NCT #: NCT04419558

# STATISTICAL ANALYSIS PLAN

| | |
|---|---|
| **STUDY TITLE:** | A Phase 3, Randomized, Double-Blind, Placebo-Controlled Efficacy and Safety Study of Pamrevlumab in Subjects with Idiopathic Pulmonary Fibrosis (IPF) |
| **PROTOCOL NUMBER:** | FGCL-3019-095 (ZEPHYRUS 2) |
| **AMENDMENT VERSION:** | 2.0 |
| **SAP VERSION:** | Final V1.0 (OLE Period) |
| **RELEASE DATE:** | 16Oct2023 |

# SIGNATURE PAGE

## Approvals

I have reviewed and accepted the information in this document to be a true and accurate representation of the Statistical Analysis Plan.

*Signature: see appended final page for 21CFR Part 11 compliant approval

**Initiator:**

[redacted] FibroGen, Inc.

**Reviewed by:**

[redacted], FibroGen, Inc.

[redacted], FibroGen, Inc.

[redacted], FibroGen, Inc.

[redacted] FibroGen, Inc.

[redacted], FibroGen, Inc.

[redacted], FibroGen, Inc.

[redacted], FibroGen, Inc.

## Signature Significance

The following significance is lent to the signatures on the Approvals page of this document.

| Signatory | Significance |
|---|---|
| **Initiator** | By signing, the author is attesting that the content of the document is complete and accurate. |
| **Reviewer** | By signing, the reviewer is attesting that the document's approach and contents are compliant with the study protocol, all appropriate, regulatory requirements, and other significant guidelines. This individual(s) has reviewed the document for accuracy and completeness. |

# CHANGE HISTORY

| Version | Date | Description |
|---|---|---|
| 1.0 | 16Oct2023 | Final Approved Version |
| | | |
| | | |
| | | |

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF ABBREVIATIONS

| Abbreviation | Explanation |
|---|---|
| ADA | Anti-Drug Antibody |
| AE | Adverse Event |
| BMI | Body Mass Index |
| CI | Confidence Interval |
| CRF | Case Report Form |
| DMC | Data Monitoring Committee |
| EDC | Electronic Data Capture |
| FVC | Forced Vital Capacity |
| FVCpp | Forced Vital Capacity percentage predicted |
| HAHA | Human Anti-Pamrevlumab Antibody |
| HRCT | High-Resolution Computed Tomography |
| IPF | Idiopathic Pulmonary Fibrosis |
| IV | Intravenous |
| L | Liter |
| mL | milliliter |
| OLE | Open Label Extension |
| PEY | Patient Exposure Years |
| PFT | Pulmonary Function Test |
| QLF | Quantitative Lung Fibrosis |
| SAF | Safety Analysis Set |
| SAP | Statistical Analysis Plan |
| TLF | Table, Listing, and Figure |

# 1. INTRODUCTION

This statistical analysis plan (SAP) is for the pre-specified reporting of study results for protocol FGCL-3019-095, amendment 2.0. This SAP documents the planned analyses only for the open-label extension (OLE) period. Specifications of tables, data listings, and figures (TLF) are contained in a separate document. A separate SAP details the statistical analyses and methods for the double-blind period of FGCL-3019-095.

In this document, only new definitions and rules will be described. Cross reference to the double-blind SAP will be made for definitions and rules that will follow the same methodology used for the double-blind period of the trial.

# 2. STUDY OBJECTIVE

The overall objective of this Phase 3 registration-enabling study is to evaluate the efficacy and safety of 30 mg/kg intravenous (IV) infusions of pamrevlumab as compared to placebo in subjects with Idiopathic Pulmonary Fibrosis (IPF). The objective of the OLE phase is to evaluate the long-term safety of pamrevlumab in subjects with IPF, though efficacy assessments are included.

# 3. STUDY DESIGN (OLE PERIOD)

## 3.1. Enrollment to OLE Period

Subjects who complete the Week 48 visit of the main study (regardless of the number of study drug infusions received) will be eligible to participate in the optional, OLE period of the study. The OLE period offers continuing access to pamrevlumab regardless of randomization assignment (pamrevlumab or placebo) in double-blind period of the study. Subjects will be considered as having entered the OLE study if they have received at least one dose of the study drug during the OLE period.

## 3.2. Randomization

Randomization is not applicable for the OLE period.

## 3.3. Study Periods

- Optional, OLE period of pamrevlumab:
  - Access to pamrevlumab will be available until the last subject completes 48 weeks of treatment in the OLE period, or pamrevlumab is commercially available for the indication of IPF, or the Sponsor decides to end the OLE period, whichever occurs first.
- Follow-up period/final safety assessments:
  - 28 days after the last dose
  - 60 days after the last dose: follow-up phone call, for a final safety assessment

# 4. STUDY ENDPOINTS (OLE)

The following efficacy and safety assessments will be collected only as standard-of-care per each investigational site's real-world practice and descriptively presented during the OLE period:

**Efficacy assessments:**

- Forced vital capacity (FVC) (L)
- FVC percent predicted (FVCpp)
- Absolute FVCpp decline of ≥10%
- All-cause mortality
- Change in high-resolution computed tomography (HRCT) Quantitative Lung Fibrosis (QLF) volume (if HRCT is performed as part of routine patient care)

**Safety assessments:**

Treatment-emergent adverse events (TEAEs) and treatment-emergent serious adverse events (TESAEs; including hypersensitivity/anaphylactic reactions) will be collected for 60 days after the last dose.

In addition, pre- and post-infusion vital signs and immunogenicity (Human Anti-Human Antibody [HAHA] formation) testing will be collected.

# 5. GENERAL STATISTICAL CONSIDERATIONS

## 5.1. Statistical Methodology

All data collected will be included in the data listings. All analyses will be performed using SAS® Version 9.3 or higher.

### 5.1.1. Analysis of Categorical Endpoints

Follow the same guidelines as described in SAP for the double-blind period.

### 5.1.2. Analysis of Continuous Endpoints

Follow the same guidelines as described in SAP for the double-blind period.

### 5.1.3. Analysis of Spirometry Data

Follow the same guidelines as described in SAP for the double-blind period.

### 5.1.4. Analysis of HRCT Data

Follow the same guidelines as described in SAP for the double-blind period.

## 5.2. Analysis Set

### 5.2.1. Safety Analysis Set - OLE

The **Safety Analysis Set - OLE (SAF-OLE)** includes all subjects who have received at least one study medication in the OLE period. Subjects will be analyzed according to the previous treatment actually received in the double-blind period.

All analysis in OLE will be performed on SAF-OLE unless otherwise specified.

### 5.2.2. Immunogenicity Analysis Set -OLE (IGS-OLE)

The IGS-OLE will include all subjects who are in SAF-OLE, have baseline antidrug antibody (ADA) assessment in the double-blind period, and have at least one evaluable ADA assessment in the OLE period.

## 5.3. General Data Handling Rules and Presentation Specifications

### 5.3.1. Analysis Period

- The primary analysis will be based on the on-study period which is defined as from Day 1 in the OLE period up to the end of the entire study, regardless of patients' treatment status.

### 5.3.2. Baseline Definitions

Baseline is defined as the last assessment prior to the first dosing in the OLE period with 2 exceptions below:

- Baseline in the double-blind period and OLE will be applied for assessments of FVC and QLF; and
- Baseline in the double-blind period will be applied for assessments of ADA.

### 5.3.3. Formulas

- Study Day Calculation
  - The day when a subject receives the first dose of study drug in the OLE period is designated as Day 1 in OLE.
  - Study day of an assessment/procedure is calculated as follows.
    - For assessments or procedures on Day 1 or later,
      *Study day = assessment/procedure date – Day 1 date + 1.*
    - For assessments or procedures earlier than Day 1,
      *Study day = assessment/procedure date – Day 1 date.*
- Follow the same formulas as described in SAP for the double-blind period:
  - Body Mass Index (BMI)
  - Body weight, height, and temperature
  - (Absolute) Change from Baseline

  - Relative change from baseline (%)
- Patient exposure years (PEY) (on pamrevlumab in OLE period):
  - PEY = Number of subjects × Average Duration of Exposure, where duration of exposure is (date of last non-zero dose on pamrevlumab – date of first dose in OLE period on pamrevlumab + 1)/365.25.

### 5.3.4. General Instructions of TLF

Follow the same guidelines as described in SAP for the double-blind period.

Summary tables will be presented for each previous actual treatment group in the double-blind period and all subjects pooled in the OLE period.

Efficacy and safety data will be summarized descriptively for subjects who entered OLE. No formal statistical inference will be performed.

### 5.3.5. Handling Dropouts and Missing Data

For handling of dropouts and missing data the same guidelines will be followed as described in SAP for the double-blind period. No imputations will be performed on post-death analysis visits.

## 5.4. Interim Analysis and Data Monitoring Committee

In addition to routine safety monitoring, an independent Data Monitoring Committee (DMC) will be established to review safety data on an ongoing basis. A DMC charter will establish the procedures, meeting frequency, and scope of responsibilities of the committee.

This study has no planned or pre-specified interim analysis for either efficacy or futility.

# 6. STATISTICAL ANALYSIS

## 6.1. Subject Accountability and Disposition

The number (%) of subjects who entered the OLE period, discontinued prematurely from treatment, discontinued prematurely from study, and reasons for premature discontinuation will be presented for each previous treatment group in the double-blind period and all subjects pooled.

Enrollment will also be summarized by study site.

## 6.2. Protocol Deviations

The number and percentage of subjects with important protocol deviations will be categorized and tabulated as appropriate. COVID-19 related important protocol deviations will be summarized in a separate table. All protocol deviations will be finalized prior to the OLE database lock.

## 6.3. Demographics

Demographic summaries will follow the same guidelines as described in SAP for the double-blind period.

## 6.4. Concomitant Medications

Concomitant medications summaries will follow the same guidelines as described in SAP for the double-blind period.

## 6.5. Treatments and Medications

### 6.5.1. Study Drug Exposure

Study treatment will be summarized for OLE period for SAF-OLE.

- Duration of exposure in the OLE period = date last non-zero dose in the OLE period – date of the first dose in the OLE period + 1.

Duration of study treatment in the OLE period will be summarized as a continuous variable and be tabulated by the categories as follows:

- <12 weeks
- 12 to <24 weeks
- 24 to <48 weeks
- 48 to <72 weeks
- 72 to <96 weeks
- ≥96 Weeks

Number of infusions is defined as the number of non-zero dose infusions.

The following dose administration parameters:

- Total number of infusions in the OLE period (1-4, 5-8, 9-16, 17-24, 25-32, ≥33)
- Overall average infusion doses (in mg/kg) in the OLE period
- Whether infusion was interrupted and reason for interruption in the OLE period

### 6.5.2. Treatment Compliance

Follow the same guidelines as described in SAP for the double-blind period.

## 6.6. Efficacy Analysis

All analyses for efficacy endpoints will be performed for the SAF-OLE during the on-study period unless noted otherwise.

### 6.6.1. Efficacy Endpoint Analysis and Estimand

#### 6.6.1.1. Estimand Strategy

The primary estimand is intended to provide a population-level assessment of the pamrevlumab on a continuous endpoint, regardless of participant compliance with the IP dosing.

#### 6.6.1.2. Population of Interest

All subjects in SAF-OLE.

#### 6.6.1.3. Intercurrent Event Handling Strategy

Treatment policy strategy will be used for any intercurrent event. No imputation on missing data.

#### 6.6.1.4. Analysis Variable

Follow guidelines in SAP for the double-blind period.

#### 6.6.1.5. Population Summary

Efficacy parameters, including change from baseline in FVC (L), FVCpp, and HRCT fibrosis score will be summarized descriptively at each scheduled visit as defined in Appendix 1.

## 6.7. Safety Analysis

The safety analyses will be performed for the SAF-OLE in the OLE period.

### 6.7.1. Adverse Events (AEs)

Follow the same guidelines as described in SAP for the double-blind period with additional PEY for OLE period on pamrevlumab will be provided.

### 6.7.2. Vital Signs

Follow the same guidelines as described in SAP for the double-blind period. The analysis visit window is defined in Appendix 1.

## 6.8. ADA Analysis

ADA analyses follow the same guidelines as described in SAP Part 1 for the double-blind period.

Considering ADA assessment was collected only on 28 days after last dose in the OLE period, ADA incidence will be summarized on IGS-OLE. A listing with details will be provided for SAF-OLE.

No subgroup analysis will be performed by ADA status.

## 6.9. PK Analysis

Not applicable for the OLE period.

# 7. CHANGES FROM PROTOCOL

Respiratory hospitalization and acute IPF exacerbation endpoints have been removed in this 095 OLE SAP, as adjudication of these events was only performed for the double-blind period of the study.

# 8. REFERENCE


- E9(R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials
- Jones PW, Forde, Y., St George's Respiratory Questionnaire Manual. http://www.healthstatus.sgul.ac.uk/SGRQ_download/SGRQ%20Manual%20June%202009.pdf. Accessed June 2, 2020
- Richeldi L, du Bois RM, Raghu G, et al. Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. N Engl J Med 2014; 370:2071-82.
- Noble PW, Albera C, Bradford WZ, Costabel U, etc.; Pirfenidone in patients with idiopathic pulmonary fibrosis (CAPACITY): two randomised trials; Lancet 2011; 377: 1760–69
- Statistical Review and Evaluation for Ofev® (nintedanib) 150 mg capsules; Indication: Treatment of Idiopathic Pulmonary Fibrosis (IPF); PDUFA: January 2, 2015
- King TE Jr., Bradford WZ, Castro-Bernardini S, etc.; A Phase 3 Trial of Pirfenidone in Patients with Idiopathic Pulmonary Fibrosis; N Engl J Med 2014;370:2083-92.

