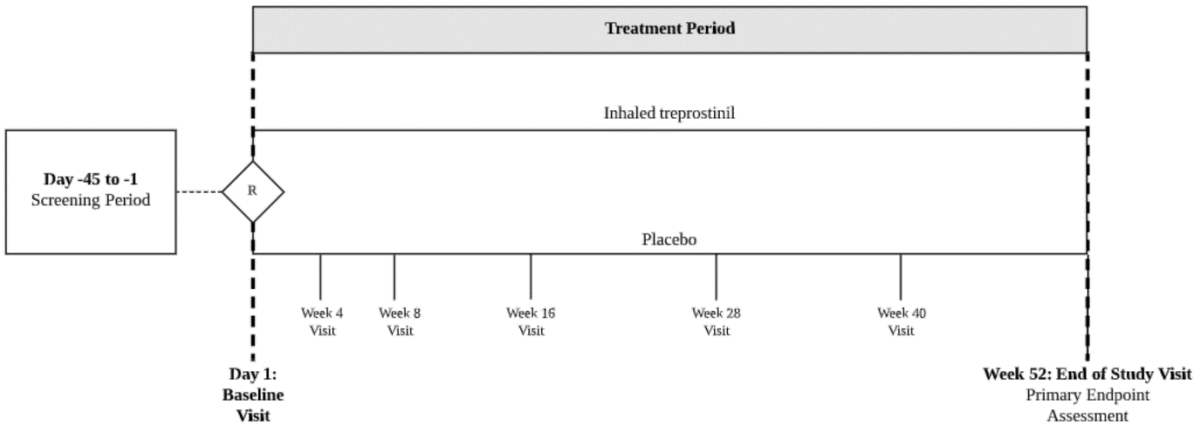
- Adverse events (AEs) and serious adverse events (SAEs)
- Clinical laboratory parameters
- Vital signs, including resting saturation of peripheral capillary oxygenation
- 12-lead electrocardiograms (ECGs)

3 STUDY DESIGN

This is a Phase 3, randomized, double-blind, placebo-controlled, efficacy and safety study of subjects with IPF treated with inhaled treprostinil over a 52-week period (Figure 3-1).

Approximately 576 eligible subjects will be randomly assigned 1:1 to inhaled treprostinil or placebo at Baseline. Randomization will be stratified by IPF background therapy (nintedanib or pirfenidone vs no background therapy). All subjects will initiate inhaled treprostinil (6 mcg/breath) or placebo at a dose of 3 breaths (18 mcg) administered 4 times daily (during waking hours) and will titrate to a target dosing regimen of 12 breaths (72 mcg) 4 times daily.

Figure 3-1 RIN-PF-301 Study Design



Abbreviation: R, randomization

4 RANDOMIZATION

Subjects will be randomized 1:1 to inhaled treprostinil (6 mcg/breath) or placebo. An interactive voice/web response system will be utilized for randomization. Randomization will occur using a centrally administered stratified permuted block randomization, stratified by IPF background therapy (nintedanib or pirfenidone vs no background therapy). Confirmation that a subject qualifies based on Screening spirometry, high-resolution computed tomography parameters, and surgical lung biopsy, if available, from the central reviewer, must be received before randomization occurs.

5 STRATIFICATION

Subjects will be stratified by IPF background therapy (nintedanib or pirfenidone versus no background therapy). If incorrect stratification information is used for randomization, the actual stratification information will be used for analyses.

6 SEQUENCE OF PLANNED ANALYSES

Interim Safety Analyses

Interim safety analyses are intended to be performed according to the Data Monitoring Committee (DMC) Charter. In particular, interim safety analyses are planned following the enrollment of approximately 50, 175, 500, 700, and 900 subjects across both the RIN-PF-301 (TETON-1) and RIN-PF-303 (TETON-2) studies, or at least annually (whichever is sooner). All analyses will be prepared by an independent external consultant and reviewed only by the DMC and external biostatistician as defined in the DMC Charter. The Sponsor will only have access to the blinded study data during this process. Interim efficacy analyses are not planned in this study.

Dry-run Analysis Prior to Database Lock (Soft Lock)

At the completion of the study enrollment and prior to the database lock, a dry-run analysis is planned. This dry-run analysis utilizes a dummy randomization schedule in a blinded fashion. The purpose of the dry-run analysis is to verify programming of the analysis data sets and the planned tables, listings, and figures, and to identify any data issues. Programming and data issues identified through the dry-run process will be resolved prior to the database lock and study unblinding.

Final Analysis after Database Lock and Study Unblinding

After the database has been quality assured and locked, the treatment assignments will be provided to the Sponsor's project statistician by the central randomization service and all planned analyses described in this document will be performed. By intent, no changes will be made to the clinical database after unblinding. However, any changes that are deemed necessary, after unblinding, will be clearly documented in the Clinical Study Report.

7 SAMPLE SIZE CONSIDERATIONS

A sample size of 472 subjects will have at least 80% power to detect a difference in means of 80 mL in change from Baseline to Week 52 in absolute FVC, assuming that the common standard deviation is 300 mL using a 2-group t-test with a 0.05 two-sided significance level. Assuming an 18% dropout rate, 576 subjects need to be randomized.

The assumptions related to sample size may be evaluated by review of blinded FVC data throughout the study and, if necessary, sample size will be recalculated using suitably modified assumptions on the standard deviation.

8 ANALYSIS POPULATIONS

The Intent-to-Treat (ITT) Population is defined as all subjects randomized into the study who received at least 1 dose of study drug. All ITT subjects will be counted in the group to which they were randomized, regardless of the study drug they were given. All efficacy analyses will be performed on this ITT Population, unless otherwise specified.

The Safety Population is defined as all subjects enrolled into the study who received at least 1 dose of study drug. All Safety Population subjects will be counted in the group corresponding to the study drug received, regardless of randomized assignment. All safety analyses will be performed on this Safety Population, unless otherwise specified.

The Per-protocol Population will include all subjects in the ITT Population, excluding subjects with protocol deviations that may have a notable impact on the primary efficacy analyses. The protocol deviations and the subject's exclusion from the Per-protocol Population will be reviewed at a blinded data review meeting and documented prior to the database lock and the study unblinding.

9 GENERAL CONSIDERATIONS FOR DATA ANALYSES

All the data collected in the electronic Case Report Form (eCRF) will be listed. In general, listings will be sorted by treatment group, subject number, and scheduled assessment (if applicable). Listings will include assessment date, assessment time (if available), study day, and all relevant data collected in the eCRFs. For data collected on a fixed schedule, the assessment identifier or the nominal time point will also be included on the listing. Repeat or redundant observations within an assessment window and observations that do not fall within any predefined assessment window (and will, therefore, be excluded from summaries) will be flagged in these listings. Subjects who are to be included in the analysis population (eg, ITT, Safety Population or Per-protocol Population) will be flagged.

In general, the data will be summarized by scheduled assessment (if applicable) within each treatment group. For continuous variables, summary statistics will include the mean, standard deviation, median, minimum, and maximum. For summaries of non-normal data such as NT-proBNP, interquartile range (lower quartile, upper quartile) may also be included. Minimums and maximums will be expressed using the level of precision in which the variable was collected. All other statistics will be rounded, using an additional decimal place than was collected. For discrete variables, summaries will include the frequency and percentage in each category. Percentages will be rounded to 1 decimal place. For all inferential analyses and descriptive comparisons, p-values will be rounded to 4 decimal places. Values less than 0.0001 will be denoted as <0.0001, and values greater than 0.9999 will be denoted as >0.9999. Whenever practical, categories of discrete variables will be ordered and labeled as they appear in the eCRF, and all categories represented on the Case Report Form will be included in summaries, even when they do not apply to any subjects in the study.

Unless otherwise specified, all statistical tests will be two-sided at alpha level 0.05. All statistical calculations will be completed using SAS® Version 9.4 or above.

9.1 COVARIATES

The primary efficacy analyses of change in FVC at Week 52 will be adjusted for IPF background therapy (with 3 levels: nintedanib or pirfenidone or no background therapy) and baseline FVC. For sensitivity analyses of the primary efficacy endpoints and analyses of the

secondary/exploratory endpoints with continuous variables, IPF background therapy (with 3 levels: nintedanib, pirfenidone, or no background therapy) and the corresponding baseline measure will always be the covariates.

9.2 EXAMINATION OF SUBGROUPS

For the primary efficacy endpoint of change in absolute FVC at Week 52, subgroup analyses will be performed, if data are available. These subgroups will include:

- IPF background therapy (nintedanib or pirfenidone vs no background therapy)
- IPF background therapy (nintedanib, pirfenidone, or no background therapy)
- Sex (male vs female)
- Age group (<65 years of age, 65 to <80 years of age, or ≥80 years of age)
- Time since IPF diagnosis (<median time since IPF diagnosis or ≥median time since IPF diagnosis)
- Obstructive sleep apnea (OSA) (yes, no, or unknown)
- Smoking history (current, former, or never) or (smoker versus never) as applicable. If the number of either former or current smokers is small, will combine former and current smokers as smokers.
- Corticosteroid use for IPF (yes or no)
- Supplemental oxygen use at Baseline (yes or no)
- Last study drug dose (<7 breaths, 7 to 9 breaths, or 10 to 12 breaths)
- Last study drug dose (<9 breaths, 9 to 12 breaths)
- Race (White, or Not White)
- Ethnicity (Hispanic or Latino, or Not Hispanic or Latino)
- Body mass index (BMI) at Baseline (<25, 25 to <30, or ≥30 kg/m2)

No Type 1 error adjustments will be performed for these subgroup analyses.

9.3 PREMATURE DISCONTINUATION AND MISSING DATA

Every effort will be made to complete all scheduled study assessments for all randomized subjects. Attempts will be made to continue to collect data after early termination of study drug. These data are being collected for analyses and descriptive summaries. On treatment analysis excluding data after early termination of study drug will be done as sensitivity analysis for FVC. More details about missing data for each efficacy endpoint are available in Section 11.

9.4 MULTIPLE COMPARISONS AND MULTIPLICITY

To control the familywise error rate at two-sided level 0.05, a gatekeeping procedure for multiple endpoints (including the primary efficacy endpoint and the group of secondary efficacy endpoints) will be used.

The primary efficacy endpoint of change in FVC from Baseline to Week 52 will be tested at two-sided alpha level 0.05. If the primary efficacy endpoint is statistically significant at two-sided alpha level 0.05, the statistical tests for secondary efficacy endpoints will be performed. Otherwise, tests will stop for the group of secondary efficacy endpoints.

The secondary efficacy endpoints will be tested by a pre-specified fixed sequence. The secondary endpoints will be tested in the following order:

1. Time to first clinical worsening (including time to death, respiratory hospitalization, or ≥10% relative decline in % predicted FVC)
2. Time to first acute exacerbation of IPF
3. Overall survival at Week 52
4. Change from baseline in % predicted FVC at Week 52
5. Change from baseline in K-BILD total score at Week 52
6. Change from baseline in DLCO at Week 52

Testing will proceed only if the preceding test is statistically significant at two-sided alpha level equal to 0.05 and will stop for any of the rest secondary efficacy endpoints if non-significance of the preceding test is reached. The 3 K-BILD domain scores (breathlessness and activities, chest symptoms, and psychological) will not be included in the pre-specified fixed sequence of the secondary efficacy endpoints for multiplicity adjustment.

