Document Type:	Statistical Analysis Plan
Official Title:	A prospective, randomized, international, multicenter, double-arm, controlled, open label study of Riociguat in patients with pulmonary arterial hypertension (PAH) who are on a stable dose of phosphodiesterase-5 inhibitors (PDE-5i) with or without endothelin receptor antagonist (ERA), but not at treatment goal
NCT Number:	NCT02891850
Document Date:	09 April 2020

A prospective, randomized, international, multicenter, double-arm, controlled, open label study of Riociguat in patients with pulmonary arterial hypertension (PAH) who are on a stable dose of phosphodiesterase-5 inhibitors (PDE-5i) with or without endothelin receptor antagonist (ERA), but not at treatment goal.

Riociguat rEplacing PDE-5i therapy evaLuated Against Continued PDE-5i thErapy

BSP study drug	BAY 63-2521/Adempas/Riociguat		
Study purpose:	To demonstrate the effectiveness of riociguat as replacement of PDE-5i therapy in PAH patients		
Clinical study phase:	IV	**Date:**	09 April 2020
Study No.:	IMPACT no. 18588	**Version:**	Final 3.0
Author:	PPD		

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

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

Statistical Analysis Plan (Amendment/Supplement) Approval Form

Study Number (Bay No./IMP no.)*	BAY 63-2521 / IMPACT no. 18588
Statistical Analysis Plan (SAP) Version and Date	Final v3.0 09-April-2020

* if no IMPACT number is available, refer to the approved Study Concept

I have read and approve the SAP / SAP Amendment referred above.

	Name	Signature	Date
Author:			
Study Statistician	PPD		
Approved by:			
Project Statistician	PPD	PPD	PPD Datum: 2020.04.09 14:37:05 +02'00'
Study Medical Expert	PPD		
Medical Writer	PPD		

Statistical Analysis Plan (Amendment/Supplement) Approval Form

Study Number (Bay No./IMP no.)*	BAY 63-2521 / IMPACT no. 18588
Statistical Analysis Plan (SAP) Version and Date	Final v3.0 09-April-2020

* if no IMPACT number is available, refer to the approved Study Concept

I have read and approve the SAP / SAP Amendment referred above.

	Name	Signature	Date
Author:			
Study Statistician	PPD	PPD	09-Apr-2020
Approved by:			
Project Statistician	PPD		
Study Medical Expert	PPD	PPD	PPD Datum: 2020.04.09 14:31:08 +02'00'
Medical Writer	PPD	PPD	PPD PPD

Table of Contents

Abbreviations

6MWD	6 minute walking distance
AE	adverse event
ALT	alanine aminotransferase
AST	aspartate aminotransferase
AUC	area under the curve
BP	blood pressure
cGMP	cyclic guanosine monophosphate
cMRI	cardiac magnetic resonance imaging
CRO	contract research organization
CTEPH	chronic thromboembolic pulmonary hypertension
ECG	electrocardiogram
eCRF	electronic case report form
EOS	End of Study
ERA	endothelin receptor antagonist
EU	European Union
FEV1/FVC	forced expiratory volume in one second/forced vital capacity
FAS	full analysis set
FU	follow-up
GCP	Good Clinical Practice
IB	Investigator's Brochure
ICF	Informed consent form
IEC	independent ethics committee
IRB	institutional review board
IXRS	interactive voice response system
LOCF	last observation carried forward
LPH	living with pulmonary hypertension
MedDRA	Medical Dictionary for Regulatory Activities
mRAP	mean right atrial pressure
mRRS	modified REVEAL Risk Score
NT-proBNP	N-terminal pro-brain natriuretic peptide
NYHA	New York Heart Association
OR	odds ratio
PAH	pulmonary arterial hypertension
PCA	prostacyclin analogues
PDE-5	phosphodiesterase 5
PDE-5i	phosphodiesterase 5 inhibitor
PH	pulmonary hypertension

PH-LVD	pulmonary hypertension due to left ventricular dysfunction
PoPH	portopulmonary PAH
PPS	per protocol set
PVR	pulmonary vascular resistance
QoL	Quality of Life
RHC(s)	right heart catheterization(s)
SAE	serious adverse event
SAP	statistical analysis plan
SAS	Statistical Analysis System
SBP	systolic blood pressure
SUSAR	suspected unexpected serious adverse reaction
TEAE	Treatment-emergent adverse event
TLC	total lung capacity
TID	*ter in die* (3 times daily)
ULN	upper limit of normal
WHO-DD	World Health Organization-Drug Dictionary
WHO FC	World Health Organization Functional Class

1. Introduction

Riociguat rEplacing PDE-5i therapy evaLuated Against Continued PDE-5i thErapy (REPLACE) is a prospective, randomized, international, multicenter, double-arm, controlled, open-label study of Riociguat in patients with pulmonary arterial hypertension (PAH) who are on a stable dose of phosphodiesterase-5 inhibitors (PDE-5i) with or without endothelin receptor antagonist (ERA), but not at treatment goal.