# APPENDIX 1. ANALYSIS VISIT WINDOWS

Analysis visits are defined by the windows that will have the widths of the corresponding assessments centered at the scheduled time. Unscheduled visits within a visit window defined below will be grouped into the closest scheduled visits based on the visit date.

**Table 1: Analysis Visit Window for Vital Signs**

| Analysis Visit | Target Day in OLE | Start Day | End Day |
|---|---|---|---|
| Baseline | See Section 5.3.2 | | |
| EX-Day 1 (postdose) [#] | 1 | 1 | 1 |
| EX-Week 3 (predose and postdose) | 22 | 2 | 32 |
| EX-Week 3*n (predose and postdose) | 21*n+1 | 21*n-9 | 21*n+11 |
| EX-28 Days post last dose [@] | | | ** |

** end of OLE period
[$] n is 2, 3, 4, etc.
[@] Same as visit collected on CRF.

**Table 2: Analysis Visit Window for PFT and Weight [#]**

| Analysis Visit | Target Day | Start Day | End Day |
|---|---|---|---|
| Baseline | See Section 5.3.2 | | |
| EX-Week 24 | 169 | 2 | 253 |
| EX-Week 24*n [$] | 24*n+1 | 24*7*n-82 | 24*7*n+85 |

[$] n is 2, 3, 4, etc.
Note: The timing of PFT is dependent on local schedules.
[#] Weight baselines follow same rules as vital signs.

**Table 3: Analysis Visit Window for HRCT**

| Analysis Visit | Target Day | Start Day | End Day |
|---|---|---|---|
| Baseline | See Section 5.3.2 | | |
| EX-Week 24 | 169 | 2 | 253 |
| EX-Week 24*n [$] | 24*n+1 | 24*7*n-82 | 24*7*n+85 |
| EX-28 days Post last dose [@] | | | ** |

** end of OLE period
[$] n is 2, 3, 4, etc.
[@] Same as visit collected on CRF.
Note: The timing of HRCT is dependent on local schedules.

# APPENDIX 2. ANALYSIS OF ASSOCIATION BETWEEN QLF AND OS

Cox proportional hazard (PH) analyses on the association between QLF and overall survival (OS) will be performed for the ITT population including both double-blind and OLE periods. The hazard ratio and its corresponding 95% CI between QLF Test Group vs QLF Reference Group. A stratified log-rank test stratified by randomization stratification factor will be used for treatment comparison.

No imputation will be performed on missing data. On-treatment analysis will be optional.

- QLF as a predictor:

1. To investigate whether baseline QLF is a significant predictor of mortality, a Cox PH model will include electronic data capture (EDC) GAP stage (I/II/III), and prior IPF treatment (Yes/No) as covariates, baseline QLF% (cutoff such as 15%, 20%, or 25%) as the predictor variable.
2. Similar analysis on baseline QLF volume (ml) will be performed as in 1), using baseline QLF volume (ml) (cutoff such as 300ml, 400ml, and 500ml) as the predictor variable.
3. In addition, Cox PH model including EDC GAP stage (I/II/III), and prior IPF treatment (Yes/No) as covariates, baseline QLF% or QLF volume (ml) (as a continuous variable) as the predictor variable, will be performed.

- Change from baseline in QLF at Week 24 as a predictor

1. In addition, to investigate the association between change from baseline in QLF% (baseline to Week 24) and OS, the Cox PH model will include EDC GAP stage (I/II/III), prior IPF treatment (Yes/No), and baseline QLF% (as a continuous variable) as covariates, and category in change from baseline in QLF% (e.g., ≥2% vs <2%) as the predictor variable.
2. Similar analysis on QLF volume (ml) change from baseline will be performed as in 4) with baseline QLF% replaced with baseline QLF volume as a continuous covariate, and change from baseline in QLF volume (ml) as the predictor variable with a cutoff (e.g., ≥40 mL vs <40mL) will be performed.

**STATISTICAL ANALYSIS PLAN**

# TITLE PAGE

| | |
|---|---|
| **STUDY TITLE:** | A Phase 3, Randomized, Double-Blind, Placebo-Controlled Efficacy and Safety Study of Pamrevlumab in Subjects with Idiopathic Pulmonary Fibrosis (IPF) |
| **PROTOCOL NUMBER:** | FGCL-3019-095 (ZEPHYRUS 2) |
| **STUDY SPONSOR:** | FibroGen, Inc.<br>409 Illinois Street<br>San Francisco, California 94158 USA |
| **STUDY DRUG:** | Pamrevlumab |
| **INDICATION:** | Idiopathic Pulmonary Fibrosis (IPF) |
| **PROTOCOL VERSION:** | Version 2.0 |
| **SAP VERSION** | Final V1.0 (Double-Blind Period) |
| **RELEASE DATE:** | 12Jun2023 |

# SIGNATURE PAGE

**Approvals**

I have reviewed and accepted the information in this document to be a true and accurate representation of the Statistical Analysis Plan.

*Signature: see appended final page for 21CFR Part 11 compliant approval

**Initiator:**

| |
|---|
| [redacted], FibroGen, Inc. |

**Reviewed by:**

| |
|---|
| [redacted], FibroGen, Inc. |
| [redacted], FibroGen, Inc. |
| [redacted] FibroGen, Inc. |
| [redacted] FibroGen, Inc. |
| [redacted], FibroGen, Inc. |
| [redacted], FibroGen, Inc. |
| [redacted], FibroGen, Inc. |
| [redacted] FibroGen, Inc. |

**Signature Significance**

The following significance is lent to the signatures on the Approvals page of this document.

| Signatory | Significance |
|---|---|
| **Initiator** | By signing, the author is attesting that the content of the document is complete and accurate. |
| **Reviewer** | By signing, the reviewer is attesting that the document's approach and contents are compliant with the study protocol, all appropriate, regulatory requirements, and other significant guidelines. This individual(s) has reviewed the document for accuracy and completeness. |

# CHANGE HISTORY

| Version | Date | Description |
|---|---|---|
| 1.0 | 12Jun2023 | Final Approved Version |

# TABLE OF CONTENTS

## LIST OF TABLES

# LIST OF FIGURES

# LIST OF ABBREVIATIONS

| Abbreviation | Explanation |
|---|---|
| ADA | Anti-Drug Antibody |
| AE | Adverse Event |
| ALP | Alkaline Phosphatase |
| ALT | Alanine Aminotransferase |
| ANC | Absolute Neutrophil Count |
| ANCOVA | Analysis of Covariance |
| ANOVA | Analysis of Variance |
| aPTT/PTT | Activated Partial Thromboplastin Time |
| AST | Aspartate Aminotransferase |
| ATC | Anatomical Therapeutic Class |
| ATS | American Thoracic Society |
| BMI | Body Mass Index |
| BTR | Best-Test Review |
| BUN | Blood Urea Nitrogen |
| CCL18 | Chemokine (C-C Motif) Ligand 18 |
| CDE | Center for Drug Evaluation |
| CI | Confidence Interval |
| CRF | Case Report Form |
| CRP | C-Reactive Protein |
| CSR | Clinical Study Report |
| CTCAE | Common Terminology Criteria for Adverse Events |
| CTGF | Connective Tissue Growth Factor |
| DBL | Database Lock |
| DLCO | Diffusion capacity of the Lungs for Carbon Monoxide |
| DMC | Data Monitoring Committee |
| ECG | Electrocardiogram |
| EDC | Electronic Data Capture |
| EOS | End of Study |
| EOT | End of Treatment |
| ERS | European Respiratory Society |
| FVC | Forced Vital Capacity |
| FVCpp | Forced Vital Capacity percentage predicted |
| GAP | GAP index and staging system for IPF: gender (G), age (A), and two pulmonary physiological parameters (P) - percentage predicted FVC [%], and percentage predicted DLCO [%] |
| GGT | Gamma-Glutamyltransferase |
| HAHA | Human Anti-Pamrevlumab Antibody |
| HRCT | High-Resolution Computed Tomography |
| ICF | Informed Consent Form |
| ICH E9 | International Conference on Harmonization Statistical Principles for Clinical Trials |
| IGS | Immunogenicity Analysis Set |

| Abbreviation | Explanation |
|---|---|
| IL-6 | Interleukin-6 |
| IL-8 | Interleukin-8 |
| INR | International Normalised Ratio |
| IP | Investigational Product |
| IPD | Important Protocol Deviations |
| IPF | Idiopathic Pulmonary Fibrosis |
| IRT | Interactive Response System |
| ITT | Intent-To-Treat |
| IV | Intravenous |
| KL-6 | Krebs von den Lungen 6 |
| L | Liter |
| LCQ | Leicester Cough Questionnaire |
| LSMean(s) | Least-Square Mean(s) |
| MAR | Missing At Random |
| MCMC | Markov Chain Monte Carlo |
| MedDRA | Medical Dictionary for Regulatory Activities |
| MI | Multiple Imputation |
| mL | milliliter |
| MMP7 | Matrix MetalloProteinase-7 |
| MMRM | Mixed Model for Repeated Measures |
| NMPA | China's National Medical Products Administration |
| MNAR | Missing Not at Random |
| OLE | Open Label Extension |
| OS | Overall Survival |
| PCS | Potentially Clinically Significant |
| PE | Physical Examination |
| PFT | Pulmonary Function Test |
| PIIINP | Type III collagen |
| PINP | Procollagen type I N-terminal Propeptide |
| PK | Pharmacokinetic |
| PopPK | Population Pharmacokinetic |
| PPS | Per-Protocol Set |
| PROs | Patient Reported Outcomes |
| PT | Prefer Term |
| Q3W | Every 3 Weeks |
| QLF | Quantitative Lung Fibrosis |
| RBC | Red Blood Cell, or Erythrocyte |
| RCM | Random Coefficient Model |
| SAE | Serious Adverse Event |
| SAF | Safety Analysis Set |
| SAP | Statistical Analysis Plan |
| SD | Standard Deviation |
| SE | Standard Error |

| Abbreviation | Explanation |
|---|---|
| SGRQ | St. George's Respiratory Questionnaire |
| SOC | System Organ Class |
| TEAE | Treatment-Emergent Adverse Event |
| TESAE | Treatment-Emergent Serious Adverse Event |
| TGF-beta | Transforming Growth Factor beta |
| TLC | Total Lung Capacity Volume |
| TLF | Table, Listing, and Figure |
| TNF | Tumor Necrosis Factor |
| UCSD-SOBQ | University of California San Diego – Shortness of Breath Questionnaire |
| ULN | Upper Limit of Normal, value provided by the laboratory |
| VEGF | Vascular Endothelial Growth Factor |
| WHODD | World Health Organization Drug Dictionary |
| WOCBP | Women Of Childbearing Potential |

# 1. INTRODUCTION

This statistical analysis plan (SAP) is for the pre-specified reporting of study results for protocol FGCL-3019-095, amendment 2.0. Specifications of tables, data listings, and figures (TLF) are contained in a separate document. The statistical analyses and summary tabulations described in this SAP will provide the basis for the results section of the clinical study report (CSR). The SAP will be finalized prior to unblinding after the study Database Lock (DBL). Any major modification to this SAP after the signoff will be documented in an SAP amendment or CSR.

This SAP describes the planned statistical analyses for the double-blind period of the study. Statistical analyses for the main study cohort of subjects are planned for inclusion in the body of the CSR.

The open-label extension period (OLE) of the study will have a separate SAP. Additionally, a population pharmacokinetic (PopPK) analysis as well as an exposure-response analysis will be defined in a separate pharmacokinetic (PK) analysis plan.

Based on regional regulatory filing needs, region-specific subset analyses may be performed following the methodology outlined in this SAP. Refer to Appendix 7 if regulatory submission in China is pursued.

# 2. STUDY OBJECTIVE

The overall objective of this trial is to evaluate the efficacy and safety of 30 mg/kg intravenous (IV) infusions of pamrevlumab as compared to placebo in subjects with Idiopathic Pulmonary Fibrosis (IPF).

# 3. STUDY DESIGN

## 3.1. Overview

This is a Phase 3, randomized, double-blind, placebo-controlled multi-center trial to evaluate the efficacy and safety of pamrevlumab in subjects with IPF. Approximately 340 subjects will be enrolled in this trial.

## 3.2. Sample Size

A sample size of 340 subjects will have at least 90% power, based on a two-sided alpha level of 0.05 for a two-sample t-test, to detect a treatment difference of 120 mL in the primary efficacy endpoint, change from baseline in forced vital capacity (FVC, in Liter: L), assuming a common standard deviation of 300 mL (allowing about 20% dropout rate).

The rationale for these sample size assumptions is as follows. In the phase 2 study FGCL-3019-067, a treatment difference of 178 mL was observed in change from baseline in FVC to Week 48. For this study, a conservative treatment difference of 120 mL is assumed. The standard deviation is estimated based on published results from other IPF studies.

## 3.3. Randomization and Treatment Assignment

Subjects will be randomized in a 1:1 ratio to one of the two study treatment arms as follows:

- Arm A: pamrevlumab, 30 mg/kg IV, Q3W (every 3 weeks)
- Arm B: matching placebo IV, Q3W

Randomization is stratified by GAP stage (I, II, III) derived from the GAP score obtained at screening (Refer to Appendix 6). Calculation of the GAP score for randomization is based on the Best-Test Review (BTR) assessment provided by the independent central spirometry vendor reader.

