NCT #: NCT03496623

Statistical Analysis Plan: RIN-PH-304

Study Title:	A Phase 3, Randomized, Placebo-controlled, Double-blind, Adaptive Study to Evaluate the Safety and Efficacy of Inhaled Treprostinil in Patients with Pulmonary Hypertension due to Chronic Obstructive Pulmonary Disease (PH-COPD)
Study Number:	RIN-PH-304
Study Phase:	3
Product Name:	Inhaled Treprostinil
Indication:	Treatment of PH-COPD
IND Number:	070362
Sponsor:	United Therapeutics Corp. [redacted]
Author:	[redacted]
Final Date:	07 January 2022
Version:	2.0

TABLE OF CONTENTS

Table of In-Text Tables

ABBREVIATIONS AND DEFINITIONS

6MWD	6-Minute Walk Distance
6MWT	6-Minute Walk Test
AE	Adverse event
BMI	Body mass index
COPD	Chronic obstructive pulmonary disease
CRF	Case Report Form
CSR	Clinical Study Report
DLCO	Diffusion capacity of lung for carbon monoxide
DSMC	Data Safety Monitoring Committee
DSP	Distance saturation product
ECG	Electrocardiogram
FAS	Full Analysis Population
FCS	Fully conditional specification
FEF	Forced expiratory flow
FEV_1	Forced expiratory volume in 1 second
FEV_6	Forced expiratory volume in 6 seconds
FVC	Forced vital capacity
HR	Heart rate
ICF	Informed Consent Form
LS	Least Square
MAR	Missing at random
MET	Metabolic equivalent task
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed-effect model for repeated measures
mPAP	Mean pulmonary artery pressure
MVPA	Moderate to vigorous physical activity
NT-proBNP	N-terminal pro-brain natriuretic peptide
PEF	Peak expiratory flow
PFT	Pulmonary function test
PGA	Patient global assessment
PH	Pulmonary hypertension
PH-COPD	Pulmonary hypertension due to chronic obstructive pulmonary disease
PR	Electrocardiographic wave of P-R interval
PT	Preferred Term
PVR	Pulmonary vascular resistance
QID	4 times daily
QOL	Quality of life
QRS	Electrocardiographic wave of QRS interval
QT	Electrocardiographic wave of QT interval
RHC	Right heart catheterization
RVSP	Right ventricular systolic pressure

SAE	Serious adverse event
SAP	Statistical Analysis Plan
SGRQ	St. George's Respiratory Questionnaire
SOC	System Organ Class
SpO_2	Saturation peripheral capillary oxygenation
TLC	Total lung capacity
UCSD SOBQ	University of California San Diego Shortness of Breath Questionnaire
WHO	World Health Organization

1 PREFACE

This plan provides further details of the planned analyses for Study RIN-PH-304 as presented in the study protocol. Based on the original Study Protocol RIN-PH-304 dated 12 January 2018 and the subsequent study protocol amendments (latest version study protocol Amendment 4 dated 20 October 2020), it provides further details of the planned analyses described in the study protocol, as well as any additional planned analyses. Additional post hoc or unplanned analyses that are not defined in this statistical analysis plan (SAP) may be performed. Such analyses will be documented in the clinical study report (CSR).

The RIN-PH-304 study is an adaptive study with a (fully blinded) decision for design modification based on enrollment and missing data metrics. This SAP pre-specifies analyses under the original design, as described in the study protocol (and associated amendments). To address the adaptive design feature of this study for a Contingent Parallel Design, as defined by the criteria in Section 10.2.1 of the study protocol, this SAP has been formally amended to provide analytic plans and details for the Contingent Parallel Design in Appendix 12.4.

[redacted]

In light of the missing data arising from COVID-19, this SAP incorporates the ability to invoke a new definition for the primary endpoint based on a fully blinded assessment for missing data. This SAP defines the decision rules, data handling, and analyses for the updated definition of the

primary endpoint. In general, the amount of missing 6MWD data potentially due to the COVID-19 pandemic interfering with onsite measurements will define if adaptation is necessary. At the conclusion of the study, a blinded analysis will be performed to assess the extent of missing primary efficacy endpoint data (ie, 6MWD data). If the absence of 6MWD data, combined across subjects, treatment periods, and key visits (Baseline and Study Week 12) to be used for the final primary endpoint analysis would likely lead to an underpowered study or yield indeterminate results, an alternate primary endpoint will be used for the final primary endpoint analysis (see Section 9.2). This adaptation intends that the 6MWD is to be used as the primary efficacy endpoint for this study, except in the case where onsite measurements for the 6MWD are not available, and will occur regardless of Adaptation 1 (ie, for either the Original Crossover Design or the Contingent Parallel Design).

2 STUDY OBJECTIVES AND ENDPOINTS

2.1 OBJECTIVES

The primary objective of this study is to demonstrate the efficacy of inhaled treprostinil compared to placebo in improving exercise ability as measured by change from baseline in 6MWD following 12 weeks of active treatment in subjects with pulmonary hypertension due to chronic obstructive pulmonary disease (PH-COPD).

All secondary, safety, and exploratory objectives can be found in Section 2 of the study protocol.

2.2 PRIMARY ENDPOINT

The primary efficacy endpoint for this study is the change in 6MWD from Baseline to Week 12 in subjects with PH-COPD.

2.3 ALTERNATIVE PRIMARY ENDPOINT

An alternative primary efficacy endpoint of 6MWD/MVPA ranks will replace the change in 6MWD from Baseline to Week 12 if pre-specified rules (see Section 9.2 of this SAP for further details of the alternative primary endpoint and Section 9.3.1.2 for details regarding MVPA derivation).

2.4 SECONDARY EFFICACY ENDPOINTS

The secondary efficacy endpoints are the comparison for the effect of inhaled treprostinil versus placebo after 12 weeks of active treatment. Specifically:

- Change in MVPA as measured by actigraphy from Baseline to Week 12
- Change in overall activity as measured by actigraphy from Baseline to Week 12
- Change in Borg dyspnea score from Baseline to Week 12
- Change in 6MWD/Borg dyspnea composite score from Baseline to Week 12
- Change in quality of life (QOL) measured by the St. George's Respiratory Questionnaire (SGRQ) and the University of California San Diego Shortness of Breath Questionnaire (UCSD SOBQ) from Baseline to Week 12
- Change in plasma concentration of N-terminal pro-brain natriuretic peptide (NT-proBNP) levels from Baseline to Week 12
- Change in PGA from Baseline to Week 12.

2.5 SAFETY ENDPOINTS

The safety of inhaled treprostinil will be evaluated compared to placebo on the following parameters in comparison to placebo:

- AEs
- Clinical laboratory parameters
- Electrocardiograms (ECGs)
- Physical examination (PE) findings
- Oxygenation
  - Pulse oximetry (saturation of peripheral capillary oxygenation [SpO_2])
  - Supplemental oxygen requirement (L/min)
- Pulmonary function tests (PFTs)
  - Forced expiratory volume in 1 second (FEV1)
  - Forced vital capacity (FVC)
  - Total lung capacity (TLC)
  - Diffusion capacity of lung for carbon monoxide (DLCO)
- Vital sign measurements
- At-home spirometry.

2.6 EXPLORATORY ENDPOINTS

[redacted]

[redacted]

[redacted]

3 STUDY DESIGN

3.1 GENERAL DESCRIPTION

This is a multicenter, randomized, double-blind, placebo-controlled, 34-week, crossover study of inhaled treprostinil (Tyvaso® [treprostinil] Inhalation Solution) in males and females 18 years or older diagnosed with WHO Group 3 pulmonary hypertension (PH) associated with chronic obstructive pulmonary disease (COPD). Subject eligibility will be based on the inclusion and exclusion criteria described in Section 4 of the protocol.

Approximately 136 eligible subjects will be randomized with half of the subjects receiving 12 weeks of active study drug (Period 1) then 12 weeks of placebo (Period 2) and the other half receiving 12 weeks of placebo (Period 1) then 12 weeks of active study drug (Period 2). Under the original study design, all randomized subjects will be treated with both active study drug and placebo.

The study will consist of the following periods:

Screening Period: At Screening Visit 1, prospective subjects (defined as having prior echocardiogram findings of right ventricular systolic pressure [RVSP] >40 mmHg and/or evidence of right ventricular hypertrophy or dysfunction), who sign the Informed Consent Form (ICF) will generally undergo all non-invasive screening assessments to evaluate eligibility, as listed in the inclusion/exclusion criteria. Potentially eligible subjects will undergo a right heart catheterization (RHC), including a vasodilator test and oxygen challenge, if possible, to confirm eligibility. If RHC results are available from a test performed within the previous 12 months (even if done without oxygen or vasodilator challenge), a repeat RHC during Screening Visit 1 is not needed. A window of up to 6 weeks is permitted to complete all assessments required for Screening Visit 1, including the start of low dose inhaled treprostinil. Prospective subjects meeting all eligibility criteria will be provided with and trained on inhalation of study drug using the Tyvaso Inhalation System and begin low dose (3 breaths QID) inhaled treprostinil for a minimum of 14 days (maximum of 18 days) to ensure tolerability and compliance. Subjects will be issued Sponsor-provided actigraphy, spirometry, and smart devices during Screening Visit 1 at the time of low dose inhaled treprostinil initiation for measurement of at-home activity and FEV_1/FVC. Subjects will be given training on the at-home use of these devices, as well as Instructions for Use. At Screening Visit 2, subjects will return to the study site following completion of low dose inhaled treprostinil exposure to undergo scheduled assessments and return unused study drug. Subjects who are intolerant of therapy (as determined by the Investigator), unable to follow the dosing regimen, or noncompliant with actigraphy, at-home spirometry, or smart device use during the Screening Period will be ineligible for randomization and deemed screen failures.

Washout Period - Washout 1: During Washout 1 (Study Week -1), prospective subjects will refrain from the use of inhaled treprostinil for a minimum of 7 days (maximum of 14 days) preceding entry into the Treatment Period 1 Baseline Visit (Study Week 1). Subjects will be contacted by telephone mid-week in Washout Period 1 to monitor safety.

Treatment Period 1: Eligible subjects approved by the Sponsor's Medical Monitor will undergo Treatment Period 1 Baseline Visit assessments (Study Week 1), be assigned to a treatment group based on the randomization schedule, and receive study drug. Target dosing with study drug (after titration) will be 12 breaths QID or to maximum tolerated dose. Two additional study visits to the study site will be conducted after the Treatment Period 1 Baseline Visit to complete study assessments (Study Weeks 6 and 12). Subjects will be contacted weekly by telephone during study drug titration (until 12 breaths QID or the maximum tolerated dose is achieved) and as needed thereafter to provide dosing instructions and to assess tolerance to study drug, AEs, and changes in concomitant medications.

Washout Period - Washout 2: Subjects will refrain from the use of study drug for a minimum of 7 days (maximum of 14 days) after the Treatment Period 1 Week 12 Visit (Study Week 12) and preceding entry into Treatment Period 2 (Study Week 14). Subjects will be contacted by telephone mid-week in Washout Period 2 to monitor safety.