9.5 DERIVED AND TRANSFORMED DATA

9.5.1 Spirometry

Spirometry will be performed using the SpiroSphere® spirometer provided by the Sponsor to all participating sites. Absolute and % predicted forced expiratory volume in 1 second (FEV_1) and FVC will be collected. At each timepoint, spirometry will be performed with the subject in a seated position. Spirometry maneuvers will be repeated (up to 8) until 3 acceptable attempts have been completed. If both pre- and post-bronchodilator testing is performed by the center, only the

pre-bronchodilator values will be recorded for purposes of the study. Efforts should be made to schedule the assessment at approximately the same time of day at each study visit. Subjects should refrain from strenuous activity for at least 12 hours and smoking for at least 30 minutes prior to testing. For subjects on bronchodilators, spirometry should be performed consistently throughout the study in relation to timing of bronchodilator administration. Spirometry must occur before DLCO. At Baseline, spirometry must occur before randomization and before administration of the first in-clinic dose of study drug. During the Treatment Period, spirometry will be scheduled to occur after the in-clinic dose of study drug. For subjects who discontinue study drug prematurely before Week 52, a Premature Study Drug Discontinuation (PSDD) Visit should be conducted as soon as possible, and as best as possible prior to the final dose of study drug. In the event a subject withdraws from the study, an Early Termination Visit will be conducted. Spirometry assessments should be performed prior to study drug discontinuation at the PSDD and/or Early Termination Visits, if at all possible.

Spirometry results will be transmitted for central review to ensure the quality of the primary endpoint measurement and to confirm eligibility at Screening. The central spirometry review will provide feedback to the investigational site on the performance of the testing as needed. Central review of spirometry at Screening may take up to 2 business days. All reviews of the spirometry data are based on American Thoracic Society guidelines (Graham 2019). Each effort will be overread and separate grades will be assigned to FVC and FEV_1 values as determined by the quality of the data. The letter test grade for each of the FVC and FEV_1 values as provided by the central spirometry review will be 1 of the following letters: A, B, C, D, E, F, or U. Per guidelines, the grading system reflects the acceptability and repeatability of spirometry efforts at each visit and is determined based on the set of measurements at each visit (Table 9-1).

Table 9-1 Grading System for FEV1 and FVC (Graded Separately)

Grade	Number of Measurements	Repeatability
A	≥ 3 acceptable	Within 150 mL
B	2 acceptable	Within 150 mL
C	≥ 2 acceptable	Within 200 mL
D	≥ 2 acceptable	Within 250 mL
E	≥ 2 acceptable OR 1 acceptable	>250 mL N/A

Grade	Number of Measurements	Repeatability
U	0 acceptable AND ≥ 1 usable	N/A
F	0 acceptable and 0 usable	N/A

Abbreviation: N/A, not applicable
NOTE: The repeatability criteria are applied to the differences between the 2 largest FVC values and the 2 largest FEV1 values.

These letter test grades will fall into 1 of the following overread grades: Acceptable (A or B), Borderline Acceptable (C, D, or E), or Unacceptable (F or U). The test with best effort at each visit for FVC and FEV_1 will also be determined and used for analyses.

9.5.2 Lung Diffusion Capacity

Measurement of absolute and % predicted DLCO (uncorrected for hemoglobin and lung volume) will be performed using each clinical site's own equipment. If several devices are available at the site, the same equipment should be used to conduct study assessments at each measurement if possible.

DLCO measurements must be performed after the spirometry assessments. Efforts should be made to schedule the DLCO assessment at subsequent visits at approximately the same time of day.

DLCO measurement data will be transmitted by the site to the designated central respiratory testing reader. The central review will provide feedback to the investigational site on performance of the testing as needed. All reviews of DLCO data are based on American Thoracic Society guidelines (Graham 2017). Based on acceptability and repeatability criteria, the quality of DLCO data will be graded into one of the following categories: Acceptable, Borderline Acceptable, and Unacceptable.

9.5.3 Clinical Worsening

Clinical worsening is a composite endpoint based on death (all causes), respiratory hospitalization, or ≥10% relative decline from baseline in % predicted FVC. Respiratory hospitalization is an adjudicated event that will be evaluated by the independent Clinical Endpoint Committee (CEC). Only those hospitalizations determined by the CEC to be respiratory-related or those classified as 'Indeterminate' will be analyzed. A subject could have

multiple incidences of respiratory-related hospitalizations during the study. All causes of deaths are collected in the eCRF, including deaths during the study (from randomization until the end of the study) or deaths captured in vital status for early terminated subjects. Only death during the study will be counted as a component of the clinical worsening event. The relative decline from baseline in % predicted FVC (with acceptable or borderline acceptable grades) will be calculated programmatically based on the centrally reviewed spirometry results.

Time to first clinical worsening will be evaluated as a secondary endpoint and calculated as the time from randomization until 1 of the following criteria are met:

- Death (all causes) during the study
- Respiratory hospitalization
- ≥10% relative decline from baseline in % predicted FVC

Time to first clinical worsening is calculated as described in Table 9-2.

Table 9-2 Derivation of Time to First Clinical Worsening

Parameter	Scenario	Formula	Status
Time to clinical worsening (weeks)	Subjects with clinical worsening event reported prior to elective lung transplantation or initiation of commercial inhaled treprostinil products, if any	= (date of the First Clinical Worsening Event – Randomization date + 1)/7	0 (event)
	Subjects with a clinical worsening event reported after initiation of commercial inhaled treprostinil products	= (Commercial inhaled treprostinil products initiation date – Randomization date + 1)/7	1 (censored)
	Subjects completed the study without clinical worsening event	= (Censoring date[a] – Randomization date + 1)/7	1 (censored)
	Subjects discontinued from the study prematurely without clinical worsening event	= (Censoring date[a] – Randomization date + 1)/7	1 (censored)

a The end of study date, elective lung transplantation date, or initiation of commercial inhaled treprostinil products date, whichever is earliest.

Time to first respiratory hospitalization will be calculated as defined in Table 9-3.

Table 9-3 Derivation of Time to First Respiratory Hospitalization

Parameter	Scenario	Formula	Status
Time to respiratory hospitalization (weeks)	Subjects with respiratory hospitalization reported prior to elective lung transplantation or initiation of commercial inhaled treprostinil products, if any	= (First Respiratory hospitalization date – Randomization date + 1)/7	0 (event)
	Subjects with respiratory hospitalization reported after initiation of commercial inhaled treprostinil products	= (Commercial inhaled treprostinil products initiation date – Randomization date + 1)/7	1 (censored)
	Subjects completed study without respiratory hospitalization	= (Censoring date[a] – Randomization date + 1)/7	1 (censored)
	Subjects discontinued from study prematurely without respiratory hospitalization	= (Censoring date[a] – Randomization date + 1)/7	1 (censored)

a The end of study date, elective lung transplantation date, or initiation of commercial inhaled treprostinil products date, whichever is earliest.

9.5.4 Acute Exacerbations of IPF

Acute exacerbation of IPF is an adjudicated event that will be evaluated by the independent CEC. Only those exacerbations confirmed by the CEC to be acute exacerbations of IPF or those classified as 'Indeterminate' will be analyzed. Evaluation of multiple acute exacerbations of IPF will follow the same definition. Time to first acute exacerbation of IPF is calculated as described in Table 9-4.

Table 9-4 Derivation of Time to First Acute Exacerbation of IPF

Parameter	Scenario	Formula	Status
Time to acute exacerbation of IPF (weeks)	Subjects with acute exacerbation of IPF reported prior to elective lung transplantation or initiation of commercial inhaled treprostinil products, if any	= (First Exacerbation date – Randomization date + 1)/7	0 (event)
	Subjects with acute exacerbation of IPF reported after initiation of commercial inhaled treprostinil products	= (Commercial inhaled treprostinil products initiation date – Randomization date + 1)/7	1 (censored)
	Subjects completed study without acute exacerbation of IPF	= (Censoring date[a] – Randomization date + 1)/7	1 (censored)
	Subjects discontinued from the study prematurely without acute exacerbation of IPF	= (Censoring date[a] – Randomization date + 1)/7	1 (censored)

Abbreviation: IPF, idiopathic pulmonary fibrosis

a The end of study date, elective lung transplantation date, or initiation of commercial inhaled treprostinil products date, whichever is earliest.

9.5.5 Overall Survival at Week 52

Time to all-cause death, including death during study and death captured in vital status for early terminated subjects, will be calculated as defined in Table 9-5.

Table 9-5 Derivation of Time to Death (Overall Survival)

Parameter	Scenario	Formula	Status
Time to death[a] (weeks)	Subjects who died during the study or with death captured in vital status for early termination subjects	= (Death date – Randomization date + 1)/7	0 (event)
	Subjects with elective lung transplantation or initiation of commercial inhaled treprostinil products	= (Censoring date[b] – Randomization date +1)/7	1 (censored)
	Subjects completed study without death	= (Last contact date – Randomization date +1)/7	1 (censored)
	Subjects discontinued from the study prematurely without death	= (Last contact date – Randomization date + 1)/7	1 (censored)

a Subjects completed or discontinued from study without death will have their time to death censored at (365 + 7) days after randomization or last contact date when known to be alive, whichever is earlier.
b The end of study date, elective lung transplantation date, or initiation of commercial inhaled treprostinil products date, whichever is earliest.

Respiratory-related death (during study or captured in vital status for early terminated subjects) is an adjudicated event that will be evaluated by the independent CEC. Deaths confirmed by the CEC as respiratory-related or of unknown cause will be included in the analysis, while those confirmed as not respiratory-related will be excluded. Time to respiratory-related death will be calculated as defined in Table 9-5 with time to death replaced by time to respiratory-related death and death date replaced by respiratory-related death date.

9.6 ASSESSMENT WINDOWS

For any data summarized by scheduled visit, an analysis visit window will be used. The scheduled visits, as recorded on the eCRFs, and the corresponding target study day and study day intervals are specified in Table 9-6. The analysis visit window will be derived based on the information specified in Table 9-6.

In the protocol, the visit window for Week 4, Week 8, and Week 16 is ±5 days, and the visit window for Week 28, Week 40, and Week 52 is ±7 days. To allow slotting of discontinuation

visits and visits outside the planned schedule, the visit windows have been expanded for analysis purposes for Week 4, Week 8, Week 16, Week 28, Week 40, and Week 52.