## 3.4. Study Periods

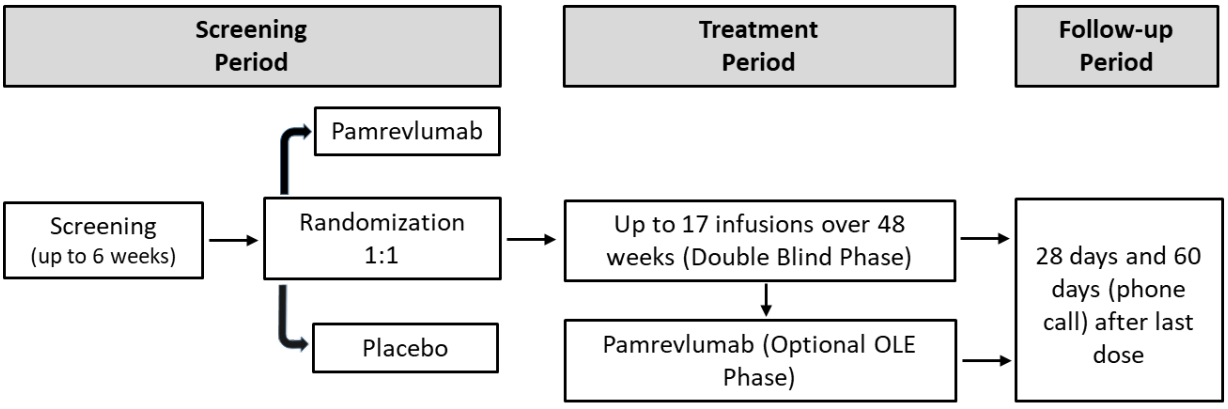
This study consists of the following study periods (Figure 1):

- Main (double-blind, placebo-controlled) study period:
  - Screening period: Up to 6 weeks
  - Treatment period: 48 weeks
- Optional, open label extension (OLE) period:

Access to pamrevlumab during the OLE period will be available until the last enrolled subject completes 48 weeks of treatment in the OLE period, or pamrevlumab is commercially available for the indication of IPF, or the Sponsor decides to end the OLE period, whichever occurs first.

- Follow-up period/final safety assessment:
  - 28 days after the last dose: scheduled visit
  - 60 days after the last dose: follow-up phone call, for a final safety assessment

**Figure 1: Study Schema**




The intent for this study is to evaluate the efficacy and safety profile of pamrevlumab as monotherapy in subjects with IPF who were previously treated with an approved therapy but

who discontinued that therapy (possible reasons for discontinuation of approved therapy could include, but are not limited to, intolerance or disease progression), unless neither treatment is available in the host country.

After randomization and during the treatment period, co-administration of an approved IPF therapy (i.e., pirfenidone or nintedanib) is permitted if clinically indicated in the Investigator's opinion, provided that the Investigator assesses the potential risks/benefits of combining approved IPF therapies with blinded study treatment.

Subjects who complete the Week 48 visit of the main study (regardless of the number of study drug infusions received and or treatment group assigned in the double-blind period) will be eligible to participate in the optional OLE period of the study that offers continuing access to pamrevlumab regardless of randomization assignment in the main study.

Subjects who discontinue study treatment for any reason are encouraged to remain in the study and be followed for all study visits and assessments.

The following assessments will be assessed centrally by independent external vendors or study committees:

- Pulmonary function tests (PFTs)
- High-resolution computed tomography (HRCT)
- Acute IPF exacerbations and respiratory hospitalizations

In addition, an independent Data Monitoring Committee (DMC) will review safety data and other clinical data (with the authority to unblind such data) on a periodic basis to monitor overall subject safety.

# 4. STUDY ENDPOINTS

## 4.1. Primary Endpoint

- Change in Forced Vital Capacity (FVC) from baseline at Week 48

## 4.2. Secondary Endpoints

- Time to disease progression, defined as absolute Forced Vital Capacity percentage predicted (FVCpp) decline of ≥10% or death, whichever occurs first
- Time to the first occurrence of any component of clinical composite endpoint: acute IPF exacerbation, respiratory hospitalization, or death, whichever occurs first
- Change in Quantitative Lung Fibrosis (QLF) volume from baseline at Week 48
- Time to first acute IPF exacerbation during study
- Time to all-cause mortality during study
- Time to first respiratory hospitalization during study

## 4.3. Exploratory Endpoints

- Time to composite of: respiratory hospitalization, absolute FVCpp decline ≥10%, or all-cause death, whichever occurs first
- Change in absolute FVCpp from baseline at Week 48
- Change in relative FVCpp from baseline at Week 48
- Change in St. George's Respiratory Questionnaire (SGRQ) score from baseline at Week 48
- Change in University of California San Diego – Shortness of Breath Questionnaire (UCSD-SOBQ) score from baseline at Week 48
- Change in Leicester Cough Questionnaire (LCQ) from baseline at Week 48

## 4.4. Safety Assessments

- Treatment-emergent adverse events (TEAEs) and treatment-emergent serious adverse events (TESAEs)
- Clinical laboratory parameters
- Vital signs
- Immunogenicity (Human Anti-Human Antibody [HAHA] formation, also named ADA: anti-drug antibody)
- Hypersensitivity/anaphylactic reactions

## 4.5. Exploratory Biomarker Endpoints

Biomarkers to be analyzed may include but are not limited to:

- Connective Tissue Growth Factor (CTGF)
- markers of inflammation such as C-reactive protein (CRP), tumor necrosis factor (TNF), interleukin-6 (IL-6), interleukin-8 (IL-8), and chemokine (C-C motif) ligand 18 (CCL18)
- markers of fibrosis and collagen synthesis such as transforming growth factor beta (TGF-beta), procollagen type I N-terminal propeptide (PINP), and type III collagen (PIIINP)
- other markers of tissue remodeling and angiogenesis such as angiopoietins, vascular endothelial growth factor (VEGF), matrix metalloproteinase-7 (MMP7), and Krebs von den Lungen 6 (KL-6).

# 5. GENERAL STATISTICAL CONSIDERATIONS

## 5.1. General Conventions

All data collected will be included in the data listings. All analyses will be performed using SAS® Version 9.3 or higher.

### 5.1.1. Analysis of Categorical Endpoints

Frequency distributions (number and percentage of subjects) will be presented for categorical variables.

### 5.1.2. Analysis of Continuous Endpoints

For continuous variables, descriptive statistics, including n, mean, standard deviation (SD), median, minimum, and maximum, will be presented.

Line graphs of group mean (and standard error: SE) and mean change from baseline (and SE) values will be plotted over visits as appropriate. Additional plots such as Least-square mean (SE) or boxplot will be provided as appropriate.

### 5.1.3. Analysis of Time-to-Event Endpoints

Kaplan-Meier curves for time-to-event endpoints will be plotted as appropriate. Comparisons of the respiratory hospitalization and acute IPF exacerbation endpoints will be based on the events identified by the Independent Adjudication Committee in accordance with the Adjudication Charter.

### 5.1.4. Analysis of Spirometry Data

All spirometry results are reviewed by the independent spirometry Over-Reader group for quality, during their Best Test Review (BTR). The Over-Reader does an overall quality rating of the spirometry data (“acceptable”, “borderline acceptable” or “unacceptable”) and determines which is the “Best” FVC (L) value and corresponding FVCpp for each test.

For each visit, only one FVC value and corresponding FVCpp that is identified by the Over-Reader as the “Best” FVC value, with an overall quality rating of “acceptable” or “borderline acceptable” will be used for analyses.

NOTE: A BTR is performed for each spirometry. In the situation where at a visit multiple spirometries were performed, the result considered for the analyses will be selected as follows:

Definition of “Best” post-BTR FVC value (and correlated FVCpp value): The “Best” value will be defined as that which meets the following criteria: (1) FVC value was selected after BTR by over-reader as the best test result; and (2) Result received an acceptable or borderline acceptable BTR rating.

- For the Day 1 assessment (Baseline), the “Best” post-BTR FVC value (and correlated FVCpp value) as defined above which also meets the following criteria will be used: (1) PFT was performed on or prior to Day 1; and (2) Result met eligibility criteria,

that is, correlated FVCpp value was within the range of >45% and <95%. The qualifying Day 1 value will be used whenever a qualifying Day 1 value is available; otherwise, the qualifying value on the last visit prior to Day 1 will be used. In cases where there is more than one set of PFTs performed on the same day, the highest of the multiple qualifying values will be used.

- For all subsequent (post-Day 1) visits, the "Best" post-BTR FVC value (and correlated FVCpp value) as defined above which correlates with the study visit will be used for FVC and FVCpp analyses.

#### 5.1.5. Analysis of HRCT Data

HRCT images that meet the quality control standards as assessed by the independent radiology imaging group will be considered for the analyses.

## 5.2. Analysis Sets

#### 5.2.1. Intent-to-Treat Population (ITT)

The intent-to-treat population includes all randomized subjects. Subjects will be analyzed according to their randomized treatment arm regardless of the actual study treatment received.

#### 5.2.2. Per-Protocol Set (PPS)

The per-protocol set is defined as all randomized subjects who have completed at least 36 weeks of treatment, with baseline and at least one post-baseline PFT assessment, and no major protocol deviation(s) that significantly impact efficacy analyses.

#### 5.2.3. Safety Analysis Set (SAF)

The safety analysis set includes all subjects who have received any study drug. Subjects will be analyzed according to the treatment actually received.

#### 5.2.4. Immunogenicity Analysis Set (IGS)

The IGS will include all subjects who are in the SAF and have a baseline evaluable immunogenicity assessment and at least 1 post-baseline evaluable (i.e., positive, negative) immunogenicity assessment.

## 5.3. General Data Handling Rules and Presentation Specifications

#### 5.3.1. Analysis Period

##### 5.3.1.1. On-Study Period

Unless otherwise specified, the efficacy analysis will be based on the on-study period, which includes all data collected during the double-blind period.

For the safety analysis, all safety assessments beyond 60 days of the last dose will be excluded from summary tables or figures.

#### 5.3.1.2. On-Treatment Period

The on-treatment analysis includes only the assessments that are observed during on-study period as defined above (Section 5.3.1.1), but within 3 weeks after the last study drug infusion in the double-blind period.

### 5.3.2. Baseline Definitions

Baseline for PFTs is defined as the last "Best" PFT (from the Over-Reader) before the first study drug infusion (or randomization date if no infusion is received) (Section 5.1.4).

Baseline for all other endpoints is defined as the last measurement prior to the first study drug infusion (or randomization date if no infusion is received). Unscheduled visits will be considered for baseline.

### 5.3.3. Formulas

- Study Day Calculation
  - The day when a subject receives the first dose of study drug after randomization is designated as Day 1. For subjects who have not received any study drug in an ITT analysis, their randomization date will be used as Day 1.
  - Study day of an assessment/procedure is calculated as follows.
    - For assessments or procedures on Day 1 or later:
      *Study day = assessment/procedure date – Day 1 date + 1*
    - For assessments or procedures earlier than Day 1:
      *Study day = assessment/procedure date – Day 1 date*
- Body Mass Index (BMI) $= \frac{Weight\ (kg)}{Height\ (cm)^2}$
- Body weight, height, and temperature will be converted using the following formula:
  - *kg = lb/2.2*
  - *cm = 2.54 × in*
  - *$^{o}C$ = (5/9) × ($^{o}F$ – 32)*
- Absolute Change from Baseline = *Value at Post-baseline visit – baseline value*
- Relative change from baseline (%) is derived with the formula below:

$$\frac{Value\ at\ Post\ baseline\ visit - Baseline\ value}{Baseline\ value} \times 100$$

### 5.3.4. General Instructions of TLF

- For continuous variables that are recorded as "<X" or ">X", the value of "X" will be used in the calculation of summary statistics.

- All reporting values (mean, median, SD, SE, LSMean, 95% CI, etc.) for the continuous variables will have 1 additional decimal place than raw data. Min and Max will have the same decimal place as raw data.
- All percentages will be rounded to one decimal place. The percentage will be suppressed when the count is zero. All durations of time will have 1 decimal place.
- All tables and listings will have a header showing "FibroGen, Inc.", the protocol number (ZEPHYRUS 2), date of data cutoff, and Page x of y. A footer will show the program file path/name, output file path/name, date of data extraction, run date, and run time.

#### 5.3.5. Handling of Dropouts or Missing Data

All assessments collected will be considered for analyses regardless of whether such data were collected during treatment or after a subject discontinued treatment. All analyses assume the missing data are missing at random (MAR), unless stated otherwise. Detailed missing data handling is described in the analysis of specific endpoints.

## 5.4. Interim Analyses and Data Monitoring Committee

In addition to routine safety monitoring, an independent DMC will be established to review safety data on an ongoing basis. A DMC charter will establish the procedures, meeting frequency, and scope of responsibilities of the committee.

This study has no planned or pre-specified interim analysis for either efficacy or futility.

## 5.5. Analysis Visit Window

Analysis visits, instead of the nominal visits from case report form (CRF), derived from visit dates and visit time windows will be used in the by-visit analyses. Unscheduled visits within a visit window (defined in Appendix 2) will be grouped into the closest scheduled visits based on the visit date. For subjects who have more than one measurement at a certain analysis visit, the last measurement will be used, with the following exceptions:

- Liver function tests, such as Alanine Aminotransferase (ALT), Aspartate Aminotransferase (AST), Gamma-Glutamyltransferase (GGT), Alkaline Phosphatase (ALP), and total bilirubin, in which the maximum value will be used.
- Spirometry data: the best value to be used for analysis as defined in Section 5.1.4.
- Vital signs: the values on the infusion days will be used if applicable.
- ADA assessment: select in the order of 1) last ADA+ assessment; 2) earlier ADA+ assessment; 3) last ADA- assessment.