This statistical analysis plan (SAP) describes the study objective, study design, study population, safety and efficacy variables, statistical methods, and study tables to be used in REPLACE study. It is based on study protocol version 4.0, dated 06 Jan 2017and eCRF version 1.0, dated 19 April 2017.

Pulmonary arterial hypertension (PAH) is a devastating, life-threatening disease that is characterized by rapid progression and a high mortality. Traditionally, oral monotherapy including PDE-5i is one recommended treatment approach for patients with PAH who are classified as World Health Organization Functional Class (WHO FC) II or III. However, a significant proportion of PAH patients fail to reach or maintain treatment goals with PDE-5i monotherapy indicating that the NO-sGC-cGMP pathway in these patients may be impaired. The clinical approach in patients demonstrating an insufficient clinical response on PDE-5i monotherapy is typically to add an endothelin receptor antagonist (ERA).

Riociguat (BAY 63-2521) is a direct stimulator of the soluble Guanylate Cyclase. The randomized, double-blind, placebo-controlled clinical Phase III study (PATENT-1) investigated the efficacy and safety of riociguat in patients with pulmonary arterial hypertension (PAH). The study met the primary endpoint and showed statistically significant improvements in several secondary endpoints. Riocigut was overall considered to be beneficial with regard to exercise capacity, cardiopulmonary hemodynamics, and symptoms and well tolerated also in previous clinical studies at multiple doses in patients with PAH and other indications, e.g., chronic thromboembolic pulmonary hypertension (CTEPH) and pulmonary hypertension due to left ventricular dysfunction (PH-LVD).

The REPLACE study was designed to examine whether the replacement of PDE-5i by riociguat in patients who are not at treatment goal but on a stable dose of PDE-5i +/- ERA will lead to a significantly higher rate of satisfactory clinical response compared to patients who remain on PDE-5i +/- ERA because of the optimization of the NO sGC-cGMP pathway provided by riociguat and its NO-independent mechanism of action.

2. Study Objectives

The primary objective is to assess the proportion of patients in each treatment arm with a satisfactory clinical response as defined by a composite primary endpoint at Week 24.

The secondary objective is to demonstrate safety and clinical effect at Week 24 indicated by change in 6 minute walking distance (6MWD), N terminal pro-brain natriuretic peptide (NT-proBNP), World Health Organization Functional Class (WHO FC) and clinical worsening from baseline in each treatment arm.

3. Study Design

3.1 General Consideration

This is a prospective, randomized, international, multicenter, double-arm, 24 week, controlled, open-label study of riociguat in patients with PAH who are on a stable dose of PDE-5i +/- ERA, but not at treatment goal. The planned sample size is 218 (109 per arm). Central randomization is planned, stratified by etiology of PAH:

• IPAH/ HPAH/ PAH drug and toxin induced

• PAH-CHD, PAH-portopulmonary PAH (PoPH)

• PAH- CTD

At Week 24 a composite endpoint of satisfactory clinical response will be assessed.

Treatment period (Duration: 24 weeks)

Patients randomized to the control arm at the baseline visit (Week 0) will continue on their current PAH-specific treatment for 24 weeks.

Patients randomized to the riociguat arm, will have a wash-out period before starting the titration period of riociguat. The wash-out period is 24 hours with previous sildenafil therapy (daily dose 60 to 300 mg) and 48 hours after previous tadalafil therapy (daily dose 20 to 40 mg). Riociguat will be administered according to the established and approved dose adjustment/ titration scheme over 8 weeks. The starting dose will be 1 mg riociguat TID. The individual riociguat dose will be adjusted every 2 weeks according to the patient's well-being and peripheral systolic blood pressure (SBP) measured at trough, preferably before intake of the next morning dose. At each titration visit for the riociguat arm (Weeks 2, 4, and 6), the investigator needs to decide, based on the patient's SBP, whether the study medication dose should be modified.

Titration period (Duration: 8 weeks)

Patients randomized to the riociguat arm will have an 8-week titration period. The dose titration will be performed with the aid of an interactive voice response system (IXRS). The individual riociguat dose will be adjusted every 2 weeks as described above. While dose up-titration can only occur at the scheduled visits, a dose decrease can be performed at any time

based on the patient's SBP and well-being. In case of intended dose decrease which can be done independently from any planned study visit, the investigator will contact the IXRS and request a dose adjustment. This visit is to be declared as an unscheduled visit.

During the titration period, a request of an increase of riociguat dose is only possible at the planned titration visits at Weeks 2, 4, and 6. The dose reached at the end of the titration period at Week 8 is considered the patient's optimal dose based on SBP and well being.