Table 9-6 Assessment Windows for Scheduled Visits

Visit	Target Study Day	Study Day Interval
Spirometry, Clinical Laboratory Assessments, Vital Signs		
Baseline	1	Study Day ≤1 (prior to the first dose)
Week 4	29	1 < Study Day ≤ 43
Week 8	57	43 < Study Day ≤ 85
Week 16	113	85 < Study Day ≤ 155
Week 28	197	155 < Study Day ≤ 239
Week 40	281	239 < Study Day ≤ 323
Week 52	365	323 < Study Day
ECGs, K-BILD, DLCO		
Baseline	1	Study Day ≤1 (prior to the first dose)
Week 28	197	1 < Study Day ≤ 281
Week 52	365	281 < Study Day
NT-proBNP		
Baseline	1	Study Day ≤1 (prior to the first dose)
Week 16	113	1 < Study Day ≤ 155
Week 28	197	155 < Study Day ≤ 239
Week 40	281	239 < Study Day ≤ 323
Week 52	365	323 < Study Day

Abbreviations: DLCO, diffusion capacity of lungs for carbon monoxide; ECG, electrocardiogram; K-BILD, King's Brief Interstitial Lung Disease Questionnaire; NT-proBNP, N-terminal pro-brain natriuretic peptide
NOTE: Study Day = (Assessment Date) − (First Dosing Date) + 1

9.6.1 Multiple Evaluations Within the Same Analysis Window

After all the observations have been slotted based on the Table 9-6, if there are multiple valid observations for an assessment within an assigned analysis visit window, only 1 of these observations will be used for summary statistics and analyses. The observation to be used is determined using the following hierarchy (in decreasing order):

- The observation closest to the target study day
- The later observation if 2 observations are equally close to the target study day

10 STUDY POPULATION

Unless otherwise specified, all efficacy analyses will be performed on the ITT Population and all safety analyses will be based on the Safety Population. In the ITT Population, the treatment assignment is based on the assignment upon randomization. In the Safety Population, the treatment assignment is based on the actual treatment the subject received.

The comparability between the 2 treatment groups will be checked for demographic and baseline characteristics. The p-values from Fisher's exact test (for discrete variables), group t-test, or Wilcoxon rank sum test (for continuous variables) will be included on summaries but are not intended to be used to test formal hypotheses. For these comparisons, missing or unknown values will be excluded from the calculations.

10.1 PATIENT ACCOUNTABILITY

Subject disposition will be listed for all subjects by individual subject number, including the analysis population the subject belongs to, premature study drug discontinuation status, primary reason for premature study drug discontinuation, premature study discontinuation status, primary reason for premature study discontinuation, number of days on the study drug, and enrollment in the open-label extension study (RIN-PF-302).

The listing of subject accountability will include the dates of informed consent, randomization, first study drug dose, last study drug dose, and last assessment; and the study week of last dose and last assessment. Whether subjects received the study drug, whether subjects completed the Weeks 4, 8, 16, 28, 40, and 52 assessments, and study discontinuation status will be summarized.

Information regarding whether each subject is included in each analysis population (see Section 8) will be listed. If a subject is not included in a particular analysis population, the reason for exclusion will be noted on the listing. Also noted on the listing will be the randomized treatment assignment. The summary will include the frequency and percentage of all subjects in each analysis population.

The stratification information used in the random assignment of subjects to treatment group (from the central randomization database) as well as the actual IPF background therapy (if mis-stratification occurred), will be listed, including date and time of randomization. Status of

the treatment blind, and if broken, date of blind broken will be listed. The number of subjects in each stratum from the central randomization database, along with the actual IPF background therapy (if mis-stratification occurred), will be summarized by treatment group and overall.

10.2 PROTOCOL DEVIATIONS

The status of the entry criteria will be listed for all subjects. The listing will include the date of the initial screening assessment, and a list of any specific entry criteria not met. This listing will also include the protocol version that the subject was enrolled under. Entry criteria violations will be summarized by treatment group and overall.

Additionally, protocol deviations will be documented throughout the study. All deviations will be reviewed by the clinical team prior to database lock and those that might affect subject safety or efficacy outcomes will be considered "critical" or "major." All other deviations will be classified as "minor." Protocol deviations will be listed, including the date of the deviation, the type of deviation, the severity of the deviation (critical/major/minor), and a description of the deviation. The protocol deviations will also be summarized by treatment group and overall. In the summary, critical and major deviations will be combined as "major".

10.3 OTHER DESCRIPTIONS OF STUDY POPULATION

10.3.1 Demographics

Demographic data will be listed for all subjects, including assessment date, date of birth, country, age, sex, ethnicity, race, weight, height, and BMI. Age, age category (<65 years of age, 65 to <80 years of age, and ≥80 years of age), sex, ethnicity, race, height, weight, and BMI will be summarized by treatment group and overall. The summary will include p-values (two-sided) from Fisher's exact test (for age category, sex, ethnicity, and race) and group t-test or Wilcoxon rank sum test (for age, weight, height, and BMI) comparing treatment groups.

10.3.2 Medical History and IPF History

Significant past or ongoing medical conditions will be listed for all subjects. The listing will include the assessment date, assessment day, Medical Dictionary for Regulatory Activities System Organ Class (SOC) and Preferred Term (PT) for each condition listed, and whether the

condition is ongoing at randomization. These medical conditions will be summarized by PT within each SOC.

Information related to subject IPF history will be listed. The listing will include the date of initial IPF diagnosis; years since IPF diagnosis; presence of OSA; OSA treatment, if any; smoking history; corticosteroid usage for IPF; and IPF background therapy, if any.

The time since IPF diagnosis, OSA (yes, no, or unknown), whether OSA is treated (yes or no) and OSA treatment (positive airway pressure device, oral appliance, or other), smoking history (current, former, or never), number of pack-years, corticosteroid usage for IPF (yes or no), IPF background therapy (nintedanib, pirfenidone, or none), and primary reason for not taking IPF background therapy will be summarized by treatment group and overall. The summary will include the p-values from Wilcoxon rank sum test or Fisher's exact test comparing treatment groups.

10.3.3 Concomitant Medications

All concomitant medications recorded on the eCRF will be mapped to a standard name and Anatomical Therapeutic Chemical (ATC) Levels 1 to 4 using the WHODrug Global dictionary. The ATC Levels 2 and 4 and verbatim term of all concomitant medications will be listed for all subjects. This listing will include the date started, date discontinued (or indication that drug was ongoing at the end of study), and the condition(s) treated/indication(s). Summary of concomitant medications present at Baseline and summary of concomitant medications added during the study will include the frequency and percentage of subjects in each treatment group receiving each drug by ATC Level 2 and Level 4. If ATC Level 4 is not available for medication, ATC Level 3 will be substituted. If ATC Level 3 is not available, ATC Level 2 will be substituted. If ATC Level 2 is not available, ATC Level 1 will be substituted.

11 EFFICACY ANALYSES

Except where otherwise noted, all efficacy analyses will only be performed on the ITT Population (see Section 8).

11.1 PRIMARY EFFICACY MEASURE

11.1.1 Primary Estimand

The primary efficacy endpoint of change in absolute FVC from Baseline to Week 52 assesses if inhaled treprostinil is superior to placebo in subjects with IPF. The null and alternative hypotheses are:

$$H_0: \mu_1 = \mu_2$$

$$H_a: \mu_1 \neq \mu_2,$$

where μ_1 and μ_2 are the respective mean change from baseline in FVC of subjects assigned to the inhaled treprostinil or placebo treatment group, regardless of the extent of their adherence to treatment or early study treatment discontinuation, but excluding the effects of elective lung transplantation and initiation of commercial inhaled treprostinil products.

The following strategies will be used for intercurrent events:

- Treatment policy for early treatment discontinuation, all values observed after early treatment discontinuation will be included.
- Treatment policy for initiation or change of IPF background therapy during the study, all values observed after initiation or change of IPF background therapy will be included.
- Treatment policy for rescue medication initiated during the study (deemed as safe and clinically necessary by the investigators), the rescue medication includes chronic concomitant medication of AZA, cyclosporine, mycophenolate mofetil, tacrolimus, OCS >20 mg/day, or the combination of OCS+AZA+N-acetylcysteine; cyclophosphamide; rituximab, endothelin receptor antagonists, phosphodiesterase type 5 inhibitors, or soluble guanylate cyclase stimulators; all values observed after these intercurrent events will be included.
- Treatment policy for initiation of prostacyclin therapy (epoprostenol, treprostinil excluding commercial inhaled treprostinil products, iloprost, or beraprost; IP receptor agonists [selexipag]) during study; all values observed after initiation of prostacyclin therapy other than commercial inhaled treprostinil products will be included.
- Composite variable strategy for death; the FVC change from baseline values at scheduled visits on or after the death date will be assigned the 2.5th percentile of all observed values (FVC change from baseline) across both treatment groups at each visit.
- Hypothetical strategy for initiation of commercial inhaled treprostinil products during study; data observed after initiation of commercial inhaled treprostinil products will be excluded.
- Hypothetical strategy for elective lung transplantation; data observed after elective lung transplantation will be excluded.

11.1.2 Primary Efficacy Analysis

The primary efficacy analysis will use all FVC values with an overread grade of Acceptable or Borderline Acceptable (with overread grade A, B, C, D or E). All FVC values that have been deemed Unacceptable (grades F or U; see Section 9.5.1) by the central review will be excluded from the primary efficacy analysis due to their lack of acceptability and repeatability per ATS guideline (Graham 2019). The total number of Unacceptable FVC values, the number of Unacceptable FVC values by visit, and the frequency of Unacceptable FVC values per subject will be summarized by treatment group. After applying different strategies to handle the FVC values after intercurrent events, the mixed-effect model for repeated measures (MMRM) will be performed to estimate the treatment difference in change from baseline in FVC at Week 52. The MMRM will include the change from baseline in FVC as the dependent variable, treatment, week, and treatment by week interaction as fixed effects, and baseline FVC and IPF background therapy (with 3 levels: nintedanib or pirfenidone or no background therapy) as covariates. An unstructured variance/covariance structure shared across treatment groups will be used to model the within-subject errors. Any missing FVC values at visits before or at Week 52 (other than those due to death) will be handled implicitly by the MMRM model. The main comparison will be a contrast between treatment groups at Week 52.

Sample SAS code for this analysis is shown below:

```
PROC MIXED;
  CLASS USUBJID TRT01P AVISITN BGTX;
  MODEL CHG = BGTX BASE TRT01P AVISITN TRT01P*AVISITN / DDFM=KR;
  REPEATED AVISITN / SUBJECT=USUBJID TYPE=UN;


 ESTIMATE '52'     TRT01PN 1 -1
                   TRT01PN * AVISITN 0 0 0 0 0 1
                                     0 0 0 0 0 -1/CL;
RUN;
```

where USUBJID is subject identifier, TRT01P is treatment group, AVISITN is study week, BGTX is background therapy, CHG is change from baseline in FVC, and BASE is baseline FVC.