# 6. STATISTICAL ANALYSIS

## 6.1. Subject Enrollment and Disposition

### 6.1.1. Eligibility Criteria

Eligibility criteria will be summarized for all screened subjects. The data will be summarized with respect to:

- number of subjects screened
- number (%) of subjects screen-failed
- number (%) of subjects for each failed inclusion/exclusion criterion

### 6.1.2. Subject Accountability and Disposition

Subject level inclusion criteria not met/exclusion criteria met listings will be provided.

The number (%) of subjects randomized (ITT), dosed (SAF), per-protocol set (PPS), Immunogenicity Analysis Set (IGS), completed the main study, entered the OLE period, and discontinued prematurely from treatment and study will be presented for each treatment group and all subjects pooled.

Reasons for premature discontinuation will be summarized by treatment group for the ITT population.

In addition, discrepancies between the randomization stratification factor derived from data in the Electronic Data Capture System (EDC) vs. data from the Interactive Response System (IRT) randomization system will be summarized. A listing will show any discrepancies between the two systems. If more than 5% discrepancies were identified, EDC randomization stratification factor will be used as fixed effects in the statistical analysis if applicable; otherwise, IRT randomization stratification factor will be used, unless otherwise specified.

## 6.2. Important Protocol Deviations

The number and percentage of subjects with important protocol deviations will be categorized and tabulated as appropriate for the ITT population. COVID-19-related important protocol deviations will be summarized in a separate table. All protocol deviations will be finalized prior to unblinding.

A subset of pre-specified major protocol deviations will exclude some patients in the PPS analyses if these deviations significantly impact efficacy analyses. These major PDs will be reviewed and finalized by following FibroGen medical review process prior to database lock and unblinding. Considerations will be given according to the following table.

**Table 1: Criteria for Major Protocol Deviations**

| Number | Major Protocol Deviation |
|---|---|
| 1 | Violation of any inclusion or exclusion criteria that may affect the assessment of efficacy or safety of the study drug* |
| 2 | Withdrawal Deviation: Subject met withdrawal criteria during the study but was not withdrawn* |
| 3 | Dosing Deviation: Subject received the wrong treatment or incorrect dose, including incorrect timing of a dose, which may affect the assessment of efficacy or safety of the study drug* |
| 4 | Administration of prohibited concomitant medication that may impact evaluation of efficacy of the study drug* |
| 5 | Significant noncompliance with study procedures that may impact evaluation of efficacy of the study drug will be evaluated case by case* |

* Major PDs will be reviewed and finalized by following FibroGen medical review process; these five categories may not cover all IPDs.

## 6.3. Demographics and Baseline Characteristics

### 6.3.1. Demographics and Baseline Characteristics

Demographics and important baseline characteristics will be summarized for the ITT by treatment arm and all pooled. These may include but are not limited to age, age group (≤64, 65-74, ≥75), sex, ethnicity, race group, weight, body mass index (BMI), smoking history, prior treatment with an approved IPF therapy (Yes/No), GAP stage (I, II, III), and years since the first diagnosis of IPF. Age is defined as the age on the day of signing the informed consent form (ICF). Years since the first diagnosis will be derived using the formula: date of ICF- date of IPF diagnosis+1. Partial missing date of IPF diagnosis will be imputed as Jan. 01 (if month is missing), or 01 (if day is missing).

Baseline values for efficacy assessments will be presented in baseline tables as appropriate.

Comparability of baseline characteristics between treatment groups will be assessed using analysis of variance (ANOVA) model for continuous variables and Chi-squared test for categorical variables. The nominal p-values will be presented for reference only.

### 6.3.2. Medical History

Medical History coded by Medical Dictionary for Regulatory Activities (MedDRA, Version 25.0 or higher) will be summarized for the ITT Population.

## 6.4. Prior and Concomitant Medications

The latest World Health Organization Drug Dictionary (WHODD Global B3 Mar2022, or higher version) will be used to classify prior and concomitant medications by therapeutic class and preferred term.

Prior medication is defined as any medication taken and stopped prior to the first dose of the study drug. Concomitant medication is defined as any medication taken after the first dose of the

study drug and before the last dose days +60 days. For subjects who are treated in the OLE, medications that started after 1st dose in the OLE period may be considered as concomitant medication for the OLE period, but not for the double-blind period. Partially or incompletely missing prior/concomitant medication start or stop dates will be imputed (Appendix 1). Table 2 provides the classification guideline when medication starting or ending dates are missing.

**Table 2: Classification of Prior and Concomitant Medications**

| End Date / Start Date | Before start of study drug administration | On or after start of study drug administration | Missing |
|---|---|---|---|
| **Before start of study drug administration** | Prior | Concomitant | Concomitant |
| **On or after start of study drug administration and before the last dose date +60 days** | - | Concomitant | Concomitant |
| **Missing** | Prior | Concomitant | Concomitant |

Both prior and concomitant medication usage will be summarized by the number (%) of subjects receiving the drug within each therapeutic class and ATC code level 3 and preferred term for the SAF. Multiple usages of the same drug by a patient will be counted only once.

Separate summaries may be provided for prior and concomitant medications of special interest such as the approved IPF medications (pirfenidone and/or nintedanib).

To evaluate the use of co-administered medications that are narrow therapeutic CYP substrates, the concomitant medication for narrow therapeutic index cytochrome P450 substrates will be listed and summarized. The list of related medications will be finalized prior to database lock.

## 6.5. Prior and Concomitant Procedures and Non-drug Therapies

Procedures and non-drug therapies were coded with MedDRA Version 25.0 or higher version. Prior/Concomitant procedures and non-drug therapies will be defined and summarized similarly to prior and concomitant medications (Details in Section 6.4). Partially or incompletely missing procedures and non-drug therapies start or stop dates will be imputed similarly to prior/concomitant medications (Appendix 1).

## 6.6. Study Drug Exposure and Treatment Compliance

### 6.6.1. Study Drug Exposure

Exposure to the study drug will be summarized in terms of treatment duration for the SAF.

Duration of study treatment in weeks is calculated as: *(last dose date – first dose date + 1)/7.*

Duration of study treatment will be summarized as a continuous variable and be tabulated by the categories as follows:

- <6 weeks
- 6 to <12 weeks
- 12 to <18 weeks

- 18 to <24 weeks
- 24 to <36 weeks
- ≥36 weeks

Number (%) of subjects by the total number of infusions (1-3, 4-6, 7-9, 10-12, ≥13) received, overall average infusion doses (in mg/kg), whether infusion was missed/interrupted and reason for missed/interruption, and patient-exposure years (PEY) will be summarized for the SAF.

### 6.6.2. Treatment Compliance

The compliance will be presented as % of the actual dose administered out of the total planned dose of infusions by end of study (EOS) or by discontinuation in subjects who discontinued.

$$Compliance\ (\%) = \frac{actual\ total\ dose\ received}{total\ planned\ dose\ while\ actively\ in\ treatment} \times 100$$

Descriptive statistics for study drug compliance will be presented by treatment group for the SAF. Treatment compliance will be summarized as a continuous variable and as a categorical variable (<70%, 70% –<80%, 80% –<90%, 90%-110%, and >110%).

## 6.7. Efficacy Analysis

The primary and secondary endpoints will be tested using a fixed sequence analysis approach to preserve the study-wide error rate of 5%. Under the sequential analysis, the primary and secondary efficacy endpoints will be tested in a defined sequence according to the order listed in Table 3 each at the usual alpha= 0.05 level of statistical significance. The testing will cease when a failure occurs in the pre-determined sequential hypothesis testing and all p-values for the subsequent testing will be considered nominal. All p-values for exploratory endpoints will be considered nominal.

All analyses for efficacy endpoints will be performed for the ITT population during the On-Study period (Section 5.3.1.1), unless noted otherwise. The placebo group will be used as the reference group for all treatment comparisons.

**Table 3: Fixed Sequence Testing Order of Primary and Secondary Endpoints**

| Testing order | Endpoints |
|---|---|
| 1 | Change in FVC (L) from baseline at Week 48 (**Primary Endpoint**) |
| 2 | Time to disease progression, defined as absolute FVCpp decline of ≥10% or death, whichever occurs first |
| 3 | Time to the occurrence of any component of the clinical composite endpoint: acute IPF exacerbation, respiratory hospitalization, or death (whichever occurs first) |
| 4 | Change in QLF volume from baseline at Week 48 |
| 5 | Time to first acute IPF exacerbation during study |
| 6 | Time to all-cause mortality during study |
| 7 | Time to first respiratory hospitalizations during study |

### 6.7.1. Primary Endpoint Analysis and Estimand

The primary endpoint is change from baseline in FVC (L) at Week 48.

The hypotheses to be tested for the primary efficacy analysis is:

$H_0$: Change from baseline in FVC (L) at week 48 for the pamrevlumab arm = Change from baseline in FVC (L) at week 48 for the placebo arm

Versus:

$H_1$: Change from baseline in FVC (L) at week 48 for the pamrevlumab arm ≠ Change from baseline in FVC (L) at week 48 for the placebo arm

#### 6.7.1.1. Primary Analysis with MMRM

##### 6.7.1.1.1. Estimand Strategy

The primary estimand is intended to provide a population level estimate of the treatment effect of the pamrevlumab on a continuous endpoint, regardless of participant compliance with the IP (investigational product) dosing.

##### 6.7.1.1.2. Population of Interest

The ITT population includes all randomized subjects during the on-study period as defined in Section 5.3.1.1. Male and female patients aged 40 or above with IPF diagnosis under American Thoracic Society/European Respiratory Society (ATS/ERS) criteria (Raghu 2018), and not currently receiving treatment with the 2 approved antifibrotic standard of care medications.

##### 6.7.1.1.3. Intercurrent Event Handling Strategy

Treatment discontinuation due to non-death reasons: Treatment policy strategy will be used. Missing data are assumed to be missing at random.

Death: Post-death FVC values will be set as the worst (lowest) post-baseline FVC values across all randomized subjects in the double-blind period. Missing study day will be imputed as the target day of each analysis visit.

#### 6.7.1.1.4. Analysis Variable

Change from baseline in all FVC assessments in the double-blind period, including scheduled, unscheduled, and available assessments after treatment discontinuation, up to week 48 during the On-Study period (Section 5.3.1.1), will be included in the analysis.

In the situation where a subject has more than one record of FVC within an analysis visit, the best value will be used for MMRM.

#### 6.7.1.1.5. Population Summary for Treatment Comparison

Treatment difference of Least-square mean (LSMean) and SE at week 48 and the corresponding 95% CI will be presented.

All post-baseline visits for the specified parameters below during the double-blind period will be included in the analysis. Change from baseline for the specified endpoints (excluding baseline visits) will be analyzed using a Mixed Model for Repeated Measures (MMRM) approach with fixed effects for treatment, visit (as a factor), treatment-by-visit interaction, randomization stratification factor, and covariates (baseline values, sex, age, race group, and height). Example SAS code in Appendix 7.1.

The following graphical displays will be provided:

- The mean (± SE) FVC over time by treatment group
- The mean (± SE) FVC change from baseline by treatment group
- The LSmean (± SE) estimated FVC change from baseline over time by treatment group (as estimated in the primary analysis)

#### 6.7.1.1.6. Covariance Structure Strategy

The unstructured covariance pattern model will be used first. The by-treatment-group option will be added to the covariance pattern to improve the model fit as appropriate. If the model using the unstructured covariance pattern does not converge, the following covariance structures will be tested in sequence: heterogeneous Toeplitz, homogeneous Toeplitz, first-order autoregressive, compound symmetry, and variance component, until the model converges.

If the model doesn't converge for all six covariance structures listed above, then the following attempts will be made: 1) sandwich covariance estimator will be used; 2) some least significant factors or interaction terms ($p>0.05$) can be dropped from the model to achieve convergence. The revised model with fewer factors or interaction terms will be tested using the same sequence as specified above.

This covariance structure strategy will apply to all analyses on continuous primary, secondary, and exploratory efficacy endpoints if appropriate.

#### 6.7.1.2. Sensitivity Analysis

The following sensitivity analysis under the same estimand framework as the primary analysis defined in Section 6.7.1 will be performed on FVC:

##### 6.7.1.2.1. Linear Slope Model (RCM)

The random coefficient linear regression model (RCM) with a random effect for the slope of time will be performed. The time in the RCM model will be calculated as the elapsed weeks (continuous) from the first infusion (or randomization if not dosed) to the assessment date. The RCM model will also include fixed effects of treatment, time, treatment-by-time interaction, randomization stratification factor, baseline FVC volume, and covariates (age, sex, race group, and height). Baseline visits will be set as 0 and included in the RCM model. In the situation where a subject has more than one record of FVC on different study days within an analysis visit, the highest qualifying FVC on each study day (Section 5.1.4) will be used for RCM.

##### 6.7.1.2.2. Jump-to-Control Analysis

The goal of the jump-to-control analysis is to address the possibility of data being missing not at random (MNAR). The missing data pattern for the pamrevlumab subjects after withdrawal from the study can be assumed to switch to the same data pattern as subjects on the placebo treatment. Subjects that discontinued from the placebo arm are assumed to have the same data pattern as placebo subjects that remain in the study (excluding post-death missing values defined in Section 6.7.1.1.3). This is often called the jump-to-control approach.