Patients in the control arm will also have home visits at Weeks 2, 4, and 6 and will attend the Week 8 visit. At the discretion of the investigators and if applicable (e.g., short distance to study center) the home visits may also be performed at the study center.

Maintenance period (Duration: 16 weeks)

Riociguat should be continued at the optimal dose as determined at the end of the titration period (Week 8) throughout the maintenance period.

Visits (for both arms) will take place at Week 16 and Week 24. The IXRS will automatically allocate the right riociguat dose at each visit.

Dose reductions or discontinuation of riociguat for safety reasons are allowed at any time. Increases or re-increases in 0.5 mg steps (maximum dose 2.5 mg) are possible at the investigator's discretion weighing the benefit with potential risks implied, e.g., hypotension during the maintenance period.

End of Treatment (Week 24 ± 4 days)

All patients randomized to either treatment arm should perform Visit 4 / end of treatment (EOT). This is the last visit of the maintenance period, when all relevant efficacy and safety measurements will be performed.

Safety follow-up Visit (30 days ± 5 days)

A safety follow-up (FU) visit 30 days (± 5 days) after discontinuation of treatment intake should be performed for both arms.

End of Study

The end of the study as a whole will be reached as soon as the last safety FU visit of the last patient according to the above definition has been reached in all participating countries (European Union [EU] and non EU).

3.2 Sample size

Sample size is determined based on results of Studies 12934/12935/16719 (PATENT-1/2, RESPITE) for riociguat. The estimate for the 'satisfactory clinical response' rate at Week 24 is 40% for the riociguat arm. Assuming a treatment effect of 50% (relative reduction, see next subsections) to the active control the estimate for a 'satisfactory clinical response' in the active control arm is 20%. The sample size of 218 patients is calculated using SAS Version 9.2 (proc power, two-sample-frequency, χ^2-test, two-sided alpha = 5%, power = 90%).

The screening failure rate is estimated at 15% and therefore a screening number of 257 patients is required. As it is planned to use LOCF for the primary analysis, no adjustment for dropouts is needed.

4. General Statistical Considerations

4.1 General Principles

The statistical evaluation will be performed by using the software package SAS release 9.3 or higher (SAS Institute Inc., Cary, NC, USA).

Unless otherwise noted, data will be analyzed by descriptive statistical methods: the number of data available, means, standard deviations, minimums, quartiles, medians, and maximums will be calculated for continuous data. Frequency tables will be generated for categorical data including number and percentage of subjects.

Definition of efficacy and safety endpoints, analyses strategies, structure of analyses datasets, and layout of analyses data displays are following Bayer Biostatistics and Statistical Programming standards.

4.2 Handling of Dropouts

A patient who discontinues study participation prematurely for any reason after randomization is defined as a "dropout". Subjects who drop out will not be replaced.

A subject who, for any reason (e.g., failure to satisfy the selection criteria), terminates the study before the time point used for the definition of "dropout" in this study is regarded a "screening failure" (which literally means usually before randomisation).

In all cases, the reason for withdrawal must be entered in the electronic case report form (eCRF) and in the subject's medical records.

4.3 Handling of Missing Data

In order to achieve the goal of a well conducted clinical trial according to International Council on Harmonisation Good Clinical Practice, every effort should be made to collect all data (i.e., if a subject misses a scheduled assessment, the site personnel should contact the subject and request him/her to come to the clinic for the visit). However, despite best efforts, it may be inevitable that missing or incomplete data are reported. All missing or partial data will be presented in the subject data listings, as they were recorded on the eCRF.

For efficacy and safety variables analyzed by visits, the values at scheduled visits will be used if available. If values at scheduled visits are missing but values are available at an associated unscheduled visit (within the analysis phase), the value from the closest unscheduled visit will be used to the scheduled visit.

If there are no values at scheduled or unscheduled visits, missing data are not imputed in statistical analyses, except as noted below (Section 4.3.1 - 4.3.3).

A value from the termination visit (for discontinued subjects) will be assigned to the next scheduled visit (for the particular assessment per protocol) after the last visit the subject has completed.

4.3.1 Handling of missing efficacy endpoint

The primary efficacy endpoint ‘satisfactory clinical response’ is defined as the composite endpoint comprising 2 major components. It is defined as ‘YES’ only if the 2 major composites (‘2 of 3 items for the first component’ and ‘No clinical worsening for the second component’) are both ’YES’. The component ‘2 of 3’ is ‘YES’ if at least 2 of its 3 sub-components are ‘YES’.

For the primary efficacy endpoint with missing data, the last observation carried forward (LOCF) will be utilized for each individual component for subjects who are still alive during the study. The composite outcome is then derived based on the LOCF imputed individual component. Where a subject dies with no termination visit, the following rules will be used: 6MWD worst possible score (0m), Modified REVEAL with worst risk score (22), LPH worst possible score (105), WHO functional class, in case of death before visit 4 the worst possible score value plus 1 (V). NT pro-BNP is the efficacy measures designed to the direct effect of study treatment on biomarker, so it is not appropriate to use imputation rule for those in case of death or missing post-baseline data.