Treatment Period 2: Subjects will undergo Treatment Period 2 Baseline Visit assessments (Study Week 14), be provided treatment based on the randomization schedule, and receive study drug for this period. Target dosing with study drug (after titration) will be 12 breaths QID or to the maximum tolerated dose. Two additional study visits to the study site (Study Weeks 19 and 25) will be conducted after the Treatment Period 2 Baseline Visit (Study Week 14) to complete study assessments. Subjects will be contacted weekly by telephone during study drug titration (until 12 breaths QID or the maximum tolerated dose is achieved) and as needed thereafter to provide dosing instructions and to assess tolerance to study drug, AEs, and changes in concomitant medications. At the completion of Treatment Period 2 (Study Week 25), subjects will be offered enrollment in the open-label extension study, RIN-PH-305.

4 SEQUENCE OF PLANNED ANALYSES

In brief, the specified primary endpoint for this study is 6MWD, which is a validated and reliable measure of exercise ability in patients with chronic respiratory diseases. Subjects perform the test at the clinical site following published guidelines and tests are conducted by trained personnel. For further details regarding the 6MWT and 6MWD, please see the study protocol.

The alternative primary endpoint for this study is a combined and standardized rank for 6MWD and MVPA, which is a measurement of subjects' time spent in moderate to vigorous activity during the course of daily living. The study protocol provides further detail about actigraphy and Section 9.3 of this SAP gives information about how blinded assessments will occur, the handling of data for actigraphy and MVPA, etc.

4.1 INTERIM ADAPTIVE ANALYSIS

Under the Original Crossover Design, at approximately 75% of subjects randomized, a blinded interim analysis will be performed to assess the extent of missing data for the primary analysis for the crossover design. The outcome of the analysis will inform a decision for the study to continue as the Original Crossover Design or switch to the Contingent Parallel Design as follows:

- If data presence ≥85% for 6MWD or actigraphy data at the end of Treatment Period 2 (Study Week 25), the study will continue as planned under the Original Crossover Design.
- If data presence <85% for both 6MWD and actigraphy data at the end of Treatment Period 2 (Study Week 25), the study will switch to the Contingent Parallel Design.

The goal of this assessment is to determine the amount of missing data due to subject dropout or otherwise lack of measurement (regardless of onsite or remote) and the outcome will be whether study continues to completion as the Original Crossover Design or the Contingent Parallel Design. Should there be a switch to the Contingent Parallel Design, the analysis plan described in Appendix 12.4 will be followed.

4.2 INTERIM EFFICACY ANALYSIS

A single interim efficacy analysis will be conducted when 75% of the subjects (ie, ~100 subjects) have completed the 26-week Treatment Period if both of the following criteria are met: a) the expected time to randomize the final (136^{th}) subject is greater than 3 months; b) at least 90% of the 6MWD data is present for both Period 1: Baseline and Week 12 and Period 2: Baseline and Week 12 for these subjects. The projected time to randomize will be based on the average recruitment rate for the prior 6 months. This plan will allow early stopping for efficacy if recruitment is slow enough to warrant an interim review and if the data are sufficiently complete.

A non-binding futility analysis will also be conducted as part of the interim analysis. Analyses will be conducted for the primary endpoint of 6MWD per Section 9.1. Analyses will use the entire Full Analysis Population (FAS), including all available data for subjects who may be still ongoing in the study. The independent external statistical consultant will provide the results of these analyses to the independent Data Safety Monitoring Committee (DSMC) for their review and recommendation. The Sponsor will remain fully blinded throughout this process.

The DSMC Charter will contain more specific details regarding the processes for the interim analyses for efficacy and futility, such as timelines for data lock, information transfer and communications. However, the following approaches will be taken for projecting recruitment and the criteria for the interim analyses for efficacy.

4.2.1 [redacted]

[redacted]

4.2.2 Formal Criteria for Interim Efficacy Analyses

If conducted, the interim analysis will apply a 2-sided alpha level of 0.010 (75% information fraction), and at the final analysis at a 2-sided alpha of 0.022 with the overall Type I error rate at 2-sided alpha of 0.05. Efficacy boundaries for early stopping of the study for efficacy are based

[1] Subjects who were active in the treatment periods of this protocol as of 13 March 2020 will be assessed on a case-by-case basis by the Sponsor for inclusion of populations for adaptive, interim, and final analyses. All decisions will be documented and presented in the clinical study report.

on a Lan-DeMets spending function, with an O'Brien-Fleming boundary (GD-INFO; EAST®, Cytel, Version 6.5). The corresponding non-binding futility boundary is provided as p-value less than or equal to 0.227 for the interim analysis (Lan-DeMets: O'Brien-Fleming).

4.3 ALTERNATIVE PRIMARY ENDPOINT

To address the possibility of missing data arising from COVID-19, this SAP incorporates the ability to determine a new definition for the primary endpoint based on a fully blinded assessment for missing data. Section 9.2 of the SAP defines the decision rules, data handling, and analyses for the alternative definition of the primary endpoint. In general, the amount of missing 6MWD data potentially due to the COVID-19 pandemic interfering with onsite measurements will define whether adaptation is necessary. At the conclusion of the study, a blinded analysis will be performed to assess the extent of missing primary efficacy endpoint data (ie, 6MWD data). If the absence of 6MWD data in the Treatment Period, combined across subjects, treatment periods, and key visits (Baseline and Study Week 12) to be used for the final primary endpoint analysis would likely lead to an underpowered study or yield indeterminate results, an alternative primary endpoint will be used for the final primary endpoint analysis (see Section 9.2 for decision rules). The 6MWD is to be used as the primary efficacy endpoint for this study, except in the case where onsite measurements for the 6MWD are not practically available, and will occur independently of Adaptation 1 (ie, for either the Original Crossover Design or the Contingent Parallel Design).

4.4 FINAL ANALYSIS

The final analysis is planned after all subjects have completed treatment or withdrawn from the study, the study database is locked, and the randomization codes have been revealed. All statistical analyses will be performed by the Sponsor's biostatistics department personnel (or appropriate designees) using SAS®, Version 9.4 (or higher) or other validated software.

5 SAMPLE SIZE CONSIDERATIONS

5.1 PLANNED STUDY

The **Original Crossover Design** is a 2x2 crossover study designed to have 90% power to detect a placebo-corrected mean change from baseline difference of 30 meters, assuming that the

standard deviation of differences is 85 meters[2] and using a 2-group t-test with a 0.01 2-sided significance level, with a sample size of 124 subjects (62 in each sequence group). To account for a discontinuation rate of approximately 10%, the total sample size will be approximately 136 subjects.

The **Contingent Parallel Design** is a parallel-group study designed to use an allocation ratio of 1:1 between inhaled treprostinil and placebo, with a sample size of 266 subjects (133 per treatment) that would provide at least 90% power at a significance level of 0.05 (2-sided hypothesis) to detect a 30 meter between-treatment difference in the change from baseline in 6MWD, assuming a standard deviation of 75 meters. To account for a discontinuation rate of approximately 15%, the total sample size will be approximately 314 subjects. Should there be a switch to the Contingent Parallel Design, the analysis plan described in Appendix 12.4 will be followed.

5.2 [redacted]

[redacted]

[2] Note. Due to the smaller sample size expected for the crossover design a more conservative standard deviation of the difference (85 meters) was assumed (compared to 75 meters for a parallel design).

[redacted]

6 ANALYSIS POPULATIONS

This study has the following 4 analysis populations[3]:

- Randomized Population (RAND): All subjects randomized to the study. Tabulations for disposition of subjects will be generated from this population.
- Full Analysis Population (FAS): All randomized subjects who received at least 1 dose of study treatment (treprostinil or placebo). All subjects will be counted in the group to which they were randomized, regardless of the study treatment they were given. All efficacy analyses will be performed on this population, unless otherwise specified.
- Per-Protocol Analysis Population (PPROT): All randomized subjects who are included in the FAS but excluding subjects with major protocol deviations that may have an impact on the primary efficacy analyses. The major protocol deviations and the subject's exclusion from the Per-Protocol Analysis Population will be reviewed at a blinded data review meeting and documented prior to the database lock and the study unblinding.
- Safety Population (SAFETY): All enrolled subjects who received at least 1 dose of treprostinil, during low dose run-in, or study treatment (treprostinil or placebo) during the randomized portion of the study. All Safety Population subjects will be counted in the group corresponding to the study treatment received, regardless of randomized

[3] Subjects who were active in the treatment periods of this protocol as of 13 March 2020 will be assessed on a case-by-case basis by the Sponsor for inclusion of populations for adaptive, interim, and final analyses. All decisions will be documented and presented in the clinical study report.

assignment. All safety analyses will be performed on this Safety Population, unless otherwise specified.

A summary table presenting the number and percentages of subjects for each analysis population will be presented by treatment group. Subjects who are excluded from any analysis population will be listed with reason for exclusion. The Sponsor must approve this list before treatment is unblinded and thus before performing final analyses.

7 GENERAL CONSIDERATIONS FOR DATA ANALYSES

Given the crossover design utilized for this study, efficacy and safety summary tables will follow the general format that baseline and post-baseline data will be presented by the assigned treatment condition when those data are observed during the treatment period. Treatments will be labeled as "Treprostinil" and "Placebo" to correspond with the assigned treatment condition.

Any efficacy and safety data that are collected during the washout period between Periods 1 and 2 will be presented for the assigned treatment condition from Period 1. Any follow-up data collected after Period 2, while on this study, will be assigned to Period 2. Subjects who are not retained will not be replaced during the study.

For continuous variables, the summary statistics will include the mean, standard deviation, median, lower quartile, upper quartile, minimum, and maximum. Statistics will be expressed using the level of precision in which the variable was collected. For discrete variables, summaries will include the frequency and percent in each category. Whenever practical, categories of discrete variables will be ordered and labeled as they appear in the case report form (CRF), and all categories represented on the CRF will be included in summaries, even when they do not apply to any patients in the study.

For data collected on a fixed schedule, the nominal time point will be used for summarization. Evaluations from Unscheduled Visits are to be mapped to Scheduled Visits to the extent possible per the protocol-specified windows. Repeat or redundant observations within an assessment window and observations that do not fall within any predefined assessment window will be excluded from summaries. If an evaluation is repeated, such as to confirm an observed aberrant laboratory value, the analyses will use the later value if within the protocol-specified window.

Unless otherwise specified, all final statistical tests will be 2-sided at alpha level 0.05 if the interim analysis is not conducted, and 0.022, if conducted. All tables will be presented in landscape format. For practical reasons, the point size may be reduced to (but not less than) 7 point for tables that contain too much information to fit into a single page. SAS® version 9.4 or higher will be used. The Sponsor's Biostatistics Department will analyze the efficacy and safety data.