The analysis of covariance (ANCOVA) model will contain terms for treatment, baseline FVC measurements, covariates (age, sex, race group, and height), and the randomization stratification factor (Example SAS code in Appendix 7.2).

##### 6.7.1.2.3. Delta-Adjusting (Tipping Point) Analysis

An alternative assumption is that the missing data for the pamrevlumab treated subjects who discontinue early have a lower expected value than the pamrevlumab subjects remaining in the study, while subjects who discontinue from the placebo arm are assumed to have the same data pattern as placebo subjects remaining in the study (excluding post-death missing values defined in Section 6.7.1.1.3). This is often called the delta-adjusting (or tipping point) approach (Example SAS code in Appendix 7.3).

ANCOVA will be performed as Section 6.7.1.2.2.

##### 6.7.1.2.4. Ranked ANCOVA Analysis

A ranked ANCOVA analysis (a non-parametric analysis) will be performed on the ranked change from baseline values in FVC at Week 48. The ranked ANCOVA was the primary efficacy endpoint analysis in the pirfenidone phase 3 trials.

Missing week 48 values will be imputed as the average value from three subjects with the smallest sum of squared differences over the visits with non-missing data between each of the three subjects and the early terminated subject.

Death: Subjects with missing data due to death were ranked worse than those who remained alive. Subjects who died were ranked according to the number of days from randomization until death, with the shortest time to death as the worst rank.

The ranked ANCOVA with terms for sex, age, race group (white and non-white), height, randomization stratification factor, and standardized rank baseline FVC value as a covariate will be performed on the ranked change from baseline in FVC at week 48 (Example SAS code in Appendix 7.4). The p-value for treatment comparison will be presented.

#### 6.7.1.3. Supplemental Analysis

The following supplemental analysis under the same estimand framework as the primary analysis defined in Section 6.7.1 will be performed on FVC:

##### 6.7.1.3.1. On-Treatment Period Analysis

An additional analysis will be performed for assessments performed during the On-treatment period as defined in Section 5.3.1.2. FVC assessments beyond 3 weeks after the last study drug infusion will be excluded from the analysis.

##### 6.7.1.3.2. Per-Protocol Set Analysis

An additional analysis will be performed for subjects in PPS as defined in Section 5.2.2. Same as Section 6.7.1.1.4. EDC randomization stratification factor will be used for analysis in PPS.

#### 6.7.1.4. Subgroup Analysis

The primary analysis will be repeated for relevant and appropriate subgroups. The LSMean of treatment difference and corresponding 95% CI will be presented in a forest plot as appropriate.

The subgroups may include but are not limited to sex, age group, race group (white vs. others), prior use of approved IPF treatments (yes/no), any use of approved IPF treatments during the study (yes/no), country/geographic region (Europe/Latin America/US/Asia Pacific), country (China/Non-China), GAP stage (I, II, III), and ADA status (Positive, Negative, or unknown; details in Section 6.10).

### 6.7.2. Secondary and Exploratory Endpoints Analysis and Estimands

The secondary endpoints during the on-study period (as defined in Section 5.3.1.1) will be analyzed in the order specified in Table 3. In addition, on-treatment analysis and subgroup analysis (both considered nominal) will also be performed for secondary efficacy endpoints. All Exploratory endpoint analyses will be considered nominal.

#### 6.7.2.1. Time-to-Event Endpoints

##### 6.7.2.1.1. Estimand Strategy

The secondary estimands are intended to provide a population level estimate of the treatment effect of the pamrevlumab on time-to-event endpoints; regardless of participant compliance with the IP dosing.

#### 6.7.2.1.2. Population of Interest

Same as Section 6.7.1.1.2

An additional analysis will be performed for all secondary endpoints during the on-treatment period as defined in Section 5.3.1.2.

#### 6.7.2.1.3. Intercurrent Event Handling Strategy

Treatment discontinuation (such as early termination due to AEs, Lost to Follow-Up, Withdrawal by Subject, Physician Decision, Protocol Deviations, etc.): Treatment policy strategy will be used.

Death: Composite strategy will be used if death is a component of the endpoint.

#### 6.7.2.1.4. Analysis Variables

##### 6.7.2.1.4.1. Time (Days) to disease progression

Time (Days) to disease progression is defined as the number of days from randomization to either the first occurrence of an absolute ≥10% decline from baseline in FVCpp or all-cause death based on observed data, whichever occurs earlier during the on-study period as described in Section 5.3.1.1. Subjects who die without any post-baseline PFT assessments will be considered as an event on the date of death.

Subjects without an event will be censored on the date of the last PFT assessment after randomization in the double-blind period. Alive/Censored subjects without any post-baseline PFT assessments will be censored on Day 1.

##### 6.7.2.1.4.2. Time (Days) to the first occurrence of any component of the clinical composite endpoint: acute IPF exacerbation, respiratory hospitalization, or death

Time (Days) to the occurrence of any component of the clinical composite endpoint is defined as the number of days from randomization to either the first occurrence of acute IPF exacerbation (including both confirmed and suspected cases), respiratory hospitalization, or death during the on-study period as described in Section 5.3.1.1.

Subjects without an event will be censored on the date of last known alive during the double-blind period.

##### 6.7.2.1.4.3. Time (days) from randomization to first adjudicated acute IPF exacerbation

Time (Days) to first acute IPF exacerbation is defined as the number of days from randomization to either the first occurrence of adjudicated acute IPF exacerbation (including both confirmed and suspected cases) as described in Section 5.3.1.1.

Subjects without an event will be censored on the date of last known alive during the double-blind period.

#### 6.7.2.1.4.4. Time (days) from randomization to all-cause mortality

Subjects without an event will be censored on the last date known alive in the double-blind period. Any survival vital status will be included for early discontinued subjects.

#### 6.7.2.1.4.5. Time (days) from randomization to first adjudicated respiratory hospitalization

Subjects without an event will be censored on the date of last known alive during the double-blind period.

#### 6.7.2.1.4.6. Time (Days) to the first occurrence of composite of adjudicated respiratory hospitalization, absolute FVCpp decline ≥10%, or all-cause death

Time (days) to the first occurrence of composite is defined as from randomization to the first occurrence, which includes adjudicated respiratory hospitalization, all-cause death, or absolute FVCpp decline ≥10%. Subjects who died without any post-baseline PFT assessments or respiratory hospitalization events will be considered as an event on the date of death.

Subjects without an event will be censored with the same rule defined in Section 6.7.2.1.4.1.

### 6.7.2.1.5. Population summary for treatment comparison

The hazard ratio and its corresponding 95% CI will be presented. P-value from stratified Log-rank test will also be presented.

All time-to-event analyses will include data collected in the double-blind period. The stratified Cox proportional hazard model including baseline FVCpp stratified by randomization stratification factor will be used to estimate the hazard ratio and its corresponding 95% CI. A stratified log-rank test stratified by randomization stratification factor will be used for treatment comparison.

For time from randomization to the first acute IPF exacerbation, all events pre-defined in the adjudication charter will be adjudicated by an independent Adjudication Committee. The adjudicated results will be summarized by categories of confirmed, suspected, or not acute respiratory exacerbation.

In addition, all acute respiratory exacerbations will be further summarized as:

- a triggered acute respiratory exacerbation
- an idiopathic acute exacerbation (i.e., no trigger identified).

## 6.7.2.2. Continuous Endpoints with MMRM Model

### 6.7.2.2.1. Estimand Strategy

Same as Section 6.7.1.1.1

### 6.7.2.2.2. Population of Interest

Same as Section 6.7.1.1.2

#### 6.7.2.2.3. Intercurrent Event Handling Strategy

Same as Section 6.7.1.1.3

For all post-death analysis visits, the missing values will be set as the worst observed post-baseline values (lowest FVCpp and LCQ score, highest QLF, UCSD-SOBQ, and SGRQ score) across all subjects in the double-blind period. Study day will be imputed as the target day of the corresponding analysis visit.

#### 6.7.2.2.4. Analysis Variables

##### 6.7.2.2.4.1. Change from baseline in QLF volume

The Quantitative Lung Fibrosis (QLF) volume is calculated as *QLF=total lung capacity volume (TLC) * % of quantitative lung fibrosis for fibrosis of the whole lung*.

##### 6.7.2.2.4.2. Absolute change from baseline in FVCpp

The absolute change from baseline in FVCpp is derived with the formula below:

$$FVCpp\ value\ at\ Post\ baseline\ visit - FVCpp\ value\ at\ Baseline$$

##### 6.7.2.2.4.3. Relative change from baseline in FVCpp

The relative change (%) from baseline in FVCpp is the change from baseline FVCpp value as a percentage of the baseline FVCpp value and is derived with the formula below:

$$\frac{FVCpp\ value\ at\ Post\ baseline\ visit - FVCpp\ value\ at\ Baseline}{Baseline\ FVCpp\ value} \times 100$$

##### 6.7.2.2.4.4. Change from baseline in SGRQ

Refer to Appendix 3 for details of scores for St. George's Respiratory Questionnaire (SGRQ). The total score of SGRQ together with the symptom, activity, and impact domain scores, will be included in the analysis.

##### 6.7.2.2.4.5. Change from baseline in the total score of UCSD-SOBQ

Refer to Appendix 4 for details of scores for University of California, San Diego – Shortness of Breath Questionnaire (UCSD-SOBQ).

##### 6.7.2.2.4.6. Change from baseline in LCQ

Refer to Appendix 5 for details of scores for Leicester Cough Questionnaire (LCQ). The total score of LCQ together with the physical, psychological, and social domains cores, will be included in the analysis.

#### 6.7.2.2.5. Population Summary for Treatment Comparison

Treatment difference of LSMeans (and SE) at each scheduled visit and the corresponding 95% CI will be presented.

All post-baseline visits for the specified parameters below during the double-blind period will be included in the analysis. Change from baseline for the specified endpoints (excluding baseline visits) will be analyzed using a Mixed Model for Repeated Measures (MMRM) approach with fixed effects for treatment, visit (as a factor), treatment-by-visit interaction, randomization stratification factor, and covariates (baseline values, sex, age, race group, and height). Covariates of sex, age, race group, and height will be excluded from the model for FVCpp since these variables have already been adjusted in the definition of FVCpp.

#### 6.7.2.3. Subgroup Analysis for Secondary Efficacy Endpoints

The same subgroup analysis defined in Section 6.7.1.4 will be performed for all secondary efficacy endpoints.

## 6.8. Safety Analysis

The safety analyses will be performed for the SAF.

### 6.8.1. Adverse Events (AEs)

Adverse events will be coded using MedDRA.

A new or worsening AE occurring on or after the first dose of study drug and within 60 days after the last dose of study drug is defined as a treatment-emergent adverse event (TEAE). AEs that started after 1st dose in the OLE period will be excluded from the main study. Partially or incompletely missing AE start/stop date/time will be imputed (Appendix 1).

If more than one event occurs with the same system organ class (SOC) and preferred term (PT) for the same patient, the patient will be counted only once for that preferred term using the most severe and most related occurrence for the summarization by common terminology criteria for adverse events (CTCAE) severity grade and relationship to the study treatment. Relationship to study treatment will be imputed as "Related" for any TEAE with a missing value for relationship.

The cumulative incidence of the following AE categories including the number (%) of subjects will be produced:

- Summary of all AEs
- TEAEs by PT
- TEAEs by SOC and PT
- TEAEs by SOC, PT, and maximum severity
- TEAEs by SOC, PT, and investigator-determined relationship
- TEAEs with severity grade ≥3 by SOC and PT
- TEAEs related to study treatment determined by the investigator by SOC and PT
- TEAEs with frequency ≥ 5% of subjects in either treatment arm by SOC and PT

- Non-serious TEAEs with frequency ≥ 5% of subjects in either treatment arm by SOC and PT (Note: this is required for ClinicalTrials.gov)
- TEAEs leading to discontinuation of study treatment by SOC and PT
- TEAEs leading to interruption of study treatment by SOC and PT
- Treatment-emergent serious AEs (TESAEs) by SOC and PT
- TESAEs by PT
- TESAEs by SOC, PT, and maximum severity
- TESAEs related to study treatment determined by the investigator by SOC and PT
- Fatal TESAEs (i.e., adverse events that have an outcome of death) by SOC and PT
- All-cause deaths

Listings of all adverse events, TESAEs, TEAEs leading to study treatment discontinuation, TEAEs leading to death, and all-cause deaths will be provided.

In addition, TEAEs and TESAEs will be presented by subgroups as described in Section 6.7.1.4 along with an additional subgroup assessment by concomitant steroid medication use.

### 6.8.2. Special Safety Events

Treatment-emergent special safety events including:

1. Hypersensitivity (any time)
2. Infusion reactions (occurred on the day or the day after any study drug infusion)
3. Anaphylactic reactions (occurred on the day or the day after any study drug infusion)

Items 1 and 2 include both hypersensitivity and angioedema events. Both items 1 and 2 will be listed and summarized similarly to TEAEs:

- Events by event type, SOC, PT, and maximum severity
- Events by event type, SOC, PT, and subgroups (same as TEAEs)

Item 3 will be listed and summarized by PT only.

The preferred term list for these special safety events will be finalized prior to database lock.