4.3.2 Safety variables

Incomplete adverse event start or stop dates and concomitant medication start or stop dates will be imputed to determine treatment emergent AE and concomitant medication as

described in Table 1. Missing or partial dates will be presented in the subject data listings as they were recorded on the CRFs.

Table 1. Imputation Rules for Incomplete Dates for AE and Concomitant Medication

	Missing	Scenario	Imputation
Start date (AE, concomitant medication)	Day	Event month&year < month&year of first dose	last day of the month
		Event month&year = month&year of first dose	first dose date
		Event month&year > month&year of first dose	first day of the month
	Day / Month	Event year < year of first dose	December 31
		Event year = year of first dose	first dose date
		Event year > year of first dose	January 01
	Complete missing		AE: first dose date Conmed: no imputation
Stop date (AE, concomitant medication)	Day	Event month&year < month&year of last known assessment date (LKAD)	last day of the month
		Event month&year = month&year of LKAD	date of LKAD
		Event month&year > month&year of LKAD	no imputation
	Day / Month	Event year < year of LKAD	December 31
		Event year = year of LKAD	date of LKAD
		Event year > year of LKAD	no imputation
	Complete missing		date of LKAD

Note: If first dose date is missing, use enrollment date for the imputation.
If the imputed start date is after the stop date, the start date will be imputed to equal the stop date.
If the imputed stop date is before the start date, the stop date will be imputed to equal the start date.

4.3.3 Laboratory values

For laboratory values below the lower limit of quantification (LLOQ) like "<xxx" or "≤xxx", or above the upper limit of quantification (ULOQ) like ">xxx" or "≥xxx", LLOQ or ULOQ (xxx) will be used for calculation of descriptive statistics. The original laboratory values ("<xxx", "≤xxx", ">xxx" or "≥xxx") are presented in the listing.

4.4 Interim Analyses and Data Monitoring

No interim analysis is planned for this study.

4.5 Data Rules

Generally, for each date stored in database a set of organizational variables will be derived to describe the temporal context of that date in the specific study: phase of treatment (pre-, during or post study treatment), the day relative to the start of study treatment, the day relative to the end of study treatment.

Refer to Section 4.3 for handling of missing data, as well as to Section 4.7, for specific endpoints data rules, for example: definition of treatment-emergent adverse events (TEAEs).

4.6 Protocol Deviation

The subjects who meet the criteria below will be flagged as having an important protocol deviation that leads to exclusion from per-protocol set.

Important protocol deviations leading to exclusion for per-protocol set are as follows:

- Subjects did not sign informed consent
- Subject was randomized but never received any study medication
- Subjects subsequently found not to satisfy the definition of PAH, including hemodynamic definition in RHC from medical history
- Baseline 6MWT less than 165 m or greater than 440 m
- Age less than 18 years at visit 1
- Concomitant use with riociguat of the following specific medications for treatment of PAH during the study: (1) PDE 5i (e.g., sildenafil, tadalafil or vardenafil) with riociguat; (2) Non-specific PDE-inhibitors (e.g., dipyridamole, theophylline) and (3) NO donors (e.g., nitrates, amyl nitrite);
- Prostacyclin analogues (PCA) and prostacyclin-receptor agonists (PRA) by any administration route within 30 days prior to and at randomization (except for vasoreactivity testing).

In addition to the deviations listed above, subjects will be excluded from analysis sets based on team review prior to database lock.

4.7 Analysis definitions, variables and subgroups

4.7.1 Analysis definitions and variables

A summary of the planned study analysis definitions and variables is provided in Table 2.

Table 2 - Summary of Analysis Definitions/Variables

Terms	Definition
Study Day 1	Day 1 is defined as randomization date. If an Adempas therapy is postponed, the day at start of wash-out is Day 1.
Study Day	Study Day = Measurement Date - Day 1 + 1, if measurement date is on or after Day 1. Study Day = Day 1 - Measurement Date, if measurement date is before Day 1.
Treatment Exposure	Treatment exposure is defined as duration of study treatment, calculated as last study treatment dose date - first study treatment dose date + 1.
Baseline	Baseline is defined as the value taken at randomization. If this is missing, the latest value before will be taken.
Change from Baseline	Change from baseline is defined as the difference between the post-baseline measurement and the baseline.
Titration Period	Patients randomized to the riociguat arm, will have a wash-out period of 24 hours with previous sildenafil therapy and 48 hours after previous tadalafil therapy before starting the 8-week titration period of riociguat. The dose titration will be performed with the aid of IXRS.
Maintenance Period	Once the subject completes titration period, he/she is entering maintenance period till the end of study.
Treatment Emergent Adverse Event (TEAE)	Treatment emergent adverse event is defined as having started or worsened after the first treatment administration up to 2 days after end of treatment.