7.1 BASELINE

Given the crossover design utilized for this study, Periods 1 and 2 each have a different baseline for efficacy measurements: Period 1: Day 1 and Period 2: Week 14. With the exception of actigraphy measurements, the baseline value will be defined as the non-missing value immediately prior to dosing with study drug for Period 1: Day 1 and Period 2: Week 14. The baseline period for actigraphy measurements will be the last 7 days during the washout period prior to the day of dosing on Day 1 and Week 14. In the event that the washout period is less than 7 days then only the non-dosing days will be used to derive a baseline average.

For safety evaluations, the baseline will be defined as the last evaluation closest, but prior to, the first dose of study drug during the low dose run-in period.

7.2 CHANGE FROM BASELINE

For parameters measured at baseline, the variables of interest are the change from baseline of the original measurements. Unless otherwise specified, all changes from baseline (for all variables where this is applicable) will be calculated as follows:

Change from baseline = Post baseline – baseline

7.3 STUDY DAY

Study day will be calculated relative to first dose of study drug (Day 1), after randomization. Specifically:

Study day before dosing = (Date of occurrence – Date of first dose of study drug).
Study day on or after dosing = (Date of occurrence – Date of first dose of study drug) + 1.

7.4 COVARIATES

Efficacy analyses will include the assigned treatment sequence and period order as covariates. Change from baseline analyses will include the baseline measurement as a continuous covariate.

7.5 EXAMINATION OF SUBGROUPS

Primary and secondary efficacy analyses will be conducted on the FAS. Subgroup analyses will be performed to explore the efficacy of treprostinil in different subject populations. The subgroups will include, but not be limited to:

- age (<65, ≥65 years);
- sex (males, females);
- race (White, Black or African-American, Asian, American Indian or Alaskan Native, Native Hawaiian or Other Pacific Islander);
- ethnicity (Hispanic or Latino, non-Hispanic or Latino);
- age at WHO Group 3 PH-COPD diagnosis (date of diagnosis minus date of birth, in years, by quartiles);
- pulmonary vascular resistance (PVR; during screening, in quartiles);
- mean pulmonary artery pressure (mPAP; during screening, in quartiles);
- baseline (Day 1) 6MWD (≤350 meters versus >350 meters);
- baseline (Day 1) Borg dyspnea score (in quartiles).

Unless otherwise prespecified, continuous covariates will be presented by quartiles. Comparative results will be presented using point estimates and associated 95% CIs both in tables and forest plots. Because the study was not designed to have high power for small or even moderate sized subgroups, these analyses will be considered exploratory. No Type 1 error adjustments will be performed for these subgroup analyses.

7.6 PREMATURE DISCONTINUATION AND MISSING DATA

Investigators and site staff are to encourage all subjects to continue their participation in the study and complete all study procedures and study visits as indicated by the study protocol unless the subject provides a written withdrawal of consent from the study. A subject's discontinuation from study drug does not mean discontinuation from the study. If study drug is discontinued for a randomized subject, the subject will be encouraged to actively continue to

return for all study visits per the study protocol, and the Investigator will continue all reasonable efforts to obtain the subject's data.

Patients may not complete the full treatment periods for the following reasons: death, lost to follow-up, study terminated by Sponsor, COVID-19 truncation, or withdrawal of consent. All available data from all patients will be used as detailed in this analysis plan. In addition, subjects may be too ill to perform assessments, such as the 6MWT or Borg, resulting in missing data for that assessment.

Methods for imputation and handling missing data, when applicable, will be described in the sections for the particular outcome.

7.7 MULTIPLE COMPARISONS AND MULTIPLICITY

A gatekeeper strategy will be employed to address multiple comparisons. Secondary endpoints will be formally tested only if the primary endpoint is statistically significant. If the primary analysis of treprostinil vs placebo change from Baseline to Week 12 is statistically significant at a 2-sided alpha level of 0.05, showing superiority of treprostinil over placebo, then the secondary endpoints will be formally tested.

[redacted]

Other secondary endpoints, including change in QOL, based on the SGRQ and UCSD SOBQ, and PGA will not have formal statistical testing.

8 STUDY POPULATION

8.1 PATIENT ACCOUNTABILITY

8.1.1 Treatment Compliance

At each study visit, all study drug returned by the subject (used and unused) will be collected and new study drug will be dispensed. The appropriate study personnel must document the number of used and unused ampoules and determine within the interactive voice/web response system.

Subject compliance with the prescribed dosage regimen will be monitored throughout the study. If it is determined that a subject is not compliant with study drug, the site personnel must re-educate the subject on proper dosing compliance and its importance.

Treatment compliance will be presented as a table summarizing the counts and summary statistics for used and unused ampoules by the assigned treatment.

8.1.2 Protocol Deviations

Significant protocol deviations occurring during the study, either critical or major, that could impact the completeness, accuracy, and/or reliability of the study data or that may affect a subject's rights, safety, or wellbeing will be determined by the Sponsor (which will be completed prior to database lock) and will be presented for all subjects randomized.

8.2 ELIGIBILITY CRITERIA

Males and females, 18 years of age and above at the time of informed consent, who have been diagnosed with WHO Group 3 PH associated with COPD are eligible for randomization to this study. A full list of inclusion and exclusion criteria can be found in the study protocol, Sections 4.1 and 4.2.

8.3 OTHER DESCRIPTIONS OF STUDY POPULATION

All summaries of baseline information will be based on the full analysis set population, unless otherwise noted.

8.3.1 Demographics

Descriptive statistics will be presented overall based on the overall population for the following demographic data collected prior to randomization:

- age (in years);
- age (<65, ≥65 years);
- sex (males, females);
- race (White, Black or African-American, Asian, American Indian or Alaskan Native, Native Hawaiian or Other Pacific Islander);
- ethnicity (Hispanic or Latino, non-Hispanic or Latino);
- baseline (Screening Visit 1 and Day 1) weight (in kg);
- baseline (Screening Visit 1) height (in m);
- baseline (Screening Visit 1 and Day 1) Body mass index (BMI) (in kg/m^2);
- country of clinical site (USA, Israel, Argentina).

8.3.2 Baseline Disease Characteristics

Descriptive statistics will be presented overall based on the randomized population for the following disease characteristics collected prior to randomization:

- age of WHO Group 3 PH-COPD diagnosis (date of diagnosis minus date of birth, in years, by quartiles);
- years of PH-COPD (date for the start of low dose run-in minus date of diagnosis, in years, by quartiles);
- baseline (Day 1) 6MWD;
- baseline (Day 1) Borg dyspnea score;
- baseline (Day 1) 6MWD and Borg dyspnea composite score;
- baseline (Days -7 to -1) actigraphy: average minutes of overall activity and MVPA;
- baseline (Day 1) SGRQ score;
- baseline (Day 1) UCSD SOBQ score;
- baseline (Day 1) plasma concentration of NT-proBNP;
- baseline (Screening Visit 1 and Day 1) FEV_1;
- baseline (Screening Visit 1 and Day 1) FEV_1/FVC;
- lowest recorded SpO_2 during baseline (Screening Visit 1 and Day 1) 6MWT;
- baseline (Day 1) PGA: categorical percentages and average scale value for fatigue and shortness of breath;
- baseline (Day 1) DSP;

- supplemental oxygen use and flow rate (in L/min) with 6MWT at baseline (Screening Visit 1 and Day 1);
- At-home spirometry values at baseline (Day 1).

8.3.3 Medications

Prescription or nonprescription medications (including vitamins and herbal products) will be coded using the most recent version of the WHO-Drug Global Dictionary prior to the data base lock.

Medications will be summarized overall and by assigned treatment - 1 for prior medications and 1 for concomitant medications. The number and percentage of subjects with any medication will be summarized by decreasing order of incidence. The number and percentage of subjects who have taken a medication within a primary drug category will be summarized by cohort under that drug category by the most specific level available.

For these summaries, a subject taking the same medication multiple times will only be counted once for the corresponding most specific level available. Similarly, a subject who took multiple medications within the same drug category will be counted only once for that drug category. Prior and concomitant medications will be presented in alphabetical order of drug category; within each drug category, the most specific level available for drugs will be presented in decreasing order of incidence.

8.3.3.1 Prior Medications

All medications or therapies reported before the first dose of treprostinil (during low dose run-in) are defined as prior medications and will be summarized overall and by assigned treatment.

8.3.3.2 Concomitant Medications

Any medication or therapy initiated after, or corresponding with, the first dose of treprostinil (during low dose run-in) will be considered a concomitant medication. These medications will be presented using the same structure as prior medications. Tabulations will allow occurrences during low dose run-in and blinded treatment to be assessed (eg, low dose treprostinil, treprostinil treated, any treprostinil, placebo treated).

8.3.4 Medical History

Medical history identified prior to ICF date will be summarized overall and by treatment. The number and percentage of subjects with any occurrence will be summarized by System Organ Class (SOC; alphabetically) and by Preferred Term (PT) in decreasing order of incidence.

For subjects experiencing the same medical condition multiple times, the summaries will count the condition only once for the corresponding PT. Similarly, the summaries will count a subject with multiple medical conditions within the same SOC only once for that SOC. SOCs will be tabulated alphabetically; within each SOC, PTs will be presented in decreasing order of incidence.

Medical conditions are coded according to the Medical Dictionary for Regulatory Activities (MedDRA) coding dictionary. The most recently available MedDRA version before the database lock date will be used to assign PTs and SOCs (the database will document the version at time of data lock).

8.3.5 Subject Disposition

All subjects screened will be used to describe subject disposition. The number and percentage of randomized subjects who completed and who discontinued (with reason for discontinuation) will be summarized overall and by treatment sequence at the time of discontinuation. In addition, the summary table will display the number and percentage of subjects in each of the following groups, as well as the reason for discontinuation and treatment assigned, where applicable:

- Screened, but never randomized;
- Randomized, but never treated;
- Discontinued during the first treatment phase;
- Completed the first treatment phase and discontinued prior to the second;
- Completed the first treatment phase and discontinued during the second treatment phase;
- Completed the study (both treatment phases).

A CONSORT diagram (Schulz 2010) may be used to display the disposition by treatment.

9 EFFICACY ANALYSES

The following describes the planned primary endpoint analysis using 6MWD if the alternative primary endpoint analysis is not enacted. For details about the alternate primary endpoint and analysis, see Section 9.2. All efficacy analyses will be performed on the FAS (see Section 8). Additional supportive analyses will be conducted on the PPROT, as outlined below.

9.1 PRIMARY EFFICACY MEASURE

The primary efficacy endpoint is the change in 6MWD measured at peak exposure (defined as 10 to 60 minutes after dosing) after 12 weeks of study drug treatment. The primary endpoint analysis will evaluate the difference in baseline-adjusted 6MWDs between active and placebo conditions. The primary endpoint analysis will test the hypothesis of an improvement in baseline-adjusted 6MWDs for active compared to placebo treatments, as randomized, to subjects with PH-COPD. The primary null hypothesis to be tested is that there are no differences in the baseline-adjusted means change in 6MWDs between active and placebo treatment groups.