The assumptions of the MMRM model include normality of residuals and homoscedasticity. They will be examined using various approaches, such as visual inspection (eg, histogram, Q-Q plots, plots of residuals against predicted values) or formal normality tests, as appropriate. If the assumptions are strongly violated, the change from baseline in absolute FVC at Week 52 will be analyzed instead by using the non-parametric aligned rank stratified Wilcoxon test and the corresponding aligned rank Hodges-Lehmann estimate. The alternative primary analysis includes the following steps:

Step 1: After applying different strategies to handle the FVC values after intercurrent events, a standard multiple imputation (Rubin, 1987) will be used to appropriately impute missing FVC values at visits before (or at) Week 52 due to non-fatal reasons. Fully Conditional Specification (FCS) regression will be used to fill the missing points in the order of timepoints using measurements calculated at the previous timepoints. Sample SAS code to generate 100 imputed datasets is shown below;

```
PROC MI DATA=FVC OUT=VIS NIMPUTE=100 SEED=162341;
  CLASS TRT01PN SEXN BGTX;
  VAR TRT01PN SEXN AGERAND BGTX BASE V4 V8 V16 V28, V40 V52;
  FCS REG (V4 = TRT01PN SEXN AGERAND BGTX BASE );
 FCS REG (V8 = TRT01PN SEXN AGERAND BGTX BASE V4);
 FCS REG (V16 = TRT01PN SEXN AGERAND BGTX BASE V4 V8);
FCS REG (V28 = TRT01PN SEXN AGERAND BGTX BASE V4 V8 V16);
FCS REG (V40 = TRT01PN SEXN AGERAND BGTX BASE V4 V8 V16 V28);

 FCS REG (V52 = TRT01PN SEXN AGERAND BGTX BASE V4 V8 V16 V28 V40);
RUN;
```

where TRT01PN is treatment group, V4, V8, V16, V28, V40 and V52 are FVC values at Week 4,8, 16, 28, 40 and 52, BGTX is background therapy, SEXN is baseline sex, AGERAND is age at randomization, and BASE is baseline FVC.

Step 2: For each imputed dataset, change from baseline in absolute FVC at Week 52 will be analyzed by aligned rank stratified Wilcoxon test. Sample SAS code is shown below;

```
PROC NPAR1WAY DATA=VIS WILCOXON ALIGN=STRATA(HL) HL CORRECT=NO;
  CLASS TRT01PN;
  VAR CHG;
 STRATA BGTX;
 BY IMPNUMBER;
 ODS OUTPUT HODGESLEHMANN=HLSTATS_A;
RUN;
```

where CHG is the change from Baseline in FVC at Week 52.

Step 3: The results for the 100 imputed datasets are combined for inference by Sample SAS code:

```
PROC MIANALYZE DATA=HLSTATS_A;
  MODELEFFECTS SHIFT;
  STDERR stderr;
 ODS OUTPUT PARAMETERESTIMATES=RESULTS;
RUN;
```

The standard multiple imputation is based on the missing at random assumption for missing data. The assumption of missing at random will be evaluated by the control-based pattern-mixture model method of Yuan (Yuan 2014). The control-based pattern-mixture model creates imputed values for the active treatment group from the observed data in the placebo treatment group. This assumes that subjects who discontinue from active treatment would behave similarly after dropout to those in the placebo group.

Similar 3-steps analyses described above will be repeated but with MNAR statement in PROC MI to follow control-based pattern-mixture model. In Step 1, treatment, baseline FVC, age at randomization, sex, background therapy, presence of OSA, occurrence of clinical worsening, and all observed values of the FVC for non-fatal subjects will be included in the multiple imputation model with 100 imputations, 200 burn-in imputations, and a noninformative prior. Sample SAS code for multiple imputation is shown below;

```
PROC MI DATA=FVC OUT=VIS NIMPUTE=100 SEED=162341;
  CLASS TRT01PN;
```

```
  VAR SEXN AGERAND BGTX OSA CW BASE V4 V8 V16 V28 V40 V52;
  MNAR MODEL (V4 V8 V16 V28 V40 V52 / MODELOBS=(TRT01PN='2'));
  FCS NBITER=200;
RUN;
```

11.1.3 Sensitivity and Subgroup Analyses

To further support the robustness and assess the sensitivity of the primary efficacy analysis of change in FVC at Week 52, multiple sensitivity and subgroup analyses for the primary endpoint will be implemented, provided that the primary analysis yields significant results. No multiplicity adjustment is planned for these analyses.

The primary efficacy analysis in Section 11.1.2 will be repeated using alternative composite strategies for death. Specifically, for post-death scheduled visits, the FVC change from baseline will be assigned using 3 different extents of penalties – the 0.1st, 5th, and 10th percentiles of all observed values (FVC changes from baseline) across both treatment groups at each visit. Two additional composite strategies will also be applied for the Week 52 absolute FVC for subjects who died before Week 52 – one assigning an extreme unfavorable value of 0, and the other assigning half of the minimum observed absolute FVC at Week 52 across both treatment groups. The sensitivity analyses based on alternative composite strategies will be analyzed using both the MMRM methodology and the non-parametric aligned stratified Wilcoxon test with the corresponding aligned rank Hodges-Lehmann estimate, as described in Section 11.1.2.

In addition, for subsequent sensitivity and subgroup analyses for the primary endpoint, similar strategies to handle intercurrent events as the primary analysis except for death, which will be handled by hypothetical strategy for simplicity in analyses. That is, apply treatment policy for early study treatment discontinuation, initiation or change of background therapy during the study, initiation of rescue medication and, and initiation of prostacyclin therapy (other than commercial inhaled treprostinil products); and hypothetical strategy for death, elective lung transplantation and initiation of commercial inhaled treprostinil products. Any missing FVC values at visits before (or at) Week 52 will be handled implicitly by the MMRM model.

After applying different strategies for intercurrent events, the MMRM analysis will be repeated using each of the following modifications (as data permit):

- The Per-protocol Population will be used instead of the ITT Population
- All FVC values with an overread grade of Acceptable or Borderline Acceptable (with overread grade A, B, C, D or E)
- Including only those FVC values deemed Acceptable (A or B)
- Including all FVC values (regardless of grades)
- On-treatment analysis, excluding data collected after termination of the study drug (that is to exclude assessment performed more than 24 hours after the last dose date)
- MMRM model adding random site effect.

Subgroup and stratified analyses of change in FVC at Week 52, if data available:

- By IPF background therapy categories (nintedanib or pirfenidone vs no background therapy)
- By IPF background therapy categories (nintedanib, pirfenidone, or no background therapy)
- By sex (male vs female)
- By age group (<65 years of age, 65 to <80 years of age, or ≥80 years of age)
- By time since IPF diagnosis (<median time since IPF diagnosis or ≥median time since IPF diagnosis)
- By OSA (yes, no, or unknown)
- By smoking history (current, former, or never) or (smoker versus never) as applicable. If the number of either former or current smokers is small, will combine former and current smokers as smokers.
- By corticosteroid use for IPF (yes or no)
- By last study drug dose (<7 breaths, 7 to 9 breaths, or 10 to 12 breaths)
- By last study drug dose (<9 breaths, 9 to 12 breaths)
- By Supplemental Oxygen Use at Baseline (Yes or No)
- By Race (White, or Not White)
- By Ethnicity (Hispanic or Latino, or Not Hispanic or Latino)
- By BMI at Baseline (<25, 25 to <30, or ≥ 30 kg/m2)

For stratified analyses, the same MMRM model will be applied on subset of the FVC data within each of the subgroups.

Analyses of FVC at Week 52 by subgroups will also be evaluated by MMRM model with change from baseline in FVC as dependent variable, 3-way (treatment, week, subgroup) full model and

background therapy as the fixed effects, and baseline FVC as covariate. The results from the subgroup analyses will be displayed as a forest plot.

As a supporting analysis, after applying the same strategies for intercurrent events as sensitivity and subgroup analyses for the primary endpoint, a random coefficient regression model including random slopes and intercepts will be constructed to estimate the rate of change in FVC over the 52-week treatment period. The model will include Baseline FVC and background therapy as covariates. The rate of change will be calculated for each subject and compared between treatment groups. A histogram of the individual subject slopes will also be provided. Sample SAS code for this analysis is shown below:

```
PROC MIXED;
  CLASS USUBJID TRT01P BGTX;
  MODEL FVC = BGTX BASE TRT01P BASE*WK TRT01P*WK / DDFM=KR;
  RANDOM INTERCEPT WK/ SUBJECT=USUBJID;
  ESTIMATE 'Active' BASE*WK meanbase TRT01P*WK 1 0/CL;
  ESTIMATE 'Placebo' BASE*WK meanbase TRT01P*WK 0 1/CL;
  ESTIMATE 'Active – Placebo' TRT01P*WK 1 -1/CL;
RUN;
```

where USUBJID is subject identifier, TRT01P is treatment group, WK is time from randomization to the date of FVC measurement, which can be derived by (date of FVC-randomization date +1)/7, BGTX is background therapy, FVC is the absolute FVC, BASE is baseline FVC, "meanbase" is the mean of baseline FVC values across both treatment groups, and INTERCEPT represents the random intercept for each subject.

Two additional supplementary analyses will be performed. First, a ranked ANCOVA will be conducted on the change from baseline in FVC at Week 52, in which patients who die are ranked below all patients who remain alive at Week 52. Among the patients who died, the patient who died earliest will be ranked lowest, with the remaining patients who died ranked in order of the length of time alive since randomization. In addition, a "While Alive" analysis will be conducted using ANCOVA on the change from baseline in FVC at Week 52, which will include all patients based on the following rules:

1. Observed values for patients with available FVC data at Week 52;
2. Multiple imputation for patients who were alive at Week 52 but had missing data;
3. The last available FVC measurement while alive for patients who died prior to Week 52.

11.2 SECONDARY EFFICACY MEASURES

Secondary efficacy endpoints include the following:

1. Time to first clinical worsening (including time to death, respiratory hospitalization, or ≥10% relative decline in % predicted FVC)
2. Time to first acute exacerbation of IPF
3. Overall survival at Week 52
4. Change from baseline in % predicted FVC at Week 52
5. Change from baseline in K-BILD score at Week 52
6. Change from baseline in DLCO at Week 52

As specified in Section 9.4, the pre-specified fixed sequence-testing procedure will be employed for this group of secondary endpoints (time to first clinical worsening event, time to first acute exacerbation of IPF, overall survival at Week 52, change in % predicted FVC at Week 52, change in K-BILD total score at Week 52 and change in DLCO at Week 52), provided that the primary analysis for primary endpoint is significant, to control the overall two-sided alpha level at 0.05.

For secondary efficacy measures (time to first clinical worsening, time to first acute exacerbation of IPF), primary analyses will include all events occurring during the study; for those subjects who discontinue study treatment prematurely, the events occurring after treatment discontinuation will be counted. For overall survival at Week 52, all deaths collected by Week 52 (including deaths during the study and deaths collected in vital status forms after early study termination) will be included.