End of Treatment	All patients randomized to either treatment arm should perform Visit 4 / EOT. This is the last visit of the maintenance period, when all relevant efficacy and safety measurements will be performed.
Safety follow-up phase	A safety follow-up (FU) visit is 30 days (± 5 days) after discontinuation of treatment intake.
End of Study	The end of the study as a whole will be reached as soon as the last safety FU visit of the last patient according to the above definition has been reached in all participating countries (European Union [EU] and non EU).

4.7.2 Subgroups

The following subgroups are defined in support of efficacy and safety analysis :

- PAH class as defined in randomization strata:
  - IPAH/ HPAH/ PAH drug and toxin induced
  - PAH-CHD, PAH-portopulmonary PAH (PoPH)
  - PAH- CTD
- Pre treatment: combination therapy with ERAs and PDE-5i vs. PDE-5i monotherapy
- ERA pre-treatment subgroups (Bosentan, Ambrisentan and Macitentan)
- Sex: male vs. female
- 6MWD at baseline: (<320 m vs. >=320 m)
- Pre treatment: tadalafil vs. sildenafil
- Age group: <65 years old vs. >= 65 years old
- Smoking status: current and former vs. never
- cMRI (subjects who performed cMRI vs. not)

5. Analysis Sets

Three analysis sets are defined as below.

Full Analysis Set (FAS)

The subjects who are randomized and take at least 1 medication will be considered for "FAS". Subjects will be analyzed as randomized.

Safety Analysis Set (SAF)

The population for safety analysis will be comprised of all subjects who received at least 1 dose of study drug. SAF set will be analyzed as treated.

All patients randomized and treated (at least 1 treatment after randomization) will be valid for the FAS, hence the number of patients in the safety and FAS will be identical. The actual treatment is defined as the majority study drug the subject received.

Per Protocol Set (PPS)

The PPS is defined as all FAS subjects who have no important protocol deviations that lead to exclusion in per-protocol set. The important protocol deviations are those protocol deviations that are considered to affect subject efficacy evaluations and subjects with such deviations should be excluded from the efficacy analysis. These important protocol deviations are defined in Section 4.6.

In addition to the criteria listed above, subjects will be excluded from analysis sets based on team review prior to database lock.

6. Statistical Methodology

The formal statistical analysis will be both descriptive and inferential. Summaries will be provided for each of the treatment groups.

6.1 Population characteristics

For this analysis, only descriptive statistics will be provided (no testing performed). The results will be displayed by treatment group as well as the total study population for FAS, SAF and PPS.

6.1.1 Disposition of subjects

Summary tables will be presented for the following subject populations:

- Subjects screened (subjects who signed informed consent)
- Screening failures (subjects who terminated the study before randomization)
- Subjects in FAS/ SAF
- Subjects in PPS
- Premature termination of treatment

- Subjects who completed the study
- Subjects who prematurely withdrawn from study medication in the study

In addition, detailed reasons for subjects who prematurely withdraw from the study or discontinue the study drug will be tabulated by treatment group.

6.1.2 Demographic and baseline characteristics

Standard descriptive statistics will be presented for the following variables of:

- Age
- Sex
- Race
- Ethnicity
- Weight and height (screening only)
- Body Mass Index
- Smoking status
- Pre treatment: combination therapy with ERAs and PDE-5i vs. PDE-5i monotherapy
- ERA pre-treatment subgroups (Bosentan, Ambrisentan and Macitentan)
- 6MWD at baseline: (<320 m vs. >=320 m)
- Pre treatment: tadalafil vs. sildenafil
- cMRI (subjects who performed cMRI vs. not)

Descriptive analysis will also be provided for baseline Dana Point classification of PH, time of first diagnosis of PAH to randomization, time from the confirmatory RHC to randomization and PAH classes at baseline.

6.1.3 Medical history

The dictionary for coding is the Medical Dictionary for Regulatory Activities (MedDRA) (version 17.1 or higher). By treatment group summary statistics (frequency and percentage) will be provided by system organ class and preferred term.

6.1.4 Prior and concomitant medications

Medications received prior to or concomitantly with treatment will be coded using the WHO Drug Dictionary Version Q3 2015, or later if updated during the study, and the Anatomical Therapeutic Chemical (ATC) Classification codes.

Prior medications are those with a stop date before the first dose date of study treatment. Concomitant medications are those with a start date on or after the first dose date of study treatment, or those with a start date before the first dose date of treatment and a stop date on or after the first dose date of study treatment.

To handle classification of medications in the event of missing or partial dates, please refer to Section 4.3.2 for imputation rules.

Prior and concomitant medications will be tabulated by treatment group using ATC codes. Frequencies of subjects having received medication from each drug category (ATC-level 2 and generic term) will be provided.