9.1.1 Hypothesis Test

The analysis of the primary efficacy endpoint of peak 6MWD assesses if inhaled treprostinil will increase the distance traversed in the peak 6MWT at Week 12 relative to baseline compared to placebo in subjects with PH-COPD. The null and alternative hypotheses are:

$H0: \mu 1 = \mu 2$
$Ha: \mu 1 > \mu 2,$

where μ1 and μ2 are the baseline-adjusted change from baseline in 6MWD of the inhaled treprostinil and placebo treatments, respectively.

9.1.2 Primary Efficacy Analysis Data

The primary efficacy analysis will be based on all available measurements of the 6MWD for the FAS as observed in the study, with the exception that for subjects whose 6MWD measures at Week 6 or Week 12 of either Period are missing due to death or being too ill to perform 6MWT, the missing values will be imputed as the worst possible score (0 meter) and change-from-baseline scores calculated. For missing data due to any other reason, the missing value will not be imputed.

9.1.3 Primary Efficacy Analysis

Descriptive statistics showing the raw values and changes from baseline for the 6MWD will be presented for each scheduled visit by treatment. Additionally, boxplots presenting the raw values and changes from baseline for each visit for the 6MWD may be presented.

A mixed-effect model for repeated measures (MMRM) will be used to estimate the treatment effect. The magnitude of treatment effect will be compared between the 2 treatments across the 12-week treatment periods, and the test of the Week 12 change from baseline and difference between treatments will be produced using contrasts.

[redacted]

[redacted]

Least square means for baseline adjusted 6MWD for each treatment group, difference of LS means between treatment groups with the associated p-value, and 95% CIs for each treatment LS mean and the difference in LS means will be summarized and presented.

9.1.4 Sensitivity Analyses

Sensitivity analyses will be carried out to assess the effect of deviations from the study protocol, the carryover effect, and the impact of missing data. All analyses will be based on the data as described in Section 9.1.2, unless otherwise noted. These analyses are mandatory if the primary endpoint analysis is statistically significant, but will otherwise be performed at Sponsor's discretion.

9.1.4.1 Deviations from Protocol

A "per protocol" analysis will be carried out to assess the effect of substantive deviations from the study protocol. Analyses will be identical to those in Section 9.1 for the primary analysis, except the PPROT analysis population will be used instead of the FAS (ie, the dataset used will remove subjects with substantive deviations). The test will examine the robustness of the total treatment effect to known departures from the protocol. The magnitude of the treatment effect and significance testing will be contrasted against the results from the primary efficacy analysis.

If results are generally similar, then the effect of substantive deviations will be considered minimal.

9.1.4.2 Carryover Effect

The carryover effect analysis will be performed using a 2-sided Wilcoxon Rank Sum Test on the sum of the changes in 6MWD for each treatment period between assigned treatment sequences. To determine the sum of the treatment effects, the change in the 6MWD from Baseline to Week 12 for the active treatment and placebo will be summed for each subject. The test will determine if the total treatment effect is different across assigned treatment sequence groups. If the result is statistically significant, the assumption that the washout period was long enough to rule out the carryover effect, will be examined using the following approach:

[redacted]

9.1.4.3 Multiple Imputation for Missing Data Due to Death and Too Ill to Walk

The primary efficacy analysis will be imputed 6MWD observations missing due to death or being too ill to perform 6MWT as the worst possible score (0 meters) and change scores calculated. An alternative sensitivity analysis will be conducted that imputes missing observations due to death or being too ill to perform 6MWD to be the worst observed change score in all subjects plus a random error component added. The error will be randomly drawn from a normal distribution with a mean of 0 and a variance equal to the residual variance estimated from the mixed model for observed values of change from baseline in 6MWD, adjusted for the same covariates and factors as pre-specified for the primary analysis.

This analysis will be identical to those in Section 9.4 for the primary analysis, except for the use of the alternatively imputed data for death or being too ill to perform 6MWT as described above.

The test will examine the robustness of the total treatment effect to an alternative imputation strategy for handling missing data arising from death or being too ill to walk.

9.1.4.4 [redacted]

9.2 ALTERNATIVE PRIMARY ENDPOINT

At the conclusion of the study, a ***fully blinded*** analysis will be performed to assess the extent of missing primary efficacy endpoint data (ie, 6MWD data). If the absence of 6MWD data in the Treatment Period, combined across subjects, treatment periods, and key visits (Baseline and Study Week 12) to be used for the final primary endpoint analysis would likely lead to an underpowered study or yield indeterminate results, an alternative primary endpoint will be used for the final primary endpoint analysis.

The alternative primary efficacy endpoint is defined as the ranked value for 6MWD measured at each study visit (if present and otherwise imputed for death or too ill to walk) or the ranked value for MVPA measured corresponding to each study visit (otherwise). These data are termed the combined 6MWD/MVPA ranks data.

[redacted]

The following sections present the pre-specified procedures to be followed.

9.2.1 Pre-specified Primary Endpoint Switching Algorithm

At the conclusion of the study, a ***blinded*** analysis will be performed to assess the extent of missing primary efficacy endpoint data. The Sponsor has estimated the power of this planned study at various sample sizes and determined that the minimum acceptable power (90% for a 2-tailed alpha=0.05) occurs at approximately 88 subjects providing evaluable primary endpoint data. With planned sample size of 136 evaluable subjects, if at least 65% of the 6MWD has been

collected (or otherwise imputed for death or too ill to walk) across subjects, periods, and key visits (Baseline and Week 12), the primary endpoint will remain, as planned, the 6MWD, regardless of actigraphy data.

If the study completes under the Original Crossover Design, the outcome of the analysis will inform a decision to adopt an alternative primary efficacy endpoint of combined 6MWD/MVPA ranks data as follows:

- If data presence ≥65% for 6MWD data (Study Weeks 1, 12, 14, and 25), regardless of data presence for actigraphy, the study will use only 6MWD data as the primary efficacy endpoint.
- If data presence <65% for 6MWD data (Study Weeks 1, 12, 14, and 25) and the data presence for actigraphy (Study Weeks 1, 12, 14, and 25) is greater than the presence for the 6MWD data, the study will switch and use combined 6MWD/MVPA ranks data as the primary efficacy endpoint.
- In the event that the data presence <65% for 6MWD data (Study Weeks 1, 12, 14, and 25) and the data presence for actigraphy (Study Weeks 1, 12, 14, and 25) is equal or lower than the presence for the 6MWD data, the study will use only 6MWD data as the primary efficacy endpoint.

If the study completes under the Contingent Parallel Design, the outcome of the analysis will inform a decision to adopt the alternative primary efficacy endpoint of actigraphy data as follows:

- If data presence ≥65% for 6MWD data (Study Weeks 1 and 12), regardless of data presence for actigraphy, the study will use only 6MWD data as the primary efficacy endpoint.
- If data presence <65% for 6MWD data (Study Weeks 1 and 12) and the data presence for actigraphy (Study Weeks 1 and 12) is greater than the presence for the 6MWD data, the study will switch and use combined 6MWD/MVPA ranks data as the primary efficacy endpoint.
- In the event that the data presence <65% for 6MWD data (Study Weeks 1 and 12) and the data presence for actigraphy (Study Weeks 1 and 12) is equal or lower than the presence for the 6MWD data, the study will use only 6MWD data as the primary efficacy endpoint.

This algorithm intends that the 6MWD is to be used as the primary efficacy endpoint for this study, except in the case where onsite measurements for the 6MWD are not available.

9.2.2 Alternative Primary Endpoint Hypothesis Test

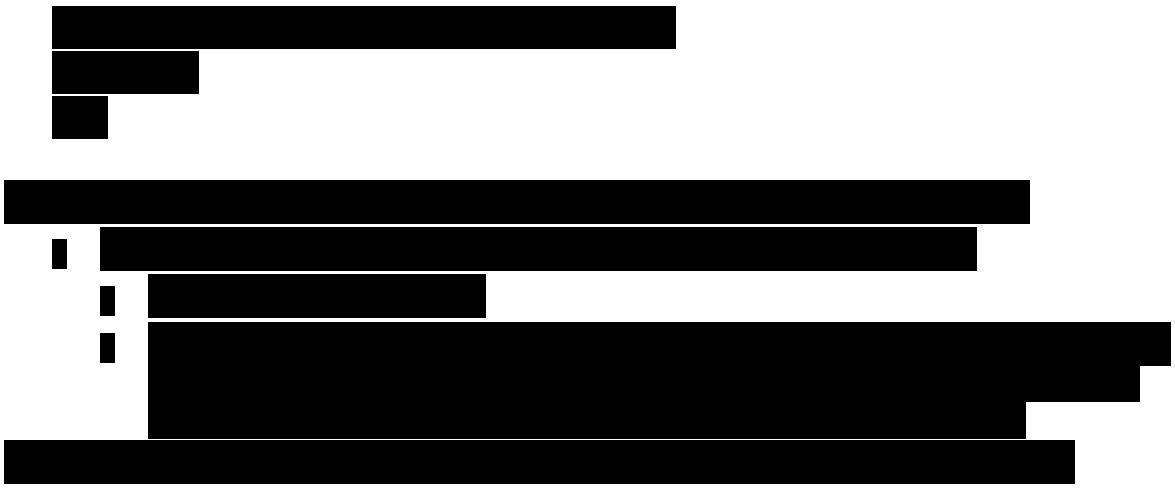
The primary efficacy endpoint of combined 6MWD/MVPA ranks assesses if inhaled treprostinil will increase the combined ranks at Week 12 over placebo in subjects with PH-COPD. The null and alternative hypotheses are:

$H0: \mu 1 = \mu 2$
$Ha: \mu 1 > \mu 2,$

where μ1 and μ2 are the baseline-adjusted change from baseline in combined 6MWD/MVPA ranks of the inhaled treprostinil and placebo treatments, respectively.

9.2.3 Alternative Primary Efficacy Endpoint Derivation

The alternative primary efficacy analysis will be primarily based on all available measurements of the 6MWD for the FAS, as observed in the study. The precursor data file for the alternative primary efficacy data set will be the data defined in Section 9.1.2 (ie, all available measurements of the 6MWD as observed in the study, with values missing due to death or being too ill to perform 6MWT being imputed as the worst possible score of 0 meter). These data will be transformed into ranks and missing 6MWD data that were otherwise measured for MVPA will be imputed for this analysis data. All other values that are missing and not present in MVPA data will not be imputed. The final data resulting will be the combined 6MWD/MVPA ranks data set.



All data sets will be retained for further use in the sensitivity analyses specified below (see Section 9.2.5).

9.2.4 Alternative Primary Efficacy Analysis

The primary efficacy analysis to be conducted will follow those specified for the planned primary efficacy endpoint (see Section 9.1.3) with the exception that specification of the variable for 6MWD (only) will be replaced with the variable created for the combined 6MWD/MVPA ranks.