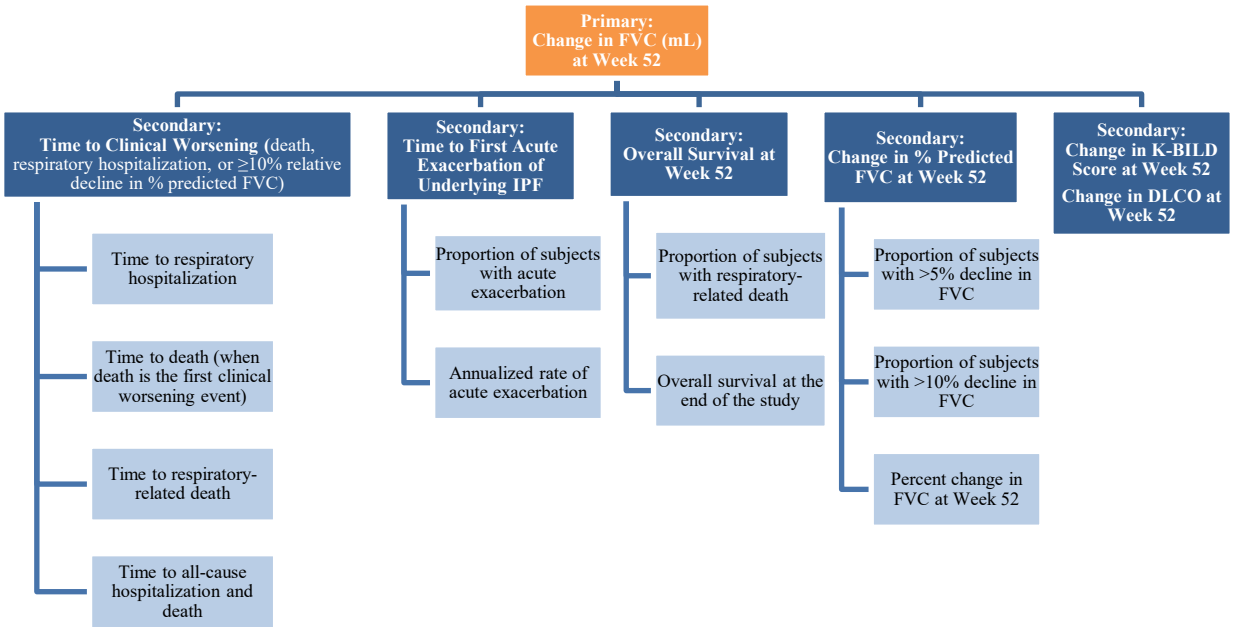
Sensitivity analyses for time to on-treatment events (if enough number of events are observed) will also be executed. Each endpoint of time to on-treatment events will be tested at a two-sided alpha level equal to 0.05. No adjustments for multiplicity of these on-treatment events will be done.

On-treatment clinical worsening events include ≥10% relative decline in % predicted FVC occurring within 7 days after the last dose date, deaths captured in SAE forms with death date within 30 days after the last dose date, and respiratory hospitalization occurring within 7 days after the last dose date. On-treatment exacerbation of IPF refers to exacerbation of IPF occurring within 7 days after the last dose date. On-treatment respiratory hospitalization refers to respiratory hospitalization occurring within 7 days after the last dose date. On-treatment death refers to death captured in SAE forms with death date within 30 days after last dose date.

For time to first on-treatment clinical worsening events, subjects without on-treatment clinical worsening events will be censored at 7 days after last dose date or end of study date, whichever is earlier (for simplicity). For time to on-treatment respiratory hospitalization and time to on-treatment exacerbation of IPF, subjects without corresponding on-treatment events will be censored at 7 days after last dose date or end of study date, whichever is earlier. For time to on-treatment death, subjects who survived without on-treatment events will be censored at 30 days after the last dose date or end of study date, whichever is earlier.

Additionally, other endpoints related to the secondary efficacy endpoints, as shown in Figure 11-1 will be evaluated. Each endpoint nested below a secondary endpoint will be tested at a two-sided alpha level equal to 0.05. No adjustments for multiplicity of these other endpoints will be done.

Figure 11-1 Other Efficacy Endpoints



Abbreviations: FVC, forced vital capacity; IPF, idiopathic pulmonary fibrosis; K-BILD, King's Brief Interstitial Lung Disease Questionnaire

Time to death (when death is the first clinical worsening event) only includes death during the study and is calculated by similar rule as Time to First Clinical Worsening Events but to censor those non-fatal clinical worsening event at their first clinical worsening date. This approach is described in Table 11-1.

Table 11-1 Derivation of Time to Death (when death is the first clinical worsening event)

Parameter	Scenario	Formula	Status
Time to Death (with death as the first clinical worsening events)	Subjects with death as the first clinical worsening events	= (Death date – Randomization date + 1)/7	0 (event)
	Subjects who died during the study but death is not the first CW event	= (First clinical worsening event date – Randomization date + 1)/7	1 (censored)
	Subjects without death during the study but with clinical worsening event other than death	= (First clinical worsening event date – Randomization date + 1)/7	1 (censored)
	Subjects without clinical worsening event during the study	= (Censoring date[a] – Randomization date + 1)/7	1 (censored)
	Subjects discontinued from the study prematurely without clinical worsening event	= (Censoring date[a] – Randomization date + 1)/7	1 (censored)

a The end of study date, elective lung transplantation date, or initiation of commercial inhaled treprostinil products date, whichever is earliest.

11.2.1 Time to Clinical Worsening

Clinical worsening events (deaths, adjudicated respiratory hospitalizations, and ≥10% relative decline in % predicted FVC) including the date (study day) of the event and category of the clinical worsening event will be listed. The time to first clinical worsening will be calculated according to the rules described in Table 9-2.

The number and percentage of subjects with any clinical worsening and with each category of clinical worsening will be summarized by treatment group. Time to first clinical worsening will be summarized by treatment group using product-limit estimates calculated by the Kaplan-Meier (K-M) method and displayed graphically as K-M curves. A tabular summary of this analysis will include the number of subjects at risk (sample size), estimated median duration, and a 95% confidence interval [CI] for the median duration for each treatment group. A Cox proportional hazards model will be fit to obtain the hazard ratio and its associated 95% CI and p-value. The model will include treatment and background therapy as explanatory variables. Sample SAS code for the K-M analysis and the Cox regression are shown below:

```
PROC LIFETEST MAXTIME=52;
  TIME AVAL*CNSR(1);
  STRATA TRT01P ORDER=INTERNAL;
RUN;


PROC PHREG;
  CLASS TRT01P BGTX;
  MODEL AVAL*CNSR(1) = TRT01P BGTX / RISKLIMITS ALPHA=0.05 TIES=EFRON;
RUN;
```

where TRT01P is treatment group, AVAL is time to clinical worsening and CNSR is censor status as defined in Table 9-2, and BGTX is background therapy.

If it is determined that the proportional hazards assumption does not hold for background therapy (but does hold for treatment), a stratified analysis will be conducted. Sample SAS code will be:

```
PROC PHREG;
```

```
  CLASS TRT01P BGTX;
  MODEL AVAL*CNSR(1) = TRT01P / RISKLIMITS ALPHA=0.05 TIES=EFRON;
  STRATA BGTX;
RUN;
```

If the proportional hazards assumption does not hold for the Cox model (neither for background therapy nor for treatment), the logrank test, adjusted for background therapy, will be used to calculate the p-value for treatment differences. Some alternative approaches may also be explored, such as, the restricted mean survival time (Royston 2013) and the "Max-Combo" test analysis (Lin 2020; Knezevic 2020).

Multiple clinical worsening events will be analyzed via Poisson regression. The model will include treatment group, background therapy, and log of time on study as an offset. The rate of clinical worsening events will be summarized by treatment group. If the data exhibit overdispersion, negative binomial regression will be used. Sample SAS code follows:

```
PROC GENMOD;
  CLASS TRT01P BGTX;
  MODEL COUNT = TRT01P BGTX / TYPE3 DIST=POISSON OFFSET=TIMEONSTY;
RUN;
```

where TRT01P is treatment group, BGTX is background therapy, COUNT is the number of adjudicated clinical worsening events for each subject and TIMEONSTY is calculated as log(end of study date – treatment start date + 1).

As described above in Figure 11-1, each endpoint nested below Time to clinical worsening will be tested at a two-sided alpha level equal to 0.05. No adjustments for multiplicity of these other endpoints will be done.

Time to death (when death is the first event of clinical worsening) and time to respiratory-related death will be analyzed via log-rank test, K-M methodology, and Cox regression.

For respiratory hospitalization, the proportion of subjects who experienced ≥1 respiratory hospitalization will be summarized by treatment group and compared with a chi-square test.

Time to first respiratory hospitalization (calculated according to the rules described in Table 9-3) will be summarized and analyzed similar to time to first clinical worsening events. It will be summarized by treatment group using product-limit estimates calculated by the K-M method and displayed graphically as K-M curves. A tabular summary of this analysis will include the number of subjects at risk (sample size), estimated median duration, and a 95% CI for the median duration for each treatment group (with p-value obtained from the log-rank test adjusted for background therapy). A Cox proportional hazards model will be fit to obtain the hazard ratio and its associated 95% CI. The model will include treatment and background therapy as explanatory variables. Multiple respiratory hospitalizations will be analyzed via Poisson regression.

For all-cause hospitalization and death (only including death during the study) as a composite endpoint, the proportion of subjects who experienced ≥1 hospitalization or death will be summarized by treatment group and compared with a chi-square test. Time to first all-cause hospitalization or death will be summarized by treatment group and analyzed via K-M methodology and Cox regression similar to time to first clinical worsening events. Multiple events of composite all-cause hospitalization and/or death will be analyzed via Poisson regression.

11.2.2 Time to Acute Exacerbation of IPF

Time to first acute exacerbation of IPF (as adjudicated by the CEC) will be summarized by treatment group and analyzed via K-M methodology, log-rank test, and Cox regression similar to that described in Section 11.2.1.

For acute exacerbation of IPF, the proportion of subjects who experienced ≥1 acute exacerbation of IPF will be summarized by treatment group and compared with a chi-square test.

Multiple exacerbations of IPF (as adjudicated by the CEC) will be summarized by treatment group and analyzed via Poisson regression in a similar manner as described in Section 11.2.1.

Annualized rate of acute exacerbation will be calculated separately for each patient by diving the number of exacerbations by the time on study in years, as well as estimated by Poisson regression with each subject's time on study as offset.

Additionally, time to first investigator-reported acute exacerbation of IPF and multiple investigator-reported exacerbations of IPF will also be summarized and analyzed.

11.2.3 Overall Survival at Week 52

Overall survival at Week 52 will be summarized by treatment group and the time to death, as defined in Section 9.5.5, will be analyzed via K-M methodology and Cox regression similar to that described in Section 11.2.1.

Overall survival at the end of the study will include all deaths collected through the end of the study and will be analyzed by treatment group. The time to death will be calculated as defined in Section 9.5.5, but will include deaths that occurred beyond (365 + 7) days after randomization, and will be analyzed using Kaplan-Meier method and Cox regression, similar to that described in Section 11.2.1.

For respiratory-related death, proportion of subjects with respiratory-related death will be summarized by treatment group and compared with a chi-square test. It will also be analyzed via K-M methodology and Cox regression.

11.2.4 Change in % Predicted FVC at Week 52

After applying the same strategies for intercurrent events as in the primary analysis for the primary endpoint, change from baseline in % predicted FVC (with acceptable or borderline acceptable grades) at Week 52 will be summarized by treatment group and analyzed in a similar manner as described in Section 11.1.2 for primary efficacy analysis; with change from baseline in % predicted FVC as the dependent variable and baseline % predicted FVC and background therapy as the covariates in MMRM model or background therapy as strata in the non-parametric aligned rank stratified Wilcoxon test.