6.2 Efficacy

6.2.1 Primary efficacy variable

6.2.1.1 Definition of 'satisfactory clinical response'

The primary efficacy endpoint 'satisfactory clinical response' is defined as the composite endpoint comprising the following major components (independent central adjudication):

• 2 of 3 must be fulfilled

- 6MWD increase by ≥ 10% or ≥ 30 m from baseline to Week 24
- WHO FC I or II at Week 24
- NT-proBNP reduction ≥ 30% from baseline to Week 24 (NT-proBNP ratio Week 24/baseline ≤ 0.7),

AND

• No clinical worsening (Defined in the protocol Section10.3.2.1)

'Satisfactory clinical response', also referred to as improvement rate, is defined as 'YES' only if the 2 major composites ('2 of 3' and 'No clinical worsening') are both 'YES'. The component '2 of 3' is 'YES' if at least 2 of its 3 sub-components are 'YES'. The biomarker is given the same importance as the clinical parameters to introduce more objectivity to the endpoint and reduce possible open-label bias.

6.2.1.2 Final analysis of primary efficacy variable

The null hypothesis tested for the primary endpoint of efficacy is that there is no difference in the satisfactory clinical response rates in terms of odds ratio (OR) when treated with riociguat

compared to patients that remain on their previous therapy (i.e., H_0: OR = 1). The two-sided alternative is that there is a difference (i.e., H_1: OR ≠ 1).

For statistical analysis of the primary endpoint with missing data, the LOCF will be utilized. The reason for LOCF is given in the following: the primary endpoint measures improvements, thus, using LOCF will (a) not yield favorable results in the riociguat arm and (b) yield favorable results in the active control arm. Because of the inclusion criterion "stable on PDE5i " one expects the first case (a) to be much more frequent. Thus, LOCF is conservative with respect to the null hypothesis. For composite endpoints each single component is independently replaced with the last available observation, respectively. The composite outcomes will then be derived based on the imputed individual items.

Both groups (riociguat vs. active control) will be compared using a stratified Mantel-Haenzel test with a two-sided alpha level of 5%. Stratification factor is the PAH classes at baseline based on eCRF. The point estimate and 95% CI for odds ratio (OR) will be derived. Both FAS and PPS will be used for this analysis. Hypothesis testing is based on p-value. In some cases MH confidence intervals might not fit to p-value due to non-continuos variable.

6.2.1.3 Sensitivity analysis of primary efficacy variable

The sensitivity analysis for the primary efficacy variable will be conducted utilizing the following 3 approaches. The first approach is to consider the data is missing completely at random and the other two approaches will consider the data is not missing at random.

The primary efficacy outcome will be considered as missing at a given visit if (1) the clinical worsening status is unknown; or (2) 2 of 3 (6MWD, WHOC FC and NT-proBNP reduction) are missing or (3) 1 of 3 (6MWD, WHOC FC and NT-proBNP reduction) is missing and the other two known items with one outcome of yes and one outcome of no.

6.2.1.3.1 Missing completely at random

To assess the influence of missing data, a supportive analysis without LOCF will be performed with a generalized estimating equations approach (GEE, binomial distribution) utilizing all adequate investigations of all primary variable components at any time from baseline up to Week 24.

This method covers missing values of type 'missing completely at random'. Data from all assessed time points (i.e., Week 8, Week 16 and Week 24) will be included as the dependent variable in the model. PAH class, treatment group, visit and interaction between treatment and visit will be included as the predictors in the model, with visit as a continuous variable. The alternating logistic regression method will be used with exchangeable log odds ratios. . The 95% CIs will be constructed for OR of improvement rate between riociguat and PDE-5i at Week 8, Week 16 and Week 24.

6.2.1.3.2 Missing not at random

In this open-label study there might also be missing values of type 'missing not at random'. To assess the influence of these missing values a sensitivity analysis is planned, multiple imputation with penalty and tipping point analysis will also be performed for the improvement rate at Week 24.

Imputation with penalty

A multiple imputation with penalty will be performed where the penalty is given by crossed improvement rates in both arms.

• Riociguat arm: imputation with a clinical improvement rate observed in active control arm

• Control arm: imputation with a clinical improvement rate observed in riociguat arm

Thus, the highest sensible rate of clinical improvement is the rate in the riociguat arm and this is applied to the control arm in the imputation process. Further, the lowest sensible rate of clinical improvement is the rate in the control arm and this is applied to the riociguat arm in the imputation process.

To be more specific, the multiple imputations will be carried out 10 times for missing data. Each time, the penalty is given by the crossed improvement rates in the other group using SAS procedure, (i.e., Proc MI). Each of these imputed datasets (which contains identical values for the non-missing data but different values for missing data) is analyzed using the logistic regression with covariates of PAH baseline class and treatment. If convergence problems were encountered during the logistic regression, the PAH category will be removed. The results from all these 10 imputed datasets are then combined together for overall inference using Proc MIANALYZE.