The Sponsor considers that the sample size is sufficiently large to rely on asymptotic theory. The combined data will be standardized (allowing the 2 separate data sets to be combined on the same scale), and the standardized ranking is a transformation towards the normal distribution. As such, a mixed model would be appropriate and generally expected to perform better under a rank transformation than an analysis based on raw values. The Sponsor believes this is consistent with published literature on rank analysis of covariance (Quade 1967; Conover 1982; Lawson 1983)

9.2.5 Alternative Primary Endpoint Sensitivity Analyses

The sensitivity analyses to be carried out for the alternative primary endpoint (the combined 6MWD/MVPA ranks) will include those specified in Section 9.1.4. Assessment for the effect of deviations from the study protocol, carryover effect, and the impact of missing data will use the combined 6MWD/MVPA ranks data instead of 6MWD data. With the use of combined

6MWD/MVPA ranks data, sensitivity analyses will be performed to assess underlying clinical meaningfulness of the rank scores, and the relationship between 6MWD and actigraphy measurements (overall activity and MVPA). In addition, sensitivity analyses will be performed with 6MWD rank data (without imputation of MVPA data), MVPA raw data (by itself, without 6MWD data), and MVPA rank data (by itself, without 6MWD), applying the primary endpoint imputation approach and method of analysis. All analyses will be based on the data, as described in Section 9.2.3, unless otherwise noted. These analyses specified in this section are mandatory if the use of the alternative primary endpoint is triggered; otherwise will only be performed at Sponsor's discretion.

9.2.5.1 [redacted]

[redacted]

- [redacted]

[redacted]

9.2.5.2 [redacted]

[redacted]

9.3 KEY SECONDARY EFFICACY MEASURES

The following secondary endpoints will be formally tested (in order) if the primary endpoint analysis is statistically significant. Each will assess if inhaled treprostinil improves the endpoint at Week 12 over placebo in subjects with PH-COPD. Improvement is defined as a decrease in the NT-proBNP plasma concentration, an increase in MVPA, an increase in overall activity, decrease in Borg dyspnea score, and an increase in the rank for the 6MWD and Borg dyspnea composite score. The null and alternative hypotheses are:

$H0: \mu 1 = \mu 2$
$Ha: \mu 1 < \mu 2$ or $\mu 1 > \mu 2$, dependent on the stated direction above,

where µ1 and µ2 are inhaled treprostinil and placebo treatments, respectively.

9.3.1 Actigraphy

Subjects wore the actigraph accelerometer CentrePoint Insight Watch (ActiGraph., Pensacola, FL) on their nondominant wrist from the onset of the low dose inhaled treprostinil run-in period until study termination. Subjects were provided an actigraph at the time of the screening run-in visit to establish their baseline level of activity. They were asked to wear the device continuously throughout the day, including while sleeping. Physical activity levels were continuously monitored throughout the blinded treatment period. Measurements were not viewable by subject via the watch or otherwise. [redacted]

9.3.1.1 Overall Activity

Summary epoch data were further processed by ActiGraph to produce daily summary data of activity. Derived from the epoch data were aggregate measures of total epochs measured in each

day, total sleep minutes (epochs with awake flag = FALSE), total awake minutes (epochs with awake flag = TRUE), total non-wear minutes (epochs with wear flag = FALSE), and total wear minutes (epochs with wear flag = TRUE). [redacted]

Physical activities can be characterized into groupings based on their energy expenditures, measured in Metabolic Equivalent of Tasks (METs), as:

Activity Intensity	Example Activities
Sedentary (<100 counts) (<1.5 MET)	• Lying • Standing • Computer work
Light (100-1,951 counts) (1.6-3.0 MET)	• Washing dishes • Washing windows • Vacuuming
Moderate (1,952-5,724 counts) (3.1-6.0 MET)	• Walking • Ascending/ descending stairs • Lawn mowing
Vigorous (>5,724 counts) (>6.0 MET)	• Slow/fast running • Intense sports

Each minute of the day can be converted into an activity intensity, allowing the amount of time in sedentary, light, moderate, and vigorous activities to be determined. Nonsedentary activity is a sum of Light, Moderate, and Vigorous activity. Average counts per day provides a direct measure of overall physical activity. MET = metabolic equivalent of task.

Adapted from Nathan 2020

These activity data were further filtered to include only activity when measured during flagged wear epochs. Summary daily data were available to the Sponsor and investigative site staff throughout the study and a final cumulative transfer of summary data to the Sponsor is to occur after database lock.

The secondary endpoint defined for analysis is the percent change in overall activity (ie, minutes spent in non-sedentary activity) during flagged wake epochs and calculated as absolute

change/baseline (eg, change overall activity [minutes]/baseline overall activity [minutes]). Data will be aggregated across the 7-day baseline to produce an average baseline measurement to minimize day to day variability. Similarly, the Weeks 6 and 12 endpoints will be constructed by averaging across the measurements on the study visit day and the measurements for the 6 days prior.

9.3.1.2 Moderate to Vigorous Physical Activity

The secondary endpoint defined for analysis is the percent change in MVPA (ie, minutes spent in moderate or vigorous activity) during flagged wake epochs and calculated as absolute change/baseline (eg, change MVPA [minutes]/baseline MVPA [minutes]). Data will be aggregated across the 7-day baseline to produce an average baseline measurement to minimize diurnal variability. Similarly, the Weeks 6 and 12 endpoints will be constructed by averaging across the study visit day measurement and the measurements for the 6 days prior.

9.3.1.3 Actigraphy Analyses

9.3.1.3.1 Main Analyses

A longitudinal mixed model, similar to that used for the primary analysis of 6MWD, will be used to describe the change from baseline in actigraphy, percent change in overall activity and MVPA, separately, over time. All subjects, including those who did not complete both treatment sequences, will be included in the model. A contrast will be calculated for Week 12 to compare the treatment groups with respect to percent change. The outcome variable will be percent change from baseline. Analyses will be conducted on the FAS and the PPROT.

9.3.1.3.2 [redacted]

9.3.1.3.3 [redacted]

[redacted]

9.3.2 Borg Dyspnea Score

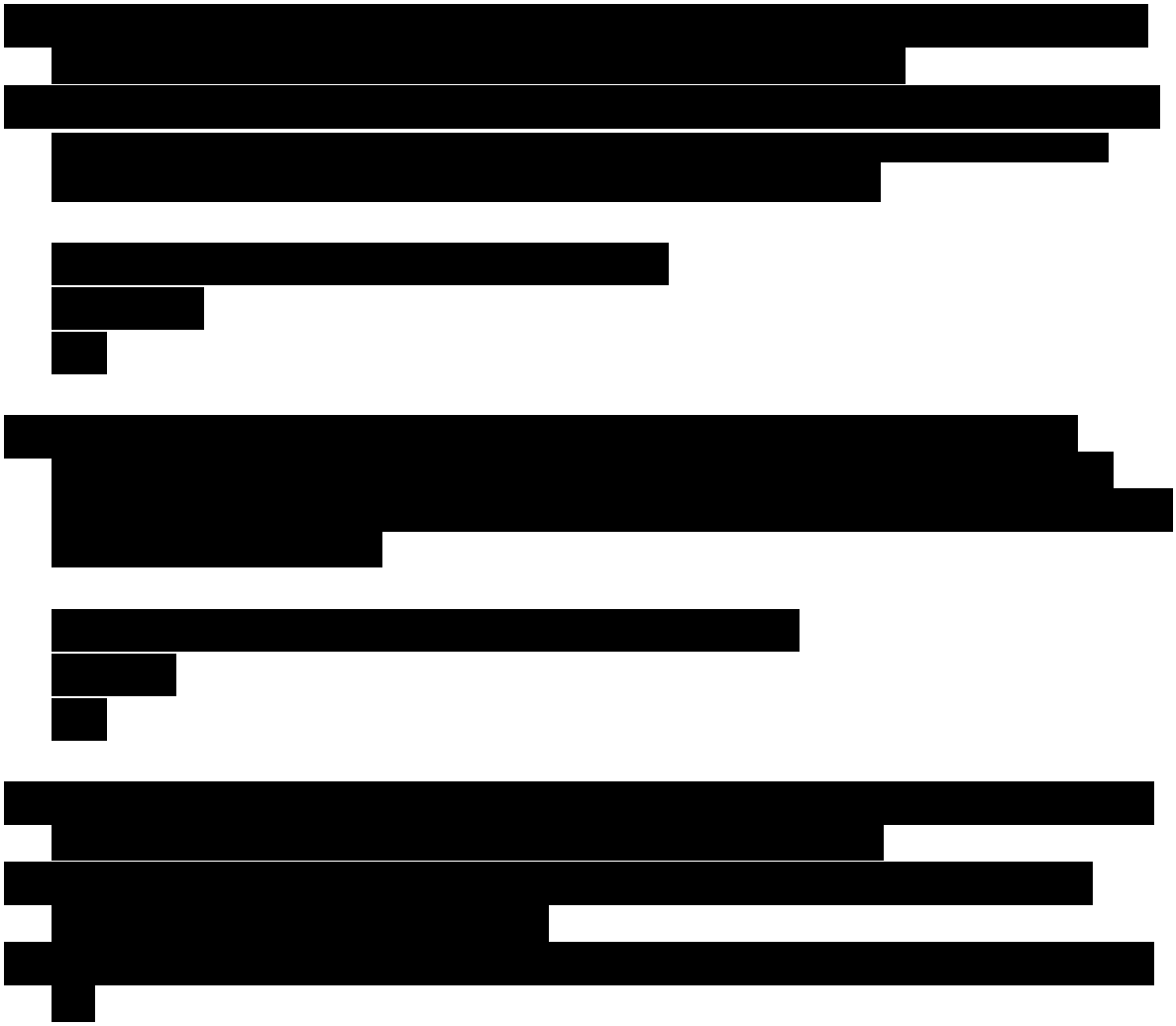
The Borg dyspnea score is an 11-point scale measuring the degree of dyspnea from 0 (no shortness of breath) to 10 (maximal shortness of breath). A score of 10 will be assigned to subjects who died or were too ill to walk for the 6MWT. Descriptive statistics of the raw and change from baseline values for Borg dyspnea score will be presented for each scheduled visit by treatment. Additionally, boxplots presenting the raw values and changes from baseline for each visit will be presented. The Borg dyspnea score will be treated as a continuous variable.

A longitudinal mixed model, similar to that used for the primary analysis of 6MWT, will be used to describe the change from baseline in Borg dyspnea score over time. All subjects, including those who did not complete both treatment sequences, will be included in the model. A contrast will be calculated for Week 12 to compare the treatment groups with respect to change. The outcome variable will be change from baseline. Independent variables will be treatment, the subject's baseline Borg dyspnea score, treatment sequence, visit number, and the interaction of visit and treatment. Analyses will be conducted on the FAS and the PPROT.

9.3.3 6MWD and Borg Dyspnea Composite Score

The intent of the 6MWT is to determine how much exercise subjects can do during the course of carrying out activities of daily living. However, the capacity of subjects to function is determined not only by what they can do when they exert themselves to the fullest, but also by how they feel when they are carrying out their usual activities of daily living. It is important not only to look at the distance traversed during the unencouraged 6MWT, but also the symptoms experienced at the end of the effort (Rao 2019).