### 6.8.3. Clinical Laboratory Parameters

Descriptive statistics for laboratory values (in SI units) and changes from baseline at selected visits will be presented for the following laboratory parameters. The number (%) of subjects by toxicity grades at selected visits will also be provided for selected lab tests. Shift tables for selected parameters from baseline to worst post-baseline will be provided.

**Table 4: Laboratory Tests**

| | |
|---|---|
| **Hematology Panel:**<br>Absolute neutrophil count (ANC), Basophils, Eosinophils, Erythrocyte count (RBC), Hematocrit %, Hemoglobin, WBCs (Leukocyte count), Lymphocytes, Mean corpuscular volume, Monocytes, Neutrophils, Platelets | **Chemistry Panel:**<br>Bicarbonate, blood urea nitrogen (BUN), Calcium, Creatinine, Chloride, Magnesium, Glucose, ALP, ALT, AST, total Bilirubin, Albumin, Phosphorous, Potassium, Sodium, Cholesterol, Total Protein, GGT, Triglycerides |
| **Coagulation Panel:**<br>International Normalised Ratio (INR), partial thromboplastin time (aPTT/PTT) | |

### 6.8.4. Liver Function Tests

Liver function test results and the number (%) of subjects above 2×ULN or 3×ULN will be summarized over selected visits.

A matrix scatter plot of liver enzymes and bilirubin showing the maximum ALT or AST vs. total bilirubin during treatment on natural-log scales with a dotted line drawn at 3×ULN for ALT or AST vs. at 2×ULN for total bilirubin will be provided.

### 6.8.5. Vital Signs

Heart rate (beat/min), diastolic and systolic blood pressure (mmHg), temperature (°C), and body weight (kg) will be descriptively summarized by treatment at selected visits.

Vital sign values are potentially clinically significant (PCS) if they meet both the observed value criteria and the change from baseline criteria listed in Table 5 below. The number and percentage of subjects with post-baseline PCS values will be tabulated by treatment group. The percentages are to be calculated relative to the number of subjects with baseline and at least one post-baseline assessment. The numerator is the total number of subjects with at least one post-baseline PCS vital sign value. A supportive listing of subjects with post-baseline PCS values will be provided including the patient ID, study center, baseline, and post-baseline values. A listing of all AEs for subjects with PCS vital signs will also be provided.

**Table 5: Criteria for Potentially Clinically Significant Vital Signs**

| Vital Sign Parameter | Flag | Criteria* | |
|---|---|---|---|
| | | Observed Value | Change from Baseline |
| Systolic Blood Pressure (mmHg) | High | ≥ 170 | Increase of ≥ 20 |
| | Low | ≤ 90 | Decrease of ≥ 20 |
| Diastolic Blood Pressure (mmHg) | High | ≥ 110 | Increase of ≥ 15 |
| | Low | ≤ 45 | Decrease of ≥ 15 |
| Heart Rate (bpm) | High | ≥ 120 | Increase of ≥ 20 |
| | Low | ≤ 50 | Decrease of ≥ 20 |
| Weight (kg) | High | - | Increase of ≥ 10% |
| | Low | - | Decrease of ≥ 10% |

*Except for body weight, a post-baseline value is considered as a PCS value if it meets both criteria for observed value and change from baseline.

### 6.8.6. Electrocardiogram (ECG)

ECG results (normal, abnormal-clinically significant, and abnormal-not clinically significant) will be summarized by treatment over visits. A shift table for ECG from baseline to worst post-baseline will be provided.

### 6.8.7. Physical Examination (PE)

Abnormal PE results will be summarized by treatment over visits. A shift table will be provided if data is appropriate. A listing of abnormal PE results will be provided.

### 6.8.8. Pregnancy Test

Pregnancy tests for women of childbearing potential (WOCBP) only: serum pregnancy test at Screening; urine pregnancy tests (pre-dose) during study visits.

A listing of positive pregnancy test results during study visits will be provided.

## 6.9. Biomarker Endpoint Analysis

All analyses for the exploratory biomarkers listed in Section 4.5 will be performed for the SAF based on observed data as appropriate. Descriptive summaries will be performed for the other exploratory biomarkers as appropriate.

A listing will be provided for Tryptase at the time of any suspected hypersensitivity/ anaphylactic reactions, as defined in the protocol.

## 6.10. Immunogenicity Analysis and PK

Summary of immunogenicity data will be based on IGS and listings will be on SAF. The analysis dataset and data listing will include all available anti-drug antibody (ADA, also named HAHA) samples. The following terms and definitions are implemented.

### 6.10.1. Terms and Definitions

#### 6.10.1.1. Sample ADA Status

- Baseline ADA-positive sample: ADA is detected in the last sample before initiation of treatment.
- Baseline ADA-negative sample: ADA is not detected in the last sample before initiation of treatment.
- **Treatment-emergent ADA-Positive**: Meets definition of treatment-induced or treatment-boosted ADA:
  - **Treatment-induced ADA-Positive**: a post-treatment positive ADA is detected in a subject for whom pre-treatment ADA assessment is either negative or not assessable
  - **Treatment-boosted ADA-Positive**: pre-existing ADA was boosted to a higher level following study treatment, i.e. pre-treatment positive ADA titer was boosted by at least 2 dilution steps (4-fold) following study treatment.
- **ADA-negative sample**: After initiation of treatment, ADA is not treatment-emergent ADA-positive.

Next, using the sample ADA status, the subject's ADA status is defined.

#### 6.10.1.2. Subject ADA Status

- Baseline ADA-positive subject: A subject with a baseline ADA-positive sample.
- Baseline ADA-negative subject: A subject with a baseline ADA-negative sample.
- ADA-positive subject: An evaluable subject with at least one treatment-emergent ADA-positive sample at any time during the study.
- Neutralizing-positive: At least one treatment-emergent ADA-positive sample with neutralizing antibodies detected (if available).
- ADA-negative subject: An evaluable patient without a treatment-emergent ADA sample during the study.
- **ADA**-unknown: Patients without evaluable baseline and/or post-baseline ADA samples will be categorized as "ADA-unknown".

### 6.10.2. Statistical Analysis for Characterization of ADA Immune Response

#### 6.10.2.1. Incidence of ADA

- Percentage of treatment-emergent ADA patients for the defined study period, where the denominator is the number of ADA evaluable patients in the respective treatment arm and/or subgroup.

- Number (%) of subjects will be reported for the following parameters based on evaluable subjects:
  - Baseline ADA-positive
  - Treatment-emergent ADA-positive (Treatment-induced, Treatment-boosted)
  - Neutralizing Positive (if available)
  - ADA-negative
- **ADA prevalence**: Percentage of treatment-emergent ADA patients at any given timepoint, where the denominator is the number of ADA evaluable patients in the respective treatment arm and/or subgroup at that timepoint.
- A listing of all ADA assessments will be provided.
- Additionally, a separate listing of ADA assessments for all neutralizing antibody (NAb)-positive subjects will be provided (if available).

#### 6.10.2.2. ADA Titer Kinetics

All ADA-positive subjects will be included in the analysis.

- Summary statistics of subject-level ADA titers using the maximum titer value within an ADA-positive subject will be presented for baseline ADA-negative subjects and baseline ADA-positive subjects. The median, interquartile range, and range of the maximum titers will be reported. For ADA-positive subjects with a baseline ADA-positive sample, the median and interquartile range of the fold increase from baseline in titer (ratio of maximum post-baseline titer to baseline titer) will also be reported. Graphical presentation of the summary data may be provided using boxplots, as appropriate.
- For sample-level ADA titers, boxplots of ADA titers at each assessment time point will be provided, as appropriate, to demonstrate whether the ADA levels tend to change over time during the treatment, along with ADA incidence at each assessment time point.
- Spider plots may be considered to show the trend of ADA titer over time.

### 6.10.3. Clinical Implication of ADA Immune Response

#### 6.10.3.1. PK

- Effect of ADA response on drug exposure will be explored by examining the drug exposures using graphical plots or simple summary statistics of observed drug concentration levels by ADA status and sampling time. Corresponding numerical values of geometric mean, arithmetic mean, and standard deviation will be displayed under the figures.

- A listing of drug concentrations will be provided. Time course of observed concentrations by study visit with identifiers for antibody response will be plotted for each subject separately.

#### 6.10.3.2. Safety and Efficacy

- AEs and the primary efficacy endpoint will be summarized by ADA status (details in Section 6.7.1.4 and Section 6.8.1).

# 7. CHANGES FROM PROTOCOL

Based on the feedback from the Food and Drug Administration on a similar Phase 3 IPF study (protocol #: FGCL-3019-091), MMRM will be used as the primary analysis for FVC and RCM will be used as a sensitivity analysis. Supplemental and subgroup analyses for FVC will also use MMRM. The exploratory efficacy endpoint FVCpp will also be analyzed using MMRM.

In addition, with consideration of FDA recommendations for the FGCL-3019-091 study, the testing order was switched for the following secondary efficacy endpoints: "Time to the first occurrence of any component of clinical composite endpoint: acute IPF exacerbation, respiratory hospitalization, or death, whichever occurs first" and "Change in QLF volume from baseline at Week 48" (Table 3).

# 8. REFERENCE


- E9(R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials
- Eakin EG, Resnikoff PM, Prewitt LM, Ries AL, Kaplan RM., Validation of a new dyspnea measure: the UCSD Shortness of Breath Questionnaire. Chest. 1998; 113:619-624
- Jones, P.W., Quirk, F.H., Baveystock, C.M., and Littlejohns, P. A, Self-complete measure of health status for chronic airflow limitation. The St. George's Respiratory Questionnaire. Am Rev Respir Dis, 1992; 145: 1321–1327
- Jones PW, Forde, Y., St George's Respiratory Questionnaire Manual. http://www.healthstatus.sgul.ac.uk/SGRQ_download/SGRQ%20Manual%20June%202009.pdf. Accessed June 2, 2020
- Richeldi L, du Bois RM, Raghu G, et al. Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. N Engl J Med 2014; 370:2071-82.
- Noble PW, Albera C, Bradford WZ, Costabel U, etc.; Pirfenidone in patients with idiopathic pulmonary fibrosis (CAPACITY): two randomised trials; Lancet 2011; 377: 1760–69

- Statistical Review and Evaluation for Ofev® (nintedanib) 150 mg capsules; Indication: Treatment of Idiopathic Pulmonary Fibrosis (IPF); PDUFA: January 2, 2015
- King TE Jr., Bradford WZ, Castro-Bernardini S, etc.; A Phase 3 Trial of Pirfenidone in Patients with Idiopathic Pulmonary Fibrosis; N Engl J Med 2014;370:2083-92
- Picozzi VJ, etc.; FG-3019, a human monoclonal antibody to connective tissue growth factor (CTGF), with gemcitabine/erlotinib (G/E) in patients with locally advanced or metastatic pancreatic ductal adenocarcinoma (PDAC). J. Cancer Clin Trials 2017, 2:1
- Raghu G, etc.; Diagnosis of Idiopathic Pulmonary Fibrosis. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. Am J Respir Crit Care Med. 2018 Sep 1;198(5):e44-e68.
- The Acceptance of Overseas Drug Clinical Trial Data. NMPA 2018 https://www.nmpa.gov.cn/zhuanti/ypqxgg/ggzhcfg/20180710151401465.html

# APPENDIX 1. HANDLING MISSING/INCOMPLETE DATES

## Appendix 1.1 Missing/Incomplete AE Onset Date

The following imputation rules apply to the case where the start date is incomplete (i.e., partially missing) for adverse events. When the start date and the stop date are both incomplete for a patient, impute the start date first.

If the stop date is complete and the imputed start date is after the stop date, then the start date will be imputed using the stop date.

- **Missing start time only**

AEs with missing start times and which occur on a study-drug-dosing day will be considered as occurred after the study treatment administration on that day. No imputation on other missing time.

- **Missing day and month**

If the year is the same as the year of the first day on double-blind study treatment, then the day and month of the start date of the double-blind study treatment will be assigned to the missing fields.

If the year is different from the year of the first day on double-blind study treatment, then January 1 will be assigned to the missing fields.

- **Missing month only**

Treat day as missing and replace both month and day according to the above procedure.

- **Missing day only**

If the month and year are the same as the year and month of the first day of double-blind study treatment, then the start date of the double-blind study treatment will be assigned to the missing day.

If the month and year are different from the year and month of the first day of double-blind study treatment, then the first day of the month will be assigned to the missing day.

- **Missing year**

No imputation.

## Appendix 1.2. Missing/Incomplete AE Stop Date

If needed, the following imputation rules apply to the case where the end date is incomplete (i.e., partially missing) for adverse events. Other partial end dates will not be imputed. If the imputed stop date is before the start date (imputed or non-imputed start date), then the imputed stop date will be replaced with the start date.

- **Missing day and month, or Missing month only**

December 31 will be assigned to the missing fields. If the imputed stop date is after the date of death, date of death will be used.

- **Missing day only**

If the month and year of the incomplete stop date are the same as the month and year of the last dose date of double-blind study treatment, then the day of the last dose date will be assigned to the missing day.

If the month and year of the incomplete stop date are different from the month and year of the last dose date of double-blind study treatment, then the minimum of (the last day of the month, end of study date, date of death) will be assigned to the missing day.

- **Missing year**

No imputation.

## Appendix 1.3. Missing/Incomplete Prior or Concomitant Medication Start Date

Same imputation rules as missing/incomplete AE onset date (Appendix 1.1).

## Appendix 1.4. Missing/Incomplete Prior or Concomitant Medication Stop Date

Same imputation rules as missing/incomplete AE stop date (Appendix 1.2).