The analysis of change from baseline in % predicted FVC will be repeated using alternative composite strategies for death. Specifically, for post-death scheduled visits, the change from baseline in % predicted FVC will be assigned using 3 different extents of penalties – the 0.1st, 5th, and 10th percentiles of all observed values (change from baseline in % predicted FVC) across both treatment groups at each visit. In addition, an extreme unfavorable value of 0 will be assigned as their % predicted FVC at Week 52 for subjects who died before Week 52. The

MMRM model and the non-parametric aligned rank stratified Wilcoxon test described in Section 11.1.2 will be applied.

As described above in Figure 11-1, each endpoint nested below Change in % Predicted FVC at Week 52 will be tested at a two-sided alpha level equal to 0.05. No adjustments for multiplicity of these other endpoints will be done. For simplicity, the analyses will just be using all observed data (including those observed after study treatment discontinuation).

Percent change from baseline in FVC to Week 52 will also be summarized and then analyzed by MMRM with background therapy as covariate.

The proportions of subjects who experience >5% decline in FVC or >10% decline in FVC will also be summarized separately by treatment group and compared between treatment groups by chi-square tests.

11.2.5 Change in K-BILD Total Score and Domain Scores at Week 52

The K-BILD is a self-administered health status questionnaire that comprises 15 items and a 7-point Likert response scale (Patel 2012). It has 3 domains: psychological, breathlessness and activities, and chest symptoms as well as a total score. The K-BILD domain and total score ranges are 0 to 100; 100 represents the best health status. The K-BILD questionnaire will be scored according to the developers' recommendations. K-BILD total and domain scores will be listed and changes from baseline in K-BILD total and domain scores at Week 52 will be summarized by treatment group.

For intercurrent events, treatment policy will be applied for early study treatment discontinuation, initiation or change of background therapy during the study, initiation of rescue medication, and initiation of prostacyclin therapy (other than commercial inhaled treprostinil products); hypothetical strategy will be applied for elective lung transplantation and initiation of commercial inhaled treprostinil products; and composite strategy will be applied for death. That is, for subjects who died before Week 52, their missing K-BILD total and 3 domain change from baseline values at scheduled visits on or after the death date will be assigned the 2.5th percentile of all observed values (change from baseline) across both treatment groups at each respective visit. Change from Baseline K-BILD total score and the 3 domain scores will be analyzed

separately either by MMRM methodology or by the non-parametric aligned rank stratified Wilcoxon test with the corresponding aligned rank Hodges-Lehmann estimate, depending on whether the assumptions of normality and homoscedasticity are met. The analysis will follow the approach described in Section 11.1.2 with the appropriate changes to the dependent variable and corresponding baseline covariate.

Sensitivity analyses of the K-BILD total and 3 domain scores will be performed using alternative composite strategies for death. Specifically, an extreme unfavorable value of 0 will be assigned as the Week 52 K-BILD total and the 3 domain scores for subjects who died before Week 52. In this case, both the MMRM approach and the non-parametric aligned rank stratified Wilcoxon test will be used. In addition, a hypothetical strategy for death will also be applied and change in K-BILD scores will be analyzed by MMRM.

11.2.6 Change from Baseline in DLCO at Week 52

Change from Baseline in DLCO at Week 52 will be summarized by treatment group and analyzed by MMRM or by the non-parametric aligned rank stratified Wilcoxon test depending on whether the assumptions of normality and homoscedasticity are met. The analyses will include all DLCO data regardless of grade as provided by the central DLCO review and will follow the approach described in Section 11.1.2 with change from Baseline in DLCO as the dependent variable and with Baseline DLCO and background therapy as the covariates. The same strategies as for K-BILD data will be used to handle intercurrent events before the analysis.

As sensitivity analyses, change from baseline in DLCO at Week 52 including only Acceptable and Borderline Acceptable DLCO data will be analyzed. Additional analyses of change from baseline in DLCO at Week 52 will be repeated using the same strategy for intercurrent events except for death, which will be handled by hypothetical strategy.

11.3 EXPLORATORY EFFICACY MEASURES

Exploratory endpoints include the following:

1. Change from baseline in absolute FVC at Weeks 16, 28, and 40
2. Change from baseline in NT-proBNP at Week 52
3. Change from baseline in resting supplemental oxygen use at Week 52
4. Evaluation of pharmacogenomics and biomarkers for subjects who consent to collection

11.3.1 Change from Baseline in Absolute FVC at Weeks 16, 28, and 40

The changes from baseline in absolute FVC at Weeks 16, 28, and 40 will be estimated and compared between treatment groups based on the MMRM model for change from baseline in FVC at Week 52 described in Section 11.1.2. The same strategies as for sensitivity analyses of K-BILD data will be used to handle intercurrent events before the analysis.

Additionally, other pulmonary function parameters, such as FEV1, % predicted FEV1, % predicted FVC, FEV1/FVC ratio, DLCO, and % predicted DLCO that are available will be summarized by treatment groups and by visits.

11.3.2 Change from Baseline in NT-proBNP at Week 52

Since it is expected that the NT-proBNP data will follow a log-normal distribution, these data will be log-transformed prior to analysis. Change from baseline in NT-proBNP at Week 52 will be summarized by treatment group and analyzed by MMRM as described in Section 11.1.2 with change from baseline in log-transformed NT-proBNP as the dependent variable and with log-transformed baseline NT-proBNP and background therapy as the covariates. For simplicity, the analyses will just be using all observed data (including those observed after study treatment discontinuation).

11.3.3 Change from Baseline in Resting Supplemental Oxygen at Week 52

Change from baseline in resting supplemental oxygen (L/min) at Week 52 will be summarized by treatment group and analyzed by MMRM as described in Section 11.1.2 with change from baseline in resting supplemental oxygen as the dependent variable and baseline resting supplemental oxygen and background therapy as the covariates. For simplicity, the analyses will just be using all observed data (including those observed after study treatment discontinuation).

Additionally, the use of supplemental oxygen (yes vs no) at each visit will be summarized by treatment group and analyzed by generalized estimating equations methodology. The model will include terms for visit, background therapy, treatment group, and interaction of the treatment group and visit assuming exchangeable correlation. Sample SAS code to implement generalized estimating equations follows:

PROC GENMOD;

CLASS TRT01P BGTX AVISITN USUBJID;

*MODEL AVALC = TRT01P BGTX AVISITN TRT01PN*AVISITN/ LINK=Logit EXP DIST=BIN;*

REPEATED SUBJECT=USUBJID / TYPE=EXCH WITHINSUBJECT=AVISITN;

RUN;

where USUBJID is subject identifier, TRT01P is treatment group, BGTX is background therapy, AVISITN is study week, and AVALC is the use of supplemental oxygen (yes or no).

11.3.4 Evaluation of Pharmacogenomics and Biomarkers

The evaluation of pharmacogenomics and biomarkers will be detailed in a separate SAP.

12 SAFETY ANALYSES

All safety analyses will be performed only on the Safety Population (see Section 8).

12.1 EXTENT OF EXPOSURE

All study drug dosing will be listed for all subjects. The listing will include initial dose (number of breaths for each session and total number of breaths per day), date of initial dose, and dose and date of each subsequent dose change. Dosing at Baseline and Weeks 4, 8, 16, 28, 40, and 52 will be summarized for subjects still receiving study drug at each of these assessments. The summary will include both the numeric summaries and the categorical summaries for number of breaths per session. If study drug is dosed differently across different sessions on the same day, the maximum number of breaths is used for summarization. Summary of overall duration of exposure will be included as well as the final dose (breaths/session) and the maximum study drug dose (breaths/session) reached for each subject (numerically and categorically).

12.2 ADVERSE EVENTS

All AEs will be coded to the appropriate PT and SOC using the current version of Medical Dictionary for Regulatory Activities. AEs will be listed by treatment group including all details recorded on the eCRF plus an indicator of whether the event was treatment-emergent or on-treatment. The AE listings will include the AE verbatim term and its corresponding PT and SOC.

Treatment-emergent AEs are those AEs with an onset date equal to or after the start date of the study drug. The non-treatment-emergent AEs (the AEs occur after signing the Informed Consent Form but before receiving study drug) will be listed but not included in summary tables. On-treatment AEs are those AEs with an onset date within 30 days after the last dose date.

For primary analyses, the AE summaries will be limited to include only treatment-emergent on-treatment AEs. All summaries will include the number and percentage of subjects experiencing each type of AE and the risk difference of AEs between treatment groups and its 95% CI, in order of risk difference and/or SOC. SAEs and non-SAEs will also be summarized by SOC and PT. No inferential analyses will be planned for AEs.

Treatment-emergent on-treatment AEs possibly or probably related to the study drug will be summarized. An AE summary by severity (mild, moderate, or severe) and by study drug dose at AE onset will also be provided. An AE summary by background therapy will also be provided.

Separate listings and summary tables will be provided for all SAEs, AEs leading to permanent discontinuation of study drug, and all deaths during the study (if data permit). For the listing of deaths, information for all subjects who die during the study period from randomization to the end of study (including those captured in the vital status form) will be included.

As sensitivity analyses, overall summary of treatment-emergent on-study AEs and summary of treatment-emergent on-study AEs by PT will also be provided. Treatment-emergent on-study AEs are any AE that occurred or worsened after the first dose of the treatment during the study. Any treatment-emergent AEs with onset date beyond 30 days of the last dose (if recorded) will be included in the summaries of on-study AEs.

12.3 CLINICAL LABORATORY PARAMETERS

Blood samples will be taken at Screening and/or Baseline, Weeks 4, 8, 16, 28, 40, and 52. All samples will be sent to a central laboratory for evaluation of clinical chemistry and hematology.

12.3.1 Clinical Chemistry

The following clinical chemistry parameters will be evaluated by the central laboratory:

Parameter	Units
Sodium	mmol/L
Potassium	mmol/L
Bicarbonate	mmol/L
Chloride	mmol/L
Total bilirubin	umol/dL
Alkaline phosphatase	U/L
Alanine aminotrasferase	U/L
Aspartate aminotransferase	U/L
Urea nitrogen	mmol/dL
Creatinine	umol/L
Calcium	mmol/L
Albumin	g/L

Values that are "High" or "Low" with respect to the reference range provided by the central laboratory will be flagged with an "H" or "L," respectively. All parameters will be listed for all subjects and assessments, along with their respective "High/Low" flags.

Values of these parameters at each visit, and their corresponding changes from Baseline will be descriptively summarized by treatment group.

For each parameter, the frequency and percentage of subjects within each treatment group who had "Low," "Normal," or "High" Baseline values, then subsequently had "Normal," "High," or "Low" follow-up values at each visit will be presented in a shift summary.

12.3.2 Hematology

The following hematology parameters will be evaluated by the central laboratory:

Parameter	Units
Hemoglobin	g/L
Hematocrit	Fraction of 1
Red blood cell count	10^{12}/L
White blood cell count	10^{9}/L
Platelet count	10^{9}/L

Values that are "High" or "Low" with respect to the reference range provided by the central laboratory will be flagged with an "H" or "L," respectively. All parameters will be listed for all subjects and assessments, along with their respective "High/Low" flags.