The 6MWD and Borg dyspnea composite score is derived based on the results of the 6MWD and the Borg dyspnea scale measuring the degree of dyspnea. The following derivation strategy will be employed to calculate the 6MWD and Borg dyspnea composite score:



Descriptive statistics of the raw and change from baseline values for composite score will be presented for each scheduled visit by treatment. The composite score will be treated as a continuous variable. A longitudinal mixed model, similar to that used for the primary analysis of 6MWD, will be used to describe the change from baseline in the composite score over time. All subjects, including those who did not complete both treatment sequences, will be included in the model. A contrast will be calculated for Week 12 to compare the treatment groups with respect to change.

9.3.4 NT-proBNP

Descriptive statistics of the raw values for NT-proBNP will be presented for each scheduled visit by treatment. Because the parameter is highly skewed, the data will be log-transformed before analysis. A mixed model will be used to describe log NT-proBNP over time. The outcome variable will be the value at each time of measurement, not the change from baseline. Independent variables will be treatment, treatment sequence, the subject's baseline log NT-proBNP, visit number, and the interaction of visit and treatment. A contrast will be calculated at Week 12 to compare the treatment with respect to level of the parameter.

The contrast will be on the log scale. In order to provide summary statistics that are easily clinically interpretable, the estimates may be back transformed to the natural scale and the change from baseline may be calculated as the difference between the mean of the back-transformed estimates and the mean of the baseline value.

Observed values for the FAS and PPROT will be used with no imputation for missing values. However, if a NT-proBNP value is reported below the lower limit of quantification, half of the lower limit of quantification will be used for summary and analysis (eg, if the result is reported as '<51 ng/L,' a numerical value of 25 ng/L will be used).

9.4 OTHER SECONDARY EFFICACY MEASURES

9.4.1 St. George's Respiratory Questionnaire

The SGRQ is a 50-item questionnaire designed to measure impact on overall health, daily life, and perceived wellbeing in patients with obstructive airways disease (Jones 1992). Scores for the SGRQ are totaled into 3 components (symptoms, activities, impacts) as well as a total score. Scores weighted and expressed as a range from 0 to 100, with higher scores indicating more limitations. The current established minimal important difference for SGRQ scores in an average COPD population is −4 units (Jorrit 2015). Calculation of component and total scores are described in the manual (Jones 2009).

Descriptive statistics for the raw and change from baseline to Week 12 values for the SGRQ the components and the total score will be presented for each scheduled visit by treatment.

Additionally, boxplots presenting the raw values and changes from baseline for each visit will be presented. The SGRQ scores will be treated as a continuous variable.

A longitudinal mixed model will be used to describe the change from baseline for the each of the scores over time. All subjects, including those who did not complete both treatment sequences, will be included in the model. A contrast will be calculated for Week 12 to compare the treatment groups with respect to change. The outcome variable will be change from baseline. Independent variables will be treatment, the subject's baseline average SGRQ score, treatment sequence, visit number, and the interaction of visit and treatment.

Observed values for the FAS and PPROT will be used with no imputation for missing values.

9.4.2 University of California San Diego Shortness of Breath Questionnaire

The UCSD SOBQ is a self-administered rating of dyspnea associated with activities of daily living (Eakin 1998). The questionnaire contains 24 items on a 6-point scale, where 0 = "not at all" and 5 = "maximal or unable to do because of breathlessness" and total score are summed across items for a range from 0 to 120. The current established minimal important difference for UCSD SOBQ scores in a chronic lung disease population is 5 units (Ries 2005). Descriptive statistics of the raw and change from baseline to Week 12 values for the UCSD SOBQ score will be presented for each scheduled visit by treatment.

A longitudinal mixed model will be used to describe the change from baseline in the total score over time. All subjects, including those who did not complete both treatment sequences, will be included in the model. A contrast will be calculated for Week 12 to compare the treatment groups with respect to change. The outcome variable will be change from baseline. Independent variables will be treatment, the subject's baseline average UCSD SOBQ score, treatment sequence, visit number, and the interaction of visit and treatment.

Observed values for the FAS and PPROT will be used with no imputation for missing values.

9.4.3 [redacted]

[redacted]

[redacted]

9.4.4 Patient Global Assessment

The PGA is a being used to rate subject fatigue and shortness of breath. Subjects will use the Sponsor-provided smart device for at-home capture of the PGA data at or just prior to each study visit. Subjects will be asked to perform the PGA at the same time as the at-home spirometry assessment at all applicable visits. Scores for each are on a 5-point response scale of: "never," "rarely," "sometimes," "often," or "always" with higher scores indicating a worse symptom rating.

Descriptive statistics for the raw and change from baseline values for the PGA ratings will be presented for each scheduled visit by treatment. Additionally, boxplots presenting the raw values and changes from baseline for each visit may be presented.

Exploratory analyses may be conducted for PGA items to examine floor and ceiling effects, reliability across repeated baseline assessments, and within subgroups to further understand the psychometric properties of PGA measurements.

9.5 EXPLORATORY EFFICACY MEASURES

Exploratory analyses will not be addressed in this SAP. Relevant analyses conducted will be presented in the CSR.

10 SAFETY ANALYSES

The safety of inhaled treprostinil compared to placebo will be evaluated on the following parameters: AEs, vital sign measurements, ECGs, clinical laboratory assessments, PFTs, spirometry, and oxygenation. All safety analyses will be based on the Safety Population (see Section 6). No formal inferential testing will be reported for routine safety analyses.

10.1 EXTENT OF EXPOSURE

Exposure to treprostinil and placebo will be calculated as the date of last dose minus the date of first dose plus 1. If the date of last dose of study drug was missing or unknown, the return date of study drug will be used as an imputed date. All analyses will be presented by treatment and period. These analyses include:

- Duration of exposure (in weeks) to each treatment
- Number of subjects exposed by time (weekly from Baseline Visit respective to treatment)
- Maintenance dose (in mcg QID) during the study (dose subject is maintained on after titration period [ie, the mode of the dose during that treatment period]).

10.2 ADVERSE EVENTS

Adverse events are recorded starting from the time each subject signs the ICF until the time a screen failure is documented, study completion, or early termination. All events recorded by the Investigators will be assigned a MedDRA PT and SOC. The most recently available MedDRA version before the database lock date will be used to assign preferred terms and system organ classes. All presentations will present data by treatment received at the time of the event. Adverse events that occur during the Washout Period following Treatment Period 1 will be presented as having occurred on the treatment taken by the subject during Treatment Period 1. Similarly, AEs that occur during post-discontinuation follow-up will be presented as having occurred on the treatment taken by the subject during Treatment Period 2. Finally, AE tabulations will allow occurrences during low dose run-in and blinded treatment to be assessed (eg, low dose treprostinil, treprostinil treated, any treprostinil, placebo treated). Non-treatment emergent AEs observed outside of any study drug exposure will not be presented.

In addition, a limited number of presentations may be tabulated for AE data by both treatments received at the time of the event and treatment sequence, including the total rates by sequence.

10.2.1 Severity of Adverse Event

The severity of AEs will be reported by the Investigator as ‘Mild,’ ‘Moderate,’ or ‘Severe.’

10.2.2 Relationship to Study Treatment

The relationship of the event to the study drug is classified by the Investigator as 'Not Related,' 'Possible,' or 'Probable.' For analytic purposes, AEs classified as 'Probable' or 'Possible' will be considered 'Related.'

10.2.3 Action Taken with Study Treatment

For each AE, the Investigator classifies the action taken with the study treatment as 1 of the following: 'Dose Not Changed,' 'Dose Increased,' 'Dose Decreased,' 'Drug Interrupted,' 'Drug Withdrawn,' 'Unknown,' or 'Not Applicable.'

10.2.4 Outcome of Treatment Emergent Adverse Event

The Investigator categorizes the outcome of each AE as 1 of the following: 'Fatal,' 'Not Recovered/Not Resolved,' 'Recovered/Resolved,' 'Recovered/Resolved with Sequelae,' 'Recovering/Resolving,' or 'Unknown.'

10.2.5 Summary of Treatment Emergent Adverse Events

An overall summary of AEs by treatment received at the time of the event and treatment will contain the following:

- Total number of AEs, total number of subjects with at least one AE
- Total number of related AEs, total number of subjects with at least one related AE
- Total number of serious adverse events (SAEs), total number of subjects with at least one SAE
- Total number of AEs leading to study withdrawal
- Total number of AEs leading to treatment discontinuation.

Table summaries of all AEs will present numbers and percentages by SOC and PT of the following by treatment received at the time of the event:

- Subjects with at least one AE
- Subjects with at least one treatment-related AE
- Subjects with at least one SAE
- Subjects with an AE leading to treatment discontinuation.

AEs will also be summarized by treatment received at the time of the event for the following:

- AEs by severity
- AEs by relationship to treatment
- AEs by action taken with treatment, separately
- Listing of all SAEs.

Finally, a table summarizing the AEs that occurred during the pre-randomization run-in phase will be presented for all screened subjects.

Adverse events will be presented in alphabetical order of SOC and by decreasing order of incidence of the PTs.

For these summaries, a subject having the same AE multiple times will only be counted once for the corresponding PT. Similarly, a subject who experienced multiple AEs within the same SOC will be counted only once for that SOC.

A subject experiencing the same AE multiple times or at multiple severities will only be counted once under that PT at the worst severity rating. Similarly, a subject who experiences the same AE multiple times or at multiple severities across multiple PTs within the same SOC will only be counted once for the worst severity rating within the SOC.

10.3 DEATHS

A listing will be provided for any deaths occurring on the study. Included in this listing, to the extent available, will be the subject's age, race, sex, treatment received at the time of death, duration of that treatment in days, whether or not the subject received the other treatment, days since randomization, days since last treatment, and the cause of death.

10.4 PHYSICAL EXAM

Any value the Investigator judges to be a clinically significant abnormal change after the subject's consent to the study was to be recorded as an AE.

10.5 CLINICAL LABORATORY EVALUATIONS

Laboratory tests for serum electrolytes, chemistry, and hematology will be assessed at each study visit. Table 10-1 provides the values that will be summarized.

Table 10-1 Clinical Laboratory Parameters

Electrolyte Panel	Chemistry Panel	Hematology Panel
Sodium	Total bilirubin	Hemoglobin
Potassium	Alkaline phosphatase	Hematocrit
Bicarbonate	Alanine aminotransferase	Red blood cell count
Chloride	Aspartate aminotransferase	White blood cell count
	Urea nitrogen	Platelet count
	Creatinine	Erythrocytes Mean Corpuscular Volume
	Calcium	
	Albumin	

Descriptive statistics for each clinical laboratory parameters (hematology, serum chemistry and electrolytes) will be presented by treatment group at baseline, each post-baseline assessment, and End of Study visits for raw data and change from respective baseline. Additionally, summary statistics will be presented for the change in each parameter from the baseline visit for Treatment Period 1 to the Baseline Visit for Treatment Period 2.

Post-baseline abnormal values, as defined by the normal limits, for hematology, serum chemistry, and electrolyte parameters will be summarized as shift tables from baseline to each post-baseline assessment and End of Study visits.