## Appendix 1.5. Missing Date Imputation for Last Dose Date

Imputed last dose date = the earliest date of (last drug dispense date + the number of days of drug dispensed, date of death, date of end of treatment (EOT)/EOS visit, and other dates as appropriate).

# APPENDIX 2. ANALYSIS VISIT WINDOWS

Analysis visits are defined by the windows that will have the widths of the corresponding assessments centered at the scheduled time. Unscheduled visits within a visit window defined below will be grouped into the closest scheduled visits based on the visit date.

**Table 6: Analysis Visit Window for Vital Signs**

| Analysis Visit | Target Day | Start Day | End Day |
|---|---|---|---|
| Baseline | Last value before the first study drug infusion (or randomization date if no infusion received) | | |
| Day 1 postdose | 1 | 1 | 1 |
| Week 3 (predose and postdose) | 22 | 2 | 32 |
| Week 6 (predose and postdose) | 43 | 33 | 53 |
| Week 9 (predose and postdose) | 64 | 54 | 74 |
| Week 12 (predose and postdose) | 85 | 75 | 95 |
| Week 15 (predose and postdose) | 106 | 96 | 116 |
| Week 18 (predose and postdose) | 127 | 117 | 137 |
| Week 21 (predose and postdose) | 148 | 138 | 158 |
| Week 24 (predose and postdose) | 169 | 159 | 179 |
| Week 27 (predose and postdose) | 190 | 180 | 200 |
| Week 30 (predose and postdose) | 211 | 201 | 221 |
| Week 33 (predose and postdose) | 232 | 222 | 242 |
| Week 36 (predose and postdose) | 253 | 243 | 263 |
| Week 39 (predose and postdose) | 274 | 264 | 284 |
| Week 42 (predose and postdose) | 295 | 285 | 305 |
| Week 45 (predose and postdose) | 316 | 306 | 326 |
| Week 48 (predose and postdose) | 337 | 327 | 351 [$] or *[@] |
| 28 Days post last dose [$] | 365 | Date of Last dosing+15 [$] | * [$] |

$ For subjects who do not enroll OLE only.
* End of the double-blind period (DBP).
@ For subjects who enroll OLE.
Note: The study day with infusions will be used for summary tables whenever possible.
Note: If raw data shows visit is “28 Days post last dose”, then set analysis visit=visit
Note: all assessments beyond 60 days post last infusion will be excluded from summary tables or figures.

**Table 7: Analysis Visit Window for PFT**

| Analysis Visit | Target Day | Start Day | End Day |
|---|---|---|---|
| Baseline | Last BTR PFT before the first study drug infusion (or randomization date if no infusion received) | | |
| Week 6 | 43 | 2 | 64 |
| Week 12 | 85 | 65 | 106 |
| Week 18 | 127 | 107 | 148 |
| Week 24 | 169 | 149 | 211 |
| Week 36 | 253 | 212 | 295 |
| Week 48 | 337 | 296 | End of DBP |

**Table 8: Analysis Visit Window for HRCT**

| Analysis Visit | Target Day | Start Day | End Day |
|---|---|---|---|
| Baseline | Last value before the first study drug infusion (or randomization date if no infusion received) | | |
| Week 24 | 169 | 2 | 253 |
| Week 48 | 337 | 254 | End of DBP |

**Table 9: Analysis Visit Window for Labs**

| Analysis Visit | Target Day | Start Day | End Day |
|---|---|---|---|
| Baseline | Last value before the first study drug infusion (or randomization date if no infusion received) | | |
| Week 12 | 85 | 2 | 127 |
| Week 24 | 169 | 128 | 211 |
| Week 36 | 253 | 212 | 295 |
| Week 48 | 337 | 296 | 351 [$] or *[@] |
| 28 days after last dose [$] | 365 | Date of Last dosing+15 [$] | * [$] |

$ For subjects who do not enroll OLE only.
* End of the double-blind period.
@ For subjects who enroll OLE.
Note: If raw data shows visit is “28 Days post last dose”, then set analysis visit=visit
Note: all assessments beyond 60 days post last infusion will be excluded from summary tables for labs.

**Table 10: Analysis Visit Window for SGRQ, UCSD-SOBQ, LCQ, PE, and Weight**

| Analysis Visit | Target Day | Start Day | End Day |
|---|---|---|---|
| Baseline | Last value before the first study drug infusion (or randomization date if no infusion received) | | |
| Week 12 | 85 | 2 | 127 |
| Week 24 | 169 | 128 | 211 |
| Week 36 | 253 | 212 | 295 |
| Week 48 | 337 | 296 | End of DBP |

Note: all assessments beyond 60 days post last infusion will be excluded from summary tables for PE.

**Table 11: Analysis Visit Window for ECG**

| Analysis Visit | Target Day | Start Day | End Day |
|---|---|---|---|
| Baseline | Last value before the first study drug infusion (or randomization date if no infusion received) | | |
| Week 48 | 337 | 2 | End of DBP |

Note: all assessments beyond 60 days post last infusion will be excluded from summary tables for ECG.

**Table 12: Analysis Visit Window for Biomarkers (Serum and Plasma)**

| Analysis Visit | Target Day | Start Day | End Day |
|---|---|---|---|
| Day 1 | 1 | 1 | 1 |
| Week 24 | 169 | 2 | 253 |
| Week 48 | 337 | 254 | 351 [$] or *[@] |
| 28 days Post last dose [$] | 365 | Date of Last dosing+15 [$] | * [$] |

$ For subjects who do not enroll OLE only.
* End of the double-blind period.
@ For subjects who enroll OLE.
Note: If raw data shows visit is “28 Days post last dose”, then set analysis visit=visit
Note: all assessments beyond 60 days post last infusion will be excluded from summary tables.

## Table 13: Analysis Visit Window for CTGF and ADA (Plasma)

| Analysis Visit | Target Day | Start Day | End Day |
|---|---|---|---|
| Day 1 | 1 | 1 | 1 |
| Week 24 | 169 | 2 | 211 |
| Week 36 | 253 | 212 | 295 |
| Week 48 | 337 | 296 | 351 [$] or *[@] |
| 28 days Post last dose [$] | 365 | Date of Last dosing+15 [$] | * [$] |

$ For subjects who do not enroll OLE only.
* End of the double-blind period.
@ For subjects who enroll OLE.
Note: If raw data shows visit is "28 Days post last dose", then set analysis visit=visit
Note: all assessments beyond 60 days post last infusion will be included in summary tables.

## Table 14: Analysis Visit Window for Pharmacokinetic Concentration (Plasma)

| Analysis Visit | Target Day | Start Day | End Day |
|---|---|---|---|
| Day 1 | 1 | 1 | 1 |
| Day 8 | 8 | 2 | 11 |
| Day 15 | 15 | 12 | 18 |
| Week 3 | 22 | 19 | 52 |
| Week 12 | 85 | 53 | 127 |
| Week 24 | 169 | 128 | 211 |
| Week 36 | 253 | 212 | 295 |
| Week 48 | 337 | 296 | 351 [$] or *[@] |
| 28 days Post last dose [$] | 365 | Date of Last dosing+15 [$] | * [$] |

$ For subjects who do not enroll OLE only.
* End of the double-blind period.
@ For subjects who enroll OLE.
Note: If raw data shows visit is "28 Days post last dose", then set analysis visit=visit
Note: all assessments beyond 60 days post last infusion will be included in summary tables.

# APPENDIX 3. SGRQ

The SGRQ is designed to measure health impairment in patients with asthma and COPD. It is also valid for use in bronchiectasis and has been used successfully in patients with kyphoscoliosis and sarcoidosis. It includes Symptoms scores (questions 1-8), Activity scores (questions 11 and 15), and Impacts scores (questions 9-10, 12-14, and 16-17). A Total score is also calculated which summarizes the impact of the disease on overall health status. Scores are expressed as a percentage of overall impairment where 100 represents the worst possible health status and 0 indicates the best possible health status.

**Questions 1 – 7, 9, 10 &17** Where a patient has ticked a box, a value of 1 is entered for the appropriate question. The empty boxes are entered as 0. It will be noted that the questionnaire requests a single response to questions 1-7, 9-10, and 17. If multiple responses are given to one of these questions, then averaging the weights for the positive responses for that question is acceptable.

**Question 8** Where a patient has ticked 'Yes' to having a worse wheeze in the morning, a value of 1 is entered for the appropriate question. All other responses are entered as 0.

**Questions 11 – 16** Where a patient has ticked 'True', a value of 1 is entered for the appropriate question, and where a patient has ticked 'False' a value of 0 is entered. Where a patient has missed a question the cell on the spreadsheet is left blank. In response to question 14, if a patient is not receiving medication, then the response is zero, otherwise, the values are missing.

**Missing Questions**

The missing (blanks) in part 1 (Symptoms) will be imputed by an average of the multiple weighted responses in part 1. This imputation will adjust for up to 24% of missing items in the questionnaire. If more than 24% of items are missing, the component score will be 'Missing'.

**Missed Items for Component Scores**

It is better not to miss items and any missing items are the fault of the experimenter, not the patient. The effect of missing items and recommendations from the scoring manual are the following:

Symptoms Score: The Symptoms component will tolerate a maximum of 2 missed items from questions 1-8.

Activity Score: The Activity component will tolerate a maximum of 4 missed items from questions 11 and 15.

Impacts Score: The Impacts component will tolerate a maximum of 6 missed items from questions 9-10, 12-14, and 16-17.

Therefore, if there are more than the recommended missed items described above the component score will be set to missing. If one of the component scores is missing, then the total score is missing.

**Outline of Scoring Algorithm**

- Principle of calculation

Each questionnaire response has a unique empirically derived 'weight'. The lowest possible 'weight' is zero and the highest 'weight' is 100.

- Each component of the questionnaire is scored separately in three steps:
  - The weights for all items with positive responses are summed.
  - The weights for missed items are deducted from the maximum possible weight for each component. The weights for all missed items are deducted from the maximum possible weight for the Total score.

Sum of maximum possible weights for each component and Total:

| Component | Questions | # of tolerated maximum missing items | Maximum Possible Weight |
|---|---|---|---|
| Symptoms | 1 – 8 | 2 | 662.5 |
| Activity | 11 and 15 | 4 | 1209.1 |
| Impacts | 9 – 10, 12 – 14, 16 – 17 | 6 | 2117.8 |
| Total | All | | 3989.4 |

(Note: these are the maximum possible weights that could be obtained for the worst possible state of the patient).

- The score is calculated by dividing the summed weights by the adjusted maximum possible weight for that component and expressing the result as a percentage:

Score = 100% x (Summed weights from positive items in that component/Sum of weights for all items in that component)

The Total score is calculated similarly:

Score = 100 x (Summed weights from positive items in the questionnaire/Sum of weights for all items in the questionnaire)

**Item Weights**

The following weights will be used to score the answers similar to the Excel-based system and other computerized scoring systems. Note that the wording of the item may not correspond exactly with the wording in the current version of the questionnaire.

- PART 1

**1) Over the past 3 months, I have coughed:**

| **Response** | **Weight** |
|---|---|
| almost every day | 80.6 |
| several days a week | 63.2 |
| a few days a month | 29.3 |
| only with respiratory infections | 28.1 |
| not at all | 0.0 |

**2) Over the past 3 months, I have brought up phlegm (sputum):**

| Response | Weight |
|---|---|
| almost every day | 76.8 |
| several days a week | 60.0 |
| a few days a month | 34.0 |
| only with respiratory infections | 30.2 |
| not at all | 0.0 |

**3) Over the past 3 months, I have had shortness of breath:**

| Response | Weight |
|---|---|
| almost every day | 87.2 |
| several days a week | 71.4 |
| a few days a month | 43.7 |
| only with respiratory infections | 35.7 |
| not at all | 0.0 |

**4) Over the past 3 months, I have had wheezing attacks:**

| Response | Weight |
|---|---|
| almost every day | 86.2 |
| several days a week | 71.0 |
| a few days a month | 45.6 |
| only with respiratory infections | 36.4 |
| not at all | 0.0 |

**5) How many times during the past 3 months have you suffered from severe or very unpleasant respiratory attacks?**

| Response | Weight |
|---|---|
| more than 3 times | 86.7 |
| 3 times | 73.5 |
| 2 times | 60.3 |
| 1 time | 44.2 |
| none of the time | 0.0 |

**6) How long did the worst attack of chest trouble last?**

| Response | Weight |
|---|---|
| a week or more | 89.7 |
| 3 or more days | 73.5 |
| 1 or 2 days | 58.8 |
| less than a day | 41.9 |

**7) Over the past 3 months, in a typical week, how many good days (with few respiratory problems) have you had?**

| Response | Weight |
|---|---|
| No good days | 93.3 |
| 1 or 2 good days | 76.6 |
| 3 or 4 good days | 61.5 |
| nearly every day was good | 15.4 |
| every day was good | 0.0 |

**8) If you wheeze, is it worse when you get up in the morning?**

| Response | Weight |
|---|---|
| No | 0.0 |
| Yes | 62.0 |

- PART 2

**9) How would you describe your respiratory condition?**

| Response | Weight |
|---|---|
| The most important problem I have | 83.2 |
| Causes me quite a lot of problem | 82.5 |
| Causes me a few problems | 34.6 |
| Causes no problem | 0.0 |

**10) If you have ever held a job:**

| Response | Weight |
|---|---|
| My respiratory problems made me stop working altogether | 88.9 |
| My respiratory problems interfere with my job or made me change my job | 77.6 |
| My respiratory problems do not affect my job | 0.0 |