Numeric values of these parameters at each visit and their corresponding changes from Baseline will be summarized by treatment group.

For each parameter, the frequency and percentage of subjects within each treatment group who have "Low," "Normal," or "High" Baseline values, then subsequently have "Normal," "High," or "Low" follow-up values at each assessment will be presented in a shift summary.

12.4 VITAL SIGNS

All vital sign assessments will be listed for all subjects. This listing will include weight, heart rate, systolic and diastolic blood pressures, respiratory rate, temperature, and saturation of peripheral capillary oxygenation. The vital sign results at each assessment and changes from baseline for each post-Baseline assessment will be descriptively summarized by treatment group.

12.5 ELECTROCARDIOGRAMS

All ECG assessments will be listed for all subjects. This listing will include the heart rate, the ECG interval from the beginning of QRS complex to end of the T wave (QT interval), the QT interval corrected for heart rate (QT interval corrected for Fridericia's formula – QTcF, and QT interval corrected for Bazett's formula - QTcB), the time between P wave and beginning of QRS complex in ECG (PR interval), the electrocardiographic wave (QRS) duration, ECG results (normal/abnormal), whether there were clinically significant changes from Baseline, and if abnormalities were present with details and comments on any abnormalities. The ECG results at Baseline and changes at Weeks 28 and 52 will be descriptively summarized by treatment group. In addition, for the QT interval corrected for heart rate using the Fridericia method, the number and percent of subjects with values ≥500 msec and the number and percent of subjects with changes from Baseline of <30 msec, 30 to <60 msec, and ≥60 msec will be presented. Additionally, each abnormality will be summarized by number and percentage of subjects reporting at each visit by treatment group.

13 REFERENCES

Graham BL, Brusasco V, Burgos F, et al. 2017 ERS/ATS standards for single-breath carbon monoxide uptake in the lung. Eur Respir J. 2017;49(1). Erratum in: Eur Respir J. 2018;52(5).

Graham BL, Steenbruggen I, Miller MR, et al. Standardization of spirometry 2019 update. An official American Thoracic Society and European Respiratory Society technical statement. Am J Respir Crit Care Med. 2019;200(8):e70-e88.

Knezevic A, Patil S. Combination weighted log-rank tests for survival analysis with non-proportional hazards. SAS Global Forum 2020. Paper 5062-2020. https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2020/5062-2020.pdf

Lin RS, Lin J, Roychoudhury S, et al. Alternative analysis methods for time to event endpoints under nonproportional hazards: a comparative analysis. Stat Biopharm Res. 2020;12(2):187-198.

Patel AS, Siegert RJ, Brignall K, et al. The development and validation of the King's Brief Interstitial Lung Disease (K-BILD) health status questionnaire. Thorax. 2012;67(9):804-810.

Royston P, Parmar MK. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. BMC Med Res Methodol. 2013;13:152.

Rubin, DB. Multiple Imputation for Nonresponse in Surveys. In: *Wiley Series in Probability and Statistics.* John Wiley & Sons, Inc; 1987:1.

Yuan Y. Sensitivity analysis in multiple imputation for missing data. SAS Institute 2014. Paper SAS270-2014. https://support.sas.com/resources/papers/proceedings14/SAS270-2014.pdf

14 APPENDICES

14.1 LIST OF TABLES

Table Number	Table Title
14.1.1	Summary of Subject Accountability
14.1.2	Summary of Analysis Population Information
14.1.3	Summary of Demographic Data
14.1.4	Summary of IPF History
14.1.5	Summary of Baseline Characteristics
14.1.6.1	Summary of Entry Criteria Violations
14.1.6.2	Summary of Protocol Deviations
14.1.7	Summary of Randomization Information
14.1.8	Summary of Medical History
14.1.9	Summary of Concomitant Medications Ongoing at Baseline
14.1.10	Summary of Concomitant Medications Added During the Study
14.1.11	Summary of Study Drug Dosing and Exposure
14.1.12	Summary of Study Drug Treatment Compliance
14.2.1.1	Summary and Analysis of FVC (mL) at Week 52 – Primary Efficacy Analysis
14.2.1.2	Summary and Analysis of FVC (mL) at Week 52 Using Control-based Pattern-mixture Model
14.2.1.3.1	Summary and Analysis of FVC (mL) at Week 52 – Sensitivity Analysis with a Penalty of 0.1st Percentile for Death
14.2.1.3.2	Summary and Analysis of FVC (mL) at Week 52 – Sensitivity Analysis with a Penalty of 5th Percentile for Death
14.2.1.3.3	Summary and Analysis of FVC (mL) at Week 52 – Sensitivity Analysis with a Penalty of 10th Percentile for Death
14.2.1.3.4	Summary and Analysis of FVC (mL) at Week 52 – Sensitivity Analysis with Extreme Penalty for Death
14.2.1.3.5	Summary and Analysis of FVC (mL) at Week 52 – Sensitivity Analysis with Half-minimum Penalty for Death
14.2.1.4	Summary and Analysis of FVC (mL) at Week 52 – Per-protocol Population
14.2.1.5.1	Summary and Analysis of FVC (mL) at Week 52 Including Acceptable or Borderline Acceptable Data
14.2.1.5.2	Summary and Analysis of FVC (mL) at Week 52 Including Only Acceptable Data (Grade A/B)
14.2.1.5.3	Summary and Analysis of FVC (mL) at Week 52 Including All Data
14.2.1.6	Summary and Analysis of On-treatment FVC (mL) at Week 52 Excluding Data Collected after Termination of Study Drug
14.2.1.7	Summary and Analysis of FVC (mL) at Week 52 – MMRM Model with Random Site Effects
14.2.1.8.1	Summary and Analysis of FVC (mL) at Week 52 by Background Therapy (Nintedanib/Pirfenidone vs No Background Therapy)
14.2.1.8.2	Summary and Analysis of FVC (mL) at Week 52 by Background Therapy (Nintedanib vs Pirfenidone vs No Background Therapy)
14.2.1.8.3	Summary and Analysis of FVC (mL) at Week 52 by Sex

Table Number	Table Title
14.2.1.8.4	Summary and Analysis of FVC (mL) at Week 52 by Age Group
14.2.1.8.5	Summary and Analysis of FVC (mL) at Week 52 by Time Since IPF Diagnosis
14.2.1.8.6	Summary and Analysis of FVC (mL) at Week 52 by Obstructive Sleep Apnea
14.2.1.8.7	Summary and Analysis of FVC (mL) at Week 52 by Smoking History
14.2.1.8.8	Summary and Analysis of FVC (mL) at Week 52 by Corticosteroid Use
14.2.1.8.9	Summary and Analysis of FVC (mL) at Week 52 by Last Study Drug Dose
14.2.1.8.10	Summary and Analysis of FVC (mL) at Week 52 by Supplemental Oxygen Use at Baseline
14.2.1.8.11	Summary and Analysis of FVC (mL) at Week 52 by Race
14.2.1.8.12	Summary and Analysis of FVC (mL) at Week 52 by Ethnicity
14.2.1.8.13	Summary and Analysis of FVC (mL) at Week 52 by BMI at Baseline
14.2.1.9	Summary and Analysis of FVC (mL) at Week 52 by Subgroups
14.2.1.10	Summary and Analysis of Annual Rate of Change in FVC (mL) Using Random Coefficients
14.2.1.11	Summary and Analysis of FVC (mL) – Ranked ANCOVA
14.2.1.12	Summary and Analysis of FVC (mL) – While-alive Analysis Using ANCOVA
14.2.1.13	Summary of the Number of Unacceptable FVC Values
14.2.2.1	Summary and Analysis of Clinical Worsening Events – Secondary Efficacy Endpoint
14.2.2.2.1	Summary and Analysis of Acute Exacerbations of IPF – Secondary Efficacy Endpoint
14.2.2.2.2	Summary and Analysis of Investigator-reported Acute Exacerbations of IPF
14.2.2.3.1	Summary and Analysis of Overall Survival at Week 52 – Secondary Efficacy Endpoint
14.2.2.3.2	Summary and Analysis of Overall Survival at the End of the Study
14.2.2.4	Summary and Analysis of % Predicted FVC at Week 52 – Secondary Efficacy Endpoint
14.2.2.5.1	Summary and Analysis of K-BILD Total Score at Week 52 – Secondary Efficacy Endpoint
14.2.2.5.2	Summary and Analysis of K-BILD Domain Scores at Week 52
14.2.2.6	Summary and Analysis of DLCO at Week 52 – Secondary Efficacy Endpoint
14.2.2.7	Summary and Analysis of On-treatment Clinical Worsening Events
14.2.2.8.1	Summary and Analysis of On-treatment Acute Exacerbation of IPF
14.2.2.8.2	Summary and Analysis of On-treatment Investigator-reported Acute Exacerbation of IPF
14.2.2.9	Summary and Analysis of On-treatment Respiratory Hospitalization
14.2.2.10	Summary and Analysis of On-treatment Survival
14.2.2.11	Summary and Analysis of Respiratory-related Hospitalizations
14.2.2.12	Summary of Analysis of Time to Death (Death as the First Clinical Worsening Event)
14.2.2.13	Summary and Analysis of Respiratory-related Deaths
14.2.2.14	Summary and Analysis of All-cause Hospitalization and Death

Table Number	Table Title
14.2.2.15.1	Summary and Analysis of % Predicted FVC at Week 52 with a Penalty of 0.1st Percentile for Death
14.2.2.15.2	Summary and Analysis of % Predicted FVC at Week 52 with a Penalty of 5th Percentile for Death
14.2.2.15.3	Summary and Analysis of % Predicted FVC at Week 52 with a Penalty of 10th Percentile for Death
14.2.2.15.4	Summary and Analysis of % Predicted FVC at Week 52 with Extreme Penalty for Death
14.2.2.16.1	Summary and Analysis of Categorized Percent Changes in FVC
14.2.2.16.2	Summary and Analysis of Percent Change from Baseline to Week 52 in FVC
14.2.2.17.1	Summary and Analysis of K-BILD Total Score at Week 52 – with Extreme Penalty for Death
14.2.2.17.2	Summary and Analysis of K-BILD Domain Score at Week 52 – with Extreme Penalty for Death
14.2.2.18.1	Summary and Analysis of K-BILD Total Score at Week 52 (Death Handled by Hypothetical Strategy)
14.2.2.18.2	Summary and Analysis of K-BILD Domain Scores at Week 52 (Death Handled by Hypothetical Strategy)
14.2.2.19	Summary and Analysis of DLCO at Week 52 – with Extreme Penalty for Death
14.2.2.20	Summary and Analysis of DLCO at Week 52 (Death Handled by Hypothetical Strategy)
14.2.2.21	Summary and Analysis of DLCO at Week 52 Including Only Acceptable and Borderline Acceptable Data (Death Handled by Hypothetical Strategy)
14.2.3.1	Summary and Analysis of FVC (mL) up to Week 52
14.2.3.2	Summary and Analysis of NT-proBNP (pg/mL) Data at Week 52
14.2.3.3.1	Summary and Analysis of Resting Supplemental Oxygen Use (L/min) at Week 52
14.2.3.3.2	Summary and Analysis of Resting Supplemental Oxygen Use (Yes/No)
14.2.4.1.1	Summary of Pulmonary Function Test Results – Observed Case Including Only Acceptable or Borderline Acceptable Data
14.2.4.1.2	Summary of Pulmonary Function Test Results – Observed Case
14.2.4.1.3	Summary of Pulmonary Function Test Results by IPF Background Therapy – Observed Case Including Acceptable or Borderline Acceptable Data
14.2.4.1.4	Summary of Pulmonary Function Test Results by IPF Background Therapy – Observed Case
14.2.4.2	Summary of K-BILD – Observed Case
14.2.4.3	Summary of NT-proBNP – Observed Case
14.2.4.4	Summary of Resting Supplemental Oxygen Use – Observed Case
14.3.1.1	Overall Summary of Adverse Events
14.3.1.2	Summary of Adverse Events by System Organ Class and Preferred Term
14.3.1.3	Summary of Adverse Events by Preferred Term
14.3.1.4	Summary of Deaths
14.3.1.5	Summary of Serious Adverse Events by System Organ Class and Preferred Term
14.3.1.6	Summary of Non-serious Adverse Events by System Organ Class and Preferred Term