Tables will contain the number and percentage of subjects by treatment received at the time of measurement with post-baseline categories (low, high, and normal for hematology, chemistry, and electrolytes) compared to the respective baseline category for each parameter.

If data from both the central lab and a local lab is available for a given lab result, then the central lab value will be used in the analyses. If data is only available from a local lab for a given lab result, then local lab results will be transformed to be consistent to the central lab results for the analyses.

If a value for a laboratory value is reported below the lower limit of quantification, half of the lower limit of quantification will be used for summary and analysis.

10.6 VITAL SIGNS

Descriptive statistics for vital signs (heart rate, systolic blood pressure, diastolic blood pressure, respiration rate, weight, BMI) will be presented for both Baseline Visits by treatment. Additionally, summary statistics will be presented for the change in each parameter from the Baseline Visit for Treatment Period 1 to the Baseline Visit for Treatment Period 2. The same statistics will also be presented for each post-baseline assessment and End of Study visits for the raw data and change from the respective baseline. Any clinically significant change from Screening Visit 1 was to be recorded as an AE.

10.7 ELECTROCARDIOGRAM EVALUATIONS

Descriptive statistics for ECG findings (sinus rhythm [present/absent], heart rate, PR interval, QRS duration, QT interval, interpretation [normal/abnormal]) will be presented by treatment group. Continuous variables will also be presented as the raw values and change from the respective baseline will be presented for each post-baseline assessment by treatment received at the time of the visit. Additionally, summary statistics for continuous variables will be presented for the change in each parameter from the Baseline Visit for Treatment Period 1 to the Baseline Visit for Treatment Period 2. Whenever more than 1 ECG evaluations are recorded on the same day, the data summaries will use the average of these readings.

In addition, a summary of subjects with treatment-emergent abnormal values will be presented by treatment received at the time of the evaluation. Presentation will be made in order of decreasing incidence by type of abnormality with the number and percentage of subjects who experienced the type of abnormality.

10.8 PULMONARY FUNCTION TESTS

Descriptive statistics will be presented by treatment group at baseline for the following parameters (absolute values and % predicted): FEV_1, FVC, TLC, and DLCO uncorrected for hemoglobin and lung volume. Additionally, descriptive statistics of the raw values and change

from the respective baseline will be presented for each post-baseline assessment by treatment received at the time of the visit.

10.9 [redacted]

[redacted]

10.10 OXYGENATION

10.10.1 Pulse Oximetry

Pulse oximetry measurements will be taken before, during, and after each 6MWT. These measurements include the collection of saturation peripheral capillary oxygenation (SpO_2) and heart rate (HR). Descriptive statistics of raw values and the change from baseline pre-walk, post-walk, and the lowest recorded SpO_2 (and associated HR) for each 6MWT will be presented by treatment received during time of assessment. The baseline value used to compute the change score of each follow-up measurement will be respective of the same timeframe of the follow-up value (ie, pre-walk baseline used for pre-walk follow-up).

11 REFERENCES

Choi L, Liu Z, Matthews CE, et al. Validation of Accelerometer Wear and Nonwear Time Classification Algorithm. Med Sci Sports Exerc. 2011;Feb;43(2):357-364.

Eakin EG, Resnikoff PM, Prewitt LM, et al. Validation of a new dyspnea measure. The UCSD shortness of breath questionnaire. Chest. 1998;113(3):619-24.

East 6 (2020). Statistical software for the design, simulation and monitoring clinical trials. Cytel Inc., Cambridge MA.

Conover WJ, Iman RL. Analysis of covariance using the rank transformation. Biometrics. 1982;38:715-724.

Jones PW, Quirk FH, Baveystock CM, et al. A self-complete measure for chronic airflow limitation - the St George's Respiratory Questionnaire. Am Rev Respir Dis. 1992;145:1321-1327.

Jones PW, Ford Y. St.George's Respiratory Questionnaire Manual. 2009.

Jorrit BA, Welling JE, Hartman NHT, et al. The minimal important difference for the St George's Respiratory Questionnaire in patients with severe COPD. Eur Respir J. 2015;46:1598-1604.

Kenward MG, Roger, JH. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics. 1997;53(3):983-997.

Lawson A. Rank Analysis of Covariance: Alternative Approaches. The Statistician. 1983;32(3):331-337.

Nathan SD, Flaherty KR, Glassberg MK, et al. A Randomized, Double-Blind, Placebo-Controlled Study of Pulsed, Inhaled Nitric Oxide in Subjects at Risk of Pulmonary Hypertension Associated With Pulmonary Fibrosis. Chest. 2020;158(2):637-645.

Quade D. Rank analysis of covariance. Journal of the American Statistical Association. 1967;62(320):1187-1200.

Rao Y, Zheng C, Nelsen A, et al. Borg Dyspnea Score Adjusted Six Minute Walk Distance in Pulmonary Arterial Hypertension Clinical Trials. PVRI 2019 abstract presented in Barcelona, Spain. 2019.

Ries AL. Minimally clinically important difference for the UCSD Shortness of Breath Questionnaire, Borg Scale, and Visual Analog Scale. COPD: Journal of Chronic Obstructive Pulmonary Disease. 2005;2:105-110.

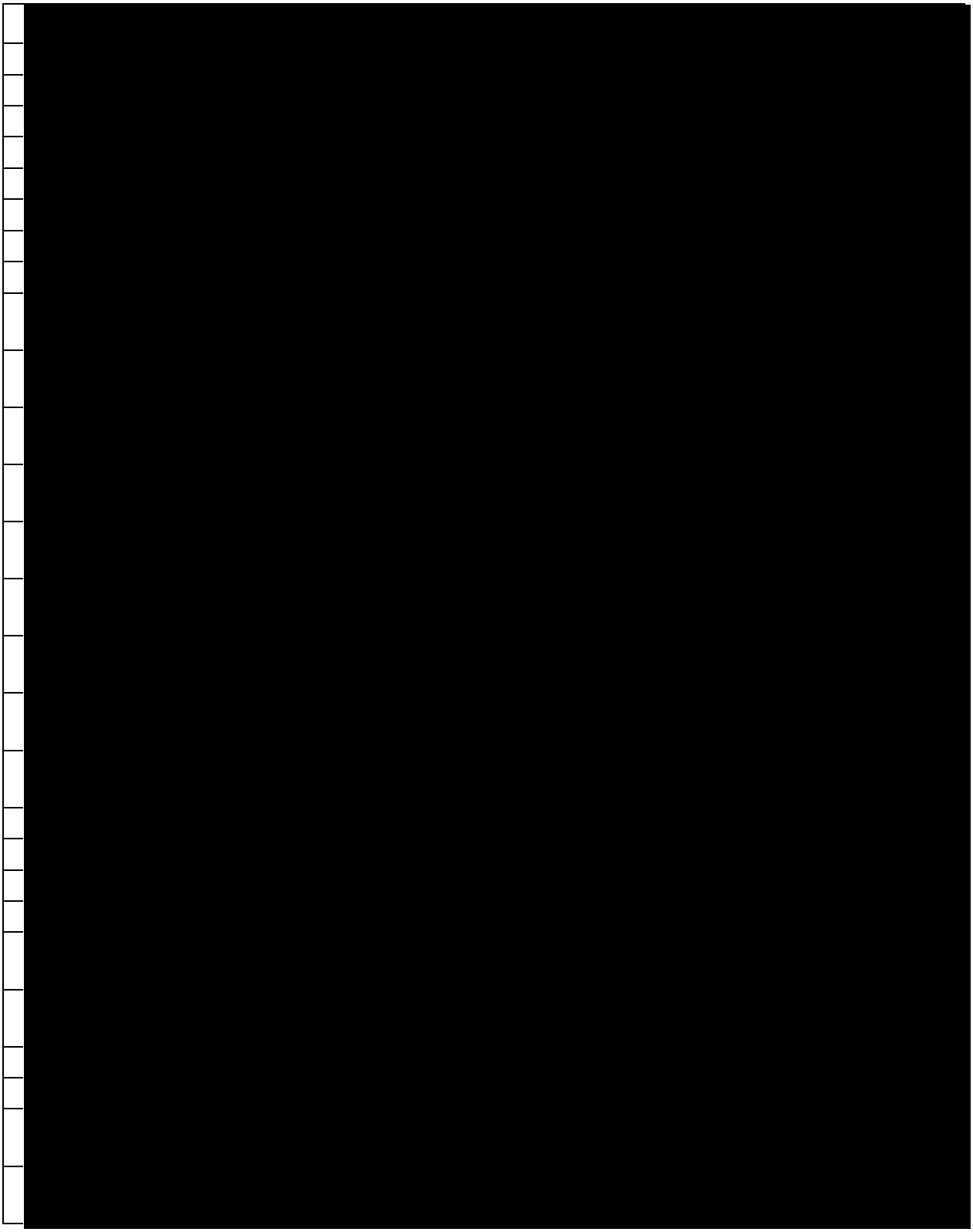
Schulz KF, Altman DG, D. Moher. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMC Medicine. 2010;8:18.

Staudenmayer J, He S, Hickey A, et al. Methods to estimate aspects of physical activity and sedentary behavior from high-frequency wrist accelerometer measurements. J Appl Physiol. 2015;119:396-403.

Tracy JD, Acra S, Chen KY, et al. Identifying bedrest using 24-h waist or wrist accelerometry in adults. PLoS ONE. 2018;13(3):e0194461. https://doi.org/10.1371/journal.pone.0194461.

12 APPENDICES

12.1 LIST OF TABLES



12.2 LIST OF LISTINGS

Listings to be provided upon request.

12.3 LIST OF FIGURES

12.4 CONTINGENT PARALLEL DESIGN ANALYSIS PLAN

The RIN-PH-304 study is an adaptive study with a (fully blinded) decision for design modification based on enrollment and missing data metrics. Following a formal and documented decision to switch from the Original Crossover Design to the Contingent Parallel Design, the following modifications will be applied to those sections specified in the main body of this SAP with the Original Crossover Design. To minimize repetition, this section will cover only those sections affected by a decision to switch to the Contingent Parallel Design, that is:

- If no modifications are required to a section in the main body of the SAP, no mention will be made below. For example, Section 2 (Study Objectives and Endpoints) requires no modification under switch to a parallel design.
- If a modification will be necessary, the affected section number and the replacement text are provided below. The updated text will then supersede the current text in the main body of the SAP. Text within an affected section that does not require modification will not be repeated below.

12.4.1 Study Design

Section 3.1: Approximately 314 eligible subjects will be randomized to either 12 weeks of active study drug or 12 weeks of placebo (Treatment Period 1 only). All subjects following the adaptive switch will be discharged from the study following 12 weeks of treatment. The second washout (Washout Period - Washout 2) and Treatment Period 2 assessments will not be conducted following the adaptive switch and any previously collected Period 2 data will be ignored for the primary analyses.

12.4.2 Interim Efficacy Analysis

Section 4.2: A single interim efficacy analysis will be conducted when 75% of the subjects (ie, ~235 subjects) have completed the 12-week Treatment Period if both of the following criteria are met: a) the expected time to randomize the final (314th) subject is greater than 6 months; b) at least 90% of the 6MWD data is present for Period 1: Baseline and Week 12 for these subjects.