**11) Questions about what activities usually make you feel short of breath these days.**

| Response | Weight |
|---|---|
| Sitting or lying still | 90.6 |
| Washing or dressing yourself | 82.8 |
| Walking around the house | 80.2 |
| Walking outside on level ground | 81.4 |
| Walking up a flight of stairs | 76.1 |
| Walking up hills | 75.1 |
| Playing sports or other physical activities | 72.1 |

**12) More questions about your cough and shortness of breath these days.**

| Response | Weight |
|---|---|
| Coughing hurts | 81.1 |
| Coughing makes me tired | 79.1 |
| I am short of breath when I talk | 84.5 |
| I am short of breath when I bend over | 76.8 |
| My coughing or breathing disturbs my sleep | 87.9 |
| I get exhausted easily | 84.0 |

**13) Questions about other effects your respiratory problems may have on you these days.**

| Response | Weight |
|---|---|
| My cough or breathing is embarrassing in public | 74.1 |
| My chest trouble is a nuisance to my family friends or neighbors | 79.1 |
| I get afraid or panic when I cannot catch my breath | 87.7 |
| I feel that I am not in control of my respiratory problems | 90.1 |
| I do not expect my respiratory problems to get any better | 82.3 |
| I have become frail or an invalid because of my respiratory problems | 89.9 |

| | |
|---|---|
| Exercise is not safe for me | 75.7 |
| Everything seems too much of an effort | 84.5 |

**14) Questions about your respiratory treatment.**

| Response | Weight |
|---|---|
| My treatment does not help me very much | 88.2 |
| I get embarrassed using my medication in public | 53.9 |
| I have unpleasant side effects from my medication | 81.1 |
| My medication interferes with my life a lot | 70.3 |

**15) Questions about how activities may be affected by your breathing.**

| Response | Weight |
|---|---|
| I take a long time to get washed or dressed | 74.2 |
| I cannot take a bath or shower, or I take a long time to do it | 81.0 |
| I walk slower than other people my age, or I stop to rest | 71.7 |
| Jobs such as household chores take a long time, or I have to stop to rest | 70.6 |
| If I walk up one flight of stairs, I have to go slowly or stop | 71.6 |
| If I hurry or walk fast, I have to stop or slow down | 72.3 |
| My breathing makes it difficult to do things such as walk-up hills, carry things upstairs, light gardening such as weeding, dance, bowl or play golf | 74.5 |
| My breathing makes it difficult to do things such as carry heavy loads, dig in the garden or shovel snow, jog or walk briskly (5 miles per hour), play tennis or swim | 71.4 |
| My breathing makes it difficult to do things such as very heavy manual work, ride a bike, run, swim fast, or play competitive sports | 63.5 |

**16) We would like to know how your respiratory problems usually affect your daily life.**

| Response | Weight |
|---|---|
| I cannot play sports or do other physical activities | 64.8 |
| I cannot go out for entertainment or recreation | 79.8 |
| I cannot go out of the house to do the shopping | 81.0 |
| I cannot do housework chores | 79.1 |
| I cannot move far from my bed or chair | 94.0 |

**17) Tick the statement which you think best describes how your chest affects you.**

| Response | Weight |
|---|---|
| It does not stop me from doing anything I would like to do | 0.0 |
| It stops me from doing one or two things I would like to do | 42.0 |
| It stops me from doing most of the things I would like to do | 84.2 |
| It stops me from doing everything I would like to do | 96.7 |

# APPENDIX 4. UCSD-SOBQ

The SOBQ asks subjects to indicate the severity of shortness of breath on a 6-point scale (0 = Not at all, to 5 = Maximally or unable to do because of breathlessness) during 21 activities of daily living associated with varying levels of exertion. Three additional questions ask about fear of harm from overexertion, limitations, and fear caused by shortness of breath, for a total of 24 items. If patients do not routinely perform an activity, they are asked to estimate their anticipated shortness of breath. The SOBQ is scored by summing non-missing responses across all 24 items to form a total score. A total sum score ranges from 0 to 120, with higher scores indicating greater dyspnea.

If 22-23 questions were answered, the total SOBQ score will be the sum of non-missing questions*24/number of questions answered assuming missing items is neutral. If 21 or fewer questions were answered, the SOBQ score will be set to missing.

# APPENDIX 5. LCQ

The LCQ is a 19 Response questionnaire that assesses cough-related QOL. It has 3 domains (physical, psychological, and social):

- Physical: Questions 1, 2, 3, 9, 10, 11, 14, 15
- Psychological: Questions 4, 5, 6, 12, 13, 16, 17
- Social: Questions 7, 8, 18, 19

The total score range is 3 to 21 and domain scores range from 1 to 7; a higher score indicates a better quality of life. Scores are calculated as a mean of each domain and the total score is calculated by adding every domain score. Only non-missing scores are considered in the mean domain score calculation. Two missing values are allowed for Physical, and 1 missing value is allowed for Social and Psychological domains. Total scores require non-missing scores for 3 domains.

# APPENDIX 6. GAP SCORE AND STAGE

GAP index and staging system for IPF: gender (G), age (A), and two pulmonary physiological parameters (P)—percentage predicted FVC [%], and percentage predicted DLCO [%].

GAP Stage I: 0-3 points; Stage II: 4-5 points; Stage III: 6-8 points.

| | Predictor | Points |
|---|---|---|
| **G: Gender** | Female | 0 |
| | Male | 1 |
| **A: Age (years)** | ≤60 years | 0 |
| | 61-65 years | 1 |
| | >65 years | 2 |
| **P: Predicted Forced Vital Capacity (FVCpp)** | >75% | 0 |
| | 50-75% | 1 |
| | <50% | 2 |
| **P: Predicted Diffusing Capacity of the Lung for Carbon Monoxide (DLCOpp)** | >55% | 0 |
| | 36-55% | 1 |
| | ≤35%* | 2 |
| | Cannot perform | 3 |
| **Total Possible Points** | | **8** |

*Value < 35.5 is rounded down to 35; value>=35.5 is rounded up to 36.

# APPENDIX 7. EXAMPLE SAS CODE

## Appendix 7.1. MMRM

Sample SAS code for MMRM:

```
PROC MIXED DATA=mmrm;
       CLASS trtp randomization_factor avisit subjid;
       MODEL chg = trtp avisit trtp*avisit covariates randomization_factor/DDFM=kr;
       REPEATED avisit/SUBJECT=subjid TYPE=UN GROUP=trtp;
       LSMEANS trtp/PDIFF CL;
       LSMEANS trtp*avisit/PDIFF CL;
RUN;
```

## Appendix 7.2. Jump-to-Control

Under the jump-to-control assumption, the analysis will be carried out in 3 steps.

Step 1 - the missing FVC data will be imputed to derive 200 imputed datasets with non-missing data according to the jump-to-control data pattern.

Sample SAS code:

```
/*generate a monotone missing data pattern*/
PROC MI DATA=aval OUT=xx1 SEED=9978 NIMPUTE=200 ROUND=0.001 MINIMUM=0.5 MAXIMUM=7;
       BY trtp;
       VAR base week6 week12 week18 week24 week36 week48;
       MCMC CHAIN=multiple IMPUTE=monotone;
RUN;

PROC SORT DATA=xx1;
       BY _imputation_ subjid;
RUN;

/*Generate MI datasets with non-missing data via a Jump-to-Control approach*/
PROC MI DATA=xx1 SEED=1482 NIMPUTE=1 OUT=mnar;
       CLASS trtp;
       BY _IMPUTATION_;
       VAR base week6 week12 week18 week24 week36 week48;
       MONOTONE REG (trtp base week6 week12 week18 week24 week36 week48);
       MNAR model (week6 week12 week18 week24 week36 Week48/
       MODELOBS=(trtp='Placebo'));
RUN;
```

Step 2 - The 200 multiple-imputation datasets with imputed and observed FVC data at Week 48 will be analyzed separately for each imputation using the ANCOVA method. The ANCOVA model will contain terms for treatment, baseline FVC measurements, covariates (age, sex, race group, and height), and the randomization stratification factor. The LSMean and corresponding SE for the change from baseline in FVC at Week 48 will be estimated.

Sample SAS code:

```
PROC MIXED data=mnar;
       BY _imputation_;
       CLASS trtp randomization_factor;
       MODEL chg_wk48 = trtp base covariates randomization_factor/SOLUTION DDFM=kr;
       LSMEANS trtp / PDIFF CL;
       ODS OUTPUT DIFFS=lsdiffs LSMEANS=lsm SOLUTIONF=parms;
```

```
RUN;
```

Step 3 - The SAS PROC MIANALYZE will be used to derive the final estimates and test statistics summarizing these 200 analysis results.

Sample SAS code:

```
PROC SORT DATA=lsm;
       BY trtp;
RUN;

PROC MIANALYZE DATA=lsm;
       BY trtp;
       MODELEFFECTS estimate;
       STDERR;
       ODS OUTPUT ParameterEstimates=lsm2;
RUN;

PROC MIANALYZE DATA=lsdiffs;
       MODELEFFECTS estimate;
       STDERR;
       ODS OUTPUT ParameterEstimates=lsdiffs2;
RUN;
```

## Appendix 7.3. Delta-Adjusting (Tipping Point) Analysis

The multiple imputations for tipping point analysis will be performed as follows.

Step 1 - the missing FVC data will be imputed to derive 200 imputed datasets with non-missing data according to the delta-adjusting data approach.

Sample SAS code:

```
/*generate a monotone missing data pattern*/
PROC MI DATA=aval OUT=xx1 SEED=9978 NIMPUTE=200 ROUND=0.001 MINIMUM=0.5 MAXIMUM=7;
       BY trtp;
       VAR base week6 week12 week18 week24 week36 week48;
       MCMC CHAIN=multiple IMPUTE=monotone;
RUN;

PROC SORT DATA=xx1;
       BY _imputation_ subjid;
RUN;

%let sft=-0.3;/*Define a macro for shift value, adjust the value in SAS code*/

/*Generate MI datasets with non-missing values using the tipping point approach*/
PROC MI DATA=xx1 SEED=1482 NIMPUTE=1 OUT=tipping;
       CLASS trtp;
       BY _IMPUTATION_;
       VAR trtp base week6 week12 week18 week24 week36 week48;
       MONOTONE REG (base week6 week12 week18 week24 week36 week48);
       MNAR ADJUST(week6 / SHIFT=&sft ADJUSTOBS=(trtp='Pamrevlumab'));
       MNAR ADJUST(week12 / SHIFT=&sft ADJUSTOBS=(trtp='Pamrevlumab'));
       MNAR ADJUST(week18 / SHIFT=&sft ADJUSTOBS=(trtp='Pamrevlumab'));
       MNAR ADJUST(week24 / SHIFT=&sft ADJUSTOBS=(trtp='Pamrevlumab'));
       MNAR ADJUST(week36 / SHIFT=&sft ADJUSTOBS=(trtp='Pamrevlumab'));
       MNAR ADJUST(week48 / SHIFT=&sft ADJUSTOBS=(trtp='Pamrevlumab'));
RUN;
```

Step 2 – The 200 multiple imputation datasets will be analyzed separately using the same ANCOVA model as Step 2 in Appendix 7.2.

Step 3 – Same as Step 3 in Appendix 7.2.

## Appendix 7.4. Ranked ANCOVA

Sample SAS code for Ranked ANCOVA:

```
PROC RANK OUT=ranked TIES=mean;
      WHERE avisit='Week 48';
      VAR base chg;/*Missing chg was imputed*/
RUN;

PROC MIXED DATA=ranked;
      CLASS sex randomization_factor;
      MODEL chg= base age height race sex randomization_factor/OUTP=residual;
RUN;

PROC FREQ DATA=residual;
      TABLES trtp*resid / NOPRINT CMH2;
      OUTPUT OUT=rancova CMH2;
RUN;
```

# APPENDIX 8. CHINA CDE REQUIREMENT

Efficacy data, mainly including overseas key clinical trial data and clinical trial data conducted in China, should not only confirm the efficacy of the study drug as a whole, but also analyze the consistency between Chinese subgroups and the overall population.

Safety data, including all domestic and foreign data used for safety evaluation, should be analyzed not only for overall safety but also for consistency between Chinese subgroup and overall population.

All TLF in China may be provided per China regulatory requirements if regulatory submission in China is pursued (NMPA guidance).

Signature Page for VV-CLIN-013733 v1.0

| | |
|---|---|
| Approval Task | [redacted] 13-Jun-2023 16:43:52 GMT+0000 |
| Approval Task | [redacted] 13-Jun-2023 16:44:35 GMT+0000 |
| Approval Task | [redacted] 13-Jun-2023 16:48:10 GMT+0000 |
| Approval Task | [redacted] 13-Jun-2023 17:13:12 GMT+0000 |
| Approval Task | [redacted] 13-Jun-2023 17:15:19 GMT+0000 |
| Approval Task | [redacted] 13-Jun-2023 17:27:54 GMT+0000 |
| Approval Task | [redacted] 13-Jun-2023 17:40:19 GMT+0000 |
| Approval Task | [redacted] 13-Jun-2023 17:59:12 GMT+0000 |

Signature Page for VV-CLIN-013733 v1.0

Signature Page for VV-CLIN-013733 v1.0

| | |
|---|---|
| Approval Task | [redacted]<br>13-Jun-2023 18:00:02 GMT+0000 |

Signature Page for VV-CLIN-013733 v1.0