Tipping point analysis

In this analysis, all the observed data will be included as non-missing. The analysis will be performed using a general 3 steps approach for each given shift (pre-specified value used for imputation).

First, multiple imputation will be conducted using PROC MI to generate 10 imputed datasets by imputing missing data using a pre-specified shift assuming monotone missing pattern. The missing values for clinical improvement rate in each treatment group will be imputed independently.

Secondly, each of the imputed datasets that contained identical values of non-missing data but different values of imputed data for missing data is analyzed using logistic regression controlling for PAH baseline class and treatment.

Last, the results from all these 10 analysis are then combined together for overall inference using PROC MIANALYZE by considering within and between variances.

More specifically, control group will be imputed with an improvement rate ranging from 1% to 99% with a step of 1% to identify the tipping point while the missing data in the treatment group will be imputed using the same rate as others in the treatment group. Note, the tipping point is defined as the data point where the conclusion based on p-value is changed. It may or may not exist depending on the amount of missing data. Due to the MCMC nature, the first p-value where the conclusion is changed will be defined as the tipping point. For example, if the treatment vs. control has a p-value of 0.001, then the tipping point will be the one where the first p-value is bigger or equal to 0.05. The analysis results will demonstrate the robustness of the primary outcome when missing data is imputed independently for two groups.

The main difference between the imputation with penalty and tipping point is that imputation with penalty will impute missing data with only 1 given improvement rate from the other group, while tipping point will impute data many times with different values and demonstrate the robustness of the results while the missing data were imputed with all the plausible values.

6.2.1.4 Subgroup analysis of primary efficacy variable

Subgroup analyses will be performed for the primary efficacy variable for the following subgroups:

- PAH categories at baseline based on eCRF
    - o IPAH/ HPAH/ PAH drug and toxin induced
    - o PAH-CHD, PAH-portopulmonary PAH (PoPH)
    - o PAH- CTD
- Pre treatment with combination therapy with ERAs and PDE-5i versus pre treatment with PDE-5i monotherapy
- ERA pre-treatment subgroups (Bosentan, Ambrisentan and Macitentan)
- Gender
- 6MWD at baseline (<320m vs. >=320m)
- Pre treatment with tadalafil or sildenafil

- Age group: <65 years old vs. >= 65 years old
- Smoking status: current and former vs. never
- cMRI: Subjects who conducted cMRI vs. not

Categorization of stratification factors (PAH) will use the values from eCRF. Subgroups may be collapsed or revised due to sparse data at a level prior to database lock.

In each subgroup analysis, both groups (riociguat vs. active control) will be compared for the primary efficacy endpoint 'satisfactory clinical response' using a Cochrane Mantel-Haenzel test with a two-sided alpha level of 5%. The point estimate and 95% CI for odds ratio (OR) will be derived.

Besides all the subgroup analysis, a forest plot will also be presented to illustrate the OR and its 95% CI for each subgroup.

6.2.2 Secondary efficacy variable

The following secondary efficacy variables will be examined at Week 8, 16 and 24. The null hypothesis tested for the secondary efficacy endpoints is that there is no difference when treated with riociguat compared to patients who remain on their previous therapy. The two-sided alternative is that there is a difference, respectively.

6.2.2.1 Change from baseline in 6MWD

The mean change and percentage change of 6MWD from baseline at each time point will be calculated for each group and total. The categories for 6MWD increase by $\geqslant$10% or $\geqslant$30m will also be displayed.

Especially, the change of 6WMD at Week 24 from baseline will be compared between riociguat and active control group using stratified Wilcoxon test with a two-sided alpha level of 5%. Mean difference and 95% CI for the mean difference obtained using tt test. The stratification factor will be PAH class at baseline. The analysis will be performed with and without missing data imputed using LOCF.

6.2.2.2 Change from baseline in NT-proBNP

The mean change and percent change of NT-proBNP from baseline at each time point will be calculated for each group and total. The number and percentage for the subjects wheather for NT-proBNP reduction by $\geqslant$30% (NT-proBNP Week 24/baseline ratio $\leq$ 0.7) will also be displayed.

Especially, the change of NT-proBNP at Week 24 from baseline will be compared between riociguat and PDE5i group using stratified Wilcoxon test and the stratification factor will be PAH classes at baseline. Mean difference and 95% CI for the mean difference obtained using tt test. The analysis will be performed with and without missing data imputed using LOCF.