Section 4.2.1: Prior to any planned interim analysis being performed, the following steps will be followed to estimate the expected time to randomize the final (314th) subject:

[redacted]

Section 4.2.2: If conducted with exactly 75% information fraction, the interim analysis will apply a 2-sided alpha level of 0.019, and at the final analysis at a 2-sided alpha of 0.044 with the overall Type I error rate at 2-sided alpha of 0.05; the actual information fraction at the time of the interim analysis will be used to calculate the 2-sided alpha level that will be applied. Efficacy boundaries for early stopping of the study for efficacy are based on an O'Brien-Fleming spending function (PASS 12, NCSS, Version 12.0). There will be no corresponding futility boundary for the interim analysis.

12.4.3 General Considerations For Data Analyses

Section 7: If the study design is changed to a parallel design, efficacy and safety summary tables will follow the general format that baseline and post-baseline data will be presented by the assigned treatment group when those data are observed. Treatments will be labeled as "Treprostinil" and "Placebo" to correspond with the assigned treatment group.

Efficacy and safety data collected after study discharge (after Period 1) will not be presented.

Section 7.1: If the study design is changed to a parallel design for this study, only the Period 1 baseline for efficacy measurements will be used. Except for actigraphy measurements, the baseline value will be defined as the non-missing value immediately prior to dosing with study drug for Period 1: Day 1. The baseline period for actigraphy measurements will be the last 7 days during the washout period prior to the day of dosing on Day 1. In the event that the washout period is less than 7 days then only the non-dosing days will be used to derive a baseline average.

12.4.4 Covariates

Section 7.4: Efficacy analyses will not include the assigned treatment sequence and period order as covariates. Change from baseline analyses will include the baseline measurement as a continuous covariate.

12.4.5 Primary Efficacy Measure

Section 9.1.3: Descriptive statistics showing the raw values and changes from baseline for the 6MWT will be presented for each scheduled visit by treatment group.

A longitudinal model with change from baseline as the outcome variable, implemented through SAS PROC MIXED, will be used to analyze the data. The model will include all measurements of 6MWT. Baseline average 6MWT will be a covariate. The model will include terms for the interaction between visit number and treatment. The contrast between treatment groups at Week 12 will be calculated from the following model:

[redacted]

LS means for baseline-adjusted 6MWT for each treatment group, LS means of treatment group differences with their associated p-values, and 95% CIs for both will be summarized and presented.

The primary sensitivity analysis will incorporate control-based multiple imputation, which assumes that patients who stop study medication will adopt the experience in the placebo group,

will be used to account for missing data. An addendum to this protocol, to be written before the data are unblinded, will present the SAS program to be used for the multiple imputation.

Section 9.1.4: Sensitivity analyses will explore the extent to which multiple imputation of Week 12 data influences the interpretation of the outcome.

The following describe the sensitivity analysis approach using multiple imputation of the primary outcomes. The analyses will be structured as follows:

Planned sensitivity analyses will be described in an addendum to this SAP, which will be completed around the Blind Data Review but prior to the breaking of the study blind.

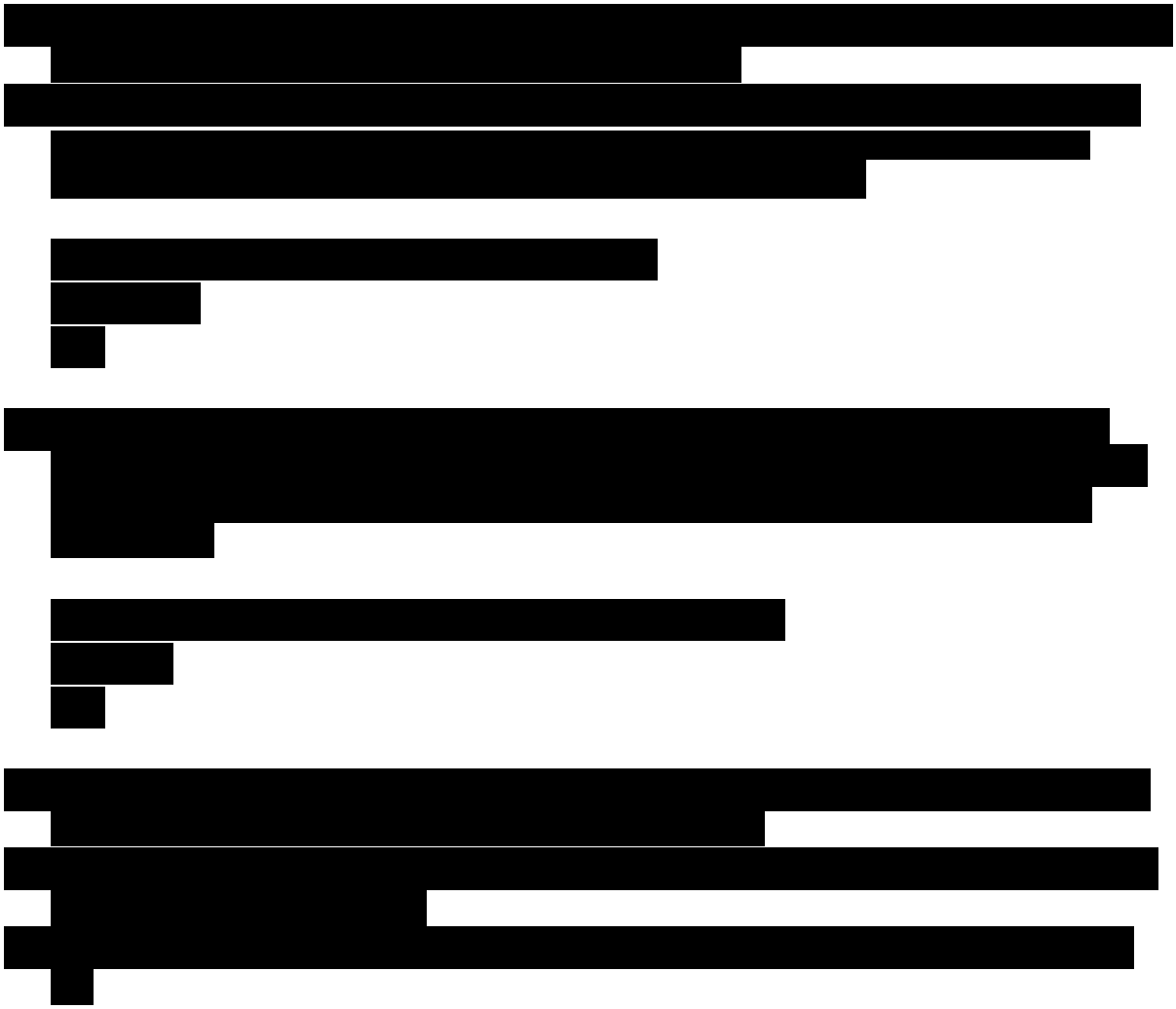
Section 9.2.1: At the conclusion of the study, a ***blinded*** analysis will be performed to assess the extent of missing primary efficacy endpoint data. The Sponsor has estimated the power of this planned study at various sample sizes and determined that the minimum acceptable power (80% for a 2-tailed alpha=0.05) occurs at approximately 200 subjects (~100 per treatment) providing evaluable primary endpoint data. With a planned sample size of 314 evaluable subjects, if at least 65% of the 6MWD has been collected (or otherwise imputed for death or too ill to walk) across subjects and key visits (Baseline and Week 12), the primary endpoint will remain, as planned, the 6MWD, regardless of actigraphy data.

Section 9.2.3:

The final data resulting will be the combined 6MWD/MVPA ranks data set.

Section 9.3.1.3.2: The associations between the observed change from baseline actigraphy measurements of MVPA and overall activity will be contrasted with the magnitude of change for PGA items. Assessments will be made in an exploratory fashion and may include bivariate frequencies and correlations to assess the associations between activity and symptoms measurements made overall and between treatments (as appropriate). A 1-category change on the 5-point PGA scale will be considered clinically meaningful.

Section 9.3.3: The 6MWD and Borg dyspnea composite score is derived based on the results of the 6MWD and the Borg dyspnea scale measuring the degree of dyspnea. The following derivation strategy will be employed to calculate the 6MWD and Borg dyspnea composite score:



Descriptive statistics of the raw and change from baseline values for composite score will be presented for each scheduled visit by treatment. The composite score will be treated as a continuous variable. A longitudinal mixed model, similar to that used for the primary analysis of 6MWD, will be used to describe the change from baseline in the composite score over time. A contrast will be calculated for Week 12 to compare the treatment groups with respect to change.

12.4.6 Secondary Efficacy Measures

Section 9.4.1: A longitudinal mixed model will be used to describe the change from baseline for the each of the scores over time. A contrast will be calculated for Week 12 to compare the treatment groups with respect to change. The outcome variable will be change from baseline. Independent variables will be treatment, the subject's baseline average SGRQ score, treatment sequence, visit number, and the interaction of visit and treatment.

Section 9.4.2: A longitudinal mixed model will be used to describe the change from baseline in the total score over time. A contrast will be calculated for Week 12 to compare the treatment groups with respect to change. The outcome variable will be change from baseline. Independent variables will be treatment, the subject's baseline average UCSD SOBQ score, treatment sequence, visit number, and the interaction of visit and treatment.

12.4.7 Safety Analyses

Section 10.1: Exposure to treprostinil and placebo will be calculated as the date of last dose minus the date of first dose plus 1. If the date of last dose of study drug was missing or unknown, the return date of study drug will be used as an imputed date. All analyses will be presented by treatment.

Section 10.2: Adverse events are recorded starting from the time each subject signs the ICF until the time a screen failure is documented, study completion, or early termination. All events recorded by the Investigators will be assigned a MedDRA PT and SOC. The most recently available MedDRA version before the database lock date will be used to assign preferred terms and system organ classes. All presentations will present data by treatment received at the time of the event. Non–treatment-emergent AEs observed outside of any study drug exposure will not be presented.

Section 10.5: Descriptive statistics for each clinical laboratory parameter (hematology, serum chemistry, and electrolytes) will be presented by treatment group at baseline, each post-baseline assessment, and End of Study visits for raw data and change from baseline.

Section 10.6: Descriptive statistics for vital signs (heart rate, systolic blood pressure, diastolic blood pressure, respiration rate, weight, BMI) will be presented for both Baseline Visits by treatment. The same statistics will also be presented for each post-baseline assessment and End of Study visits for the raw data and change from the respective baseline. Any clinically significant change from Screening Visit 1 was to be recorded as an AE.

Section 10.7: Descriptive statistics for ECG findings (sinus rhythm [present/absent], heart rate, PR interval, QRS duration, QT interval, interpretation [normal/abnormal]) will be presented by treatment group. Continuous variables will also be presented as the raw values and change from the respective baseline will be presented for each post-baseline assessment by treatment received at the time of the visit. Whenever more than one ECG evaluations are recorded on the same day, the data summaries will use the average of these readings.