Table Number	Table Title
14.3.1.7	Summary of Adverse Events Probably or Possibly Related to Study Drug by System Organ Class and Preferred Term
14.3.1.8	Summary of Adverse Events Leading to Permanent Discontinuation of Study Drug by Preferred Term
14.3.1.9	Summary of Adverse Events by System Organ Class, Preferred Term, and Severity
14.3.1.10	Summary of Adverse Events by System Organ Class, Preferred Term, and Study Drug Dose at AE Onset
14.3.1.11	Overall Summary of Adverse Events by IPF Background Therapy
14.3.1.12	Summary of Adverse Events by Preferred Term by IPF Background Therapy
14.3.1.13	Overall Summary of On Study Adverse Events
14.3.1.14	Summary of On Study Adverse Events by Preferred Term
14.3.2.1	Listing of Deaths
14.3.2.2	Listing of Serious Adverse Events
14.3.4.1	Summary of Laboratory Data – Clinical Chemistry
14.3.4.2	Summary of Clinical Chemistry Shift from Baseline
14.3.4.3	Summary of Laboratory Data – Hematology
14.3.4.4	Summary of Hematology Shift from Baseline
14.3.4.5	Summary of Vital Signs
14.3.4.6	Summary of Electrocardiogram Parameters
14.3.4.7	Summary of Electrocardiogram Categorical Results

14.2 LIST OF LISTINGS

Listing Number	Listing Title
16.2.1.1	Subject Disposition
16.2.1.2	Subject Accountability
16.2.2	Protocol Deviations
16.2.3.1	Subject Randomization
16.2.3.2	Analysis Population Information
16.2.3.3	Entry Criteria
16.2.4.1	Demographics
16.2.4.2	IPF History
16.2.4.3	Medical History
16.2.4.4	Concomitant Medications
16.2.5	Study Drug Dosing
16.2.6.1.1	Spirometry
16.2.6.1.2	Lung Diffusion Capacity
16.2.6.2	Clinical Worsening Events
16.2.6.3	Exacerbations of IPF
16.2.6.4	Deaths
16.2.6.5	King’s Brief Interstitial Lung Disease Questionnaire
16.2.6.6	Plasma NT-proBNP (pg/mL)

Listing Number	Listing Title
16.2.6.7	Supplemental Oxygen Use
16.2.7.1	Adverse Events
16.2.7.2	Adverse Events Leading to Discontinuation of Study Drug
16.2.8.1	Laboratory Results – Clinical Chemistry
16.2.8.2	Laboratory Results – Hematology
16.2.8.3	Pregnancy Status
16.2.8.4	Vital Signs
16.2.8.5	ECG Results

14.3 LIST OF FIGURES

Figure Number	Figure Title
14.2.1.1	Mean Change from Baseline in FVC (mL) by Visit
14.2.1.2	Diagnostics for Residuals from MMRM for Change from Baseline in FVC (mL) at Week 52 – Primary Efficacy Analysis
14.2.1.3	Diagnostics for Residuals from MMRM for Change from Baseline in FVC (mL) at Week 52 – Sensitivity Analysis with a Penalty of 10th Percentile for Death
14.2.2	Histogram of Annual Rate of Change in FVC (mL)
14.2.3	Forest Plot on Subgroup Analyses of FVC (mL) at Week 52
14.2.4	Kaplan-Meier Plot of Time to First Clinical Worsening
14.2.5.1	Kaplan-Meier Plot of Time to First Acute Exacerbation of IPF
14.2.5.2	Kaplan-Meier Plot of Time to First Investigator-reported Acute Exacerbation of IPF
14.2.6	Kaplan-Meier Plot of Time to First Respiratory Hospitalization
14.2.7.1	Kaplan-Meier Plot of Time to Death at Week 52
14.2.7.2	Kaplan-Meier Plot of Time to Death at the End of the Study
14.2.8	Kaplan-Meier Plot of Time to Respiratory-related Death
14.2.9.1	Mean Change from Baseline in % Predicted FVC by Visit
14.2.9.2	Diagnostics for Residuals from MMRM for Change from Baseline in % Predicted FVC at Week 52
14.2.10.1	Mean Change from Baseline in K-BILD Total Score by Visit
14.2.10.2	Diagnostics for Residuals from MMRM for Change from Baseline in K-BILD Total Score at Week 52
14.2.11.1	Mean Change from Baseline in DLCO by Visit
14.2.11.2	Diagnostics for Residuals from MMRM for Change from Baseline in DLCO at Week 52

15 SUMMARY OF CHANGES FROM PREVIOUS VERSIONS

Version	Version Date	Section	Description of Change(s)
Amendment 3	Aug 2025	Section 2.1	Updated the composite strategy for the intercurrent event of death for the primary endpoint. For subjects who died before Week 52, their missing FVC change from baseline values at scheduled visits on or after the death date will be assigned the 2.5th percentile of all observed values (FVC change from baseline) across both treatment groups at each visit.
		Section 2.2	Updated the composite strategy for the intercurrent event of death for the continuous secondary endpoints to align with the primary endpoint.
		Section 8	Update Per-protocol Population
		Section 9.5.5	Clarified Table 9-5 that subjects completed or discontinued from study without death will have their time to death censored at (365+7) days after randomization.
		Section 11.1.1	Updated the composite strategy for the intercurrent event of death for the primary endpoint. For subjects who died before Week 52, their missing FVC change from baseline values at scheduled visits on or after the death date will be assigned the 2.5th percentile of all observed values (FVC change from baseline) across both treatment groups at each visit.
		Section 11.1.3	Added sensitivity analyses for FVC using 3 different extents of penalties for death - the 0.1st, 5th, and 10th percentiles. Changed the composite strategy of assigning 0 (previously primary strategy for death) to a sensitivity analysis. Added 2 supplementary analyses, a ranked ANCOVA and a “While Alive” analysis.
		Section 11.2	Added “Overall survival at the end of the study” that will include all deaths collected through the end of the study.
		Section 11.2.4	Added sensitivity analyses for % predicted FVC using 3 different extents

Version	Version Date	Section	Description of Change(s)
			of penalties – the 0.1st, 5th, and 10th percentiles for death. Added a composite strategy of assigning 0 for death as a sensitivity analysis.
		Section 11.2.5	Updated the composite strategy for the intercurrent event of death for the K-BILD total and domain scores using 2.5th percentile of all observed change from baseline values (to align with the primary endpoint). Updated a composite strategy of assigning 0 as sensitivity analysis.
Amendment 2	21 APR 2025	Section 2.2	Changed the hypothetical strategy for the intercurrent event of death for change from baseline in K-BILD score at Week 52 and change from baseline in DLCO at Week 52 to a composite strategy (same as for primary efficacy endpoint).
		Section 9.2	Added BMI at Baseline to subgroup analyses.
		Section 9.5.1	Added a more detailed description of the FVC grading system per American Thoracic Society guidelines in new Table 9-1.
		Section 9.5.2	New section. Added a description of lung diffusion capacity per protocol.
		Sections 9.5.3, 9.5.4, and 9.5.5	Updated the censoring rules in the tables of derivation of time to events.
		Section 11.1.2	Updated the strategies for intercurrent events from using hypothetical strategy for initiation of prostacyclin therapy to using hypothetical strategy for initiation of commercial inhaled treprostinil products. Added the summary of the number of Unacceptable FVC values. Specified which normality tests will be used after MMRM model.
		Section 11.1.3	Added BMI at Baseline to subgroup analyses.
		Section 11.2	Added analyses for all-cause hospitalization and death and annualized rate of acute exacerbation to Figure 11-1.

Version	Version Date	Section	Description of Change(s)
			Added the analysis of investigator-report acute exacerbation of IPF.
		Section 11.2.1	Added analyses of all-cause hospitalization and death.
		Section 11.2.2	Added the calculation of annualized rate of acute exacerbation.
		Section 11.2.5	Added text specifying that either the MMRM methodology or the non-parametric aligned rank stratified Wilcoxon test will be used, depending on whether the assumptions of normality and homoscedasticity are met. Added sensitivity analyses for change from baseline K-BILD total score and domain scores, with death handled by hypothetical strategy.
		Section 11.2.6	Added text specifying that either the MMRM methodology or the non-parametric aligned rank stratified Wilcoxon test will be used, depending on whether the assumptions of normality and homoscedasticity are met. Added sensitivity analyses for change from baseline in DLCO at Week 52 including only Acceptable and Borderline Acceptable DLCO data. Added sensitivity analyses for change from baseline DLCO using the same strategies applied in the sensitivity analyses of K-BILD.
		Section 15	New section. Added Summary of Changes section.
Amendment 1	12 DEC 2023	General	Updated the SAP to align up with latest protocol amendment 4 (26JULY2023).
		Sections 2.1, 2.2 and Section 11.1.1	Added details on estimand for efficacy endpoints and the intercurrent event strategies.
		Section 7	Updated the sample size to be consistent with study protocol amendment 4.
		Sections 9.1, 9.2	Specified 3 levels of background therapy to be used as covariate.

Version	Version Date	Section	Description of Change(s)
			Clarified the categories for the subgroup analyses.
		Section 9.4	Clarified the multiplicity adjustment procedure for secondary endpoints.
		Section 11.1.2	Added alternative methods for primary analyses in case assumptions for the primary MMRM model are not satisfied.
		Section 11.1.3	Added random coefficient model as sensitivity analysis for FVC.
		Section 11.2	Clarified the exploratory analyses for secondary endpoints will not be included in the multiplicity adjustment procedure.
		Section 12.2	Clarified on-treatment treatment-emergent AE to differentiate from on-study AEs.
Original	3 MAR 2023	General	Based on protocol amendment 3 (25APR2022).