6.2.2.3 Change from baseline in WHO FC

The status for WHO FC will be described at each time point. It is envisaged that WHO functional class will either remain the same, improve by 1 or 2 categories or deteriorate by 1 or 2 categories. A change score from baseline could go from -2 (class III at baseline and class I at the end of study) and +2 (class III at baseline and class V, death at the end of study). The change from baseline in WHO FC between two groups and total will be analyzed using Wilcoxon test stratified by PAH classes at baseline. Mean difference and 95% CI for the mean difference obtained using tt test. The analysis will be performed with and without missing data imputed using LOCF.

6.2.2.4 Clinical worsening

The adjudicated outcome for clinical worsening at Week 24 will be analyzed descriptively. Stratified Mantel-Haenzel test with a two-sided alpha level of 5% will be used to compare these two groups. Stratification factor is the PAH classes at baseline based on eCRF. The point estimate and 95% CI for OR will be presented.

The distribution of time from date of first dose of study drug to clinical worsening for each treatment group will be generated using Kaplan-Meier (KM) estimate. Subjects who do not have clinical worsening reported will be censored at the EOS. For each percentile (25th, 50th, 75th) if reached, an estimate of time of clinical worsening from date of first dose and its associated 95% CI using Brookmeyer-Crowley methodology will be presented for each treatment group. The survival rates at Week 8, 16 and 24 for each treatment group will also be presented.

6.2.2.5 Multiplicity

A multiplicity correction is not necessary for the primary endpoint because only one primary endpoint is defined. Secondary endpoints are tested hierarchically strictly according to the order listed in Table 3. An alpha adjustment is therefore not necessary for secondary and exploratory endpoints.

Table 3: Hypothesis testing order

Order	Endpoints
1	Satisfactory clinical response at Week 24
2	Change from baseline in 6MWD
3	Change from baseline in NT-proBNP
4	Change from baseline in WHO FC
5	Clinical worsening

6.2.3 Exploratory variable

Summary tables will be provided for the exploratory variables at each visit (number and percentage of subjects for categorical variables and descriptive statistics of absolute values, as well as the change from baseline values for continuous variables):

6.2.3.1 cMRI (core laboratory, reported separately)

cMRI is performed on a subset of patients. Descriptive analysis will be provided for the information including right ventricular function and all the other variables included in cMRI.

6.2.3.2 Living with Pulmonary Hypertension questionnaire (LPH)

The LPH questionnaire is designed to measure the effects of Pulmonary Hypertension (PH) and PH specific treatments on an individual's quality of life. The questionnaire and scoring method for LPH is defined in the protocol Section 16.5. In order to minimize the missing value for the total score, if the answer for a subquestion is missing, the LOCF for that particular question will be utilized to derive the total score if the previous score exists. Otherwise, the worst score, 5 will be used in order to derive the total score. The descriptive analysis for LPH total score at each time point and the change from baseline will be presented. A stratified Wilcoxon test for the change of LPH score at Week 24 from baseline will be used to determine if there is a difference between two treatment groups.

6.2.3.3 Modified REVEAL Risk Score (mRRS) and mRRS category

The Registry to Evaluate Early and Long-Term PAH Disease Management (REVEAL) risk score calculator is a validated assessment tool that has been developed to predict the 1-year survival of patients with PAH. The scoring method for mRRS is defined in protocol Section 16.4 with a range from 0 (lowest risk) to 22 (highest risk). In order to minimize the missing value for the total score, if the answer for a subquestion is missing, the LOCF for that particular question will be utilized to derive the total score if the previous score exists. Otherwise, the worst score for that particular question will be used in order to derive the total score.

There are 5 risk groups defined for mRRS in Table 4:

Table 4. Risk scores and predicted 1-year survival for 5 specified risk strata for mRRS:

Risk score ranges	Risk group	Predicted 1-year survival
1 – 7	low-risk	95% to 100%
8	average-risk	90% to <95%
9	moderately high-risk	85% to <90%
10 – 11	high-risk	70% to <85
≥ 12	very high-risk	<70%

The descriptive analysis will be presented for total mRRS score, change and percentage change from baseline and mRRS risk group at each time point. A stratified Wilcoxon test for the change from baseline of the mRRS score at Week 16 and Week 24 from baseline will be used to compare if there is a difference between the two treatment groups. A shift table will be also presented to demonstrate the change of risk category from baseline to Week 24.

All the efficacy analyses will be performed on both FAS and PPS. During the analysis, if convergence problem is encountered, the stratification factor PAH category will be removed from the model.

6.2.3.4 Change in other biomarkers (ADMA, cGMP, GDF 15 [plasma], ST 2).

The descriptive analysis will be presented for all the other biomarkers parameters (AMDA, cGMP, cGMP, GDF-15 and DT-2), including change and percentage change from baseline at each time point.

6.3 Pharmacokinetics / pharmacodynamics

Not applicable.

6.4 Safety

Analyses of safety variables will be performed on the SAF, unless otherwise specified. The summaries will be provided by treatment group and overall. No statistical comparison will be performed between the two treatment groups.

6.4.1 Extent of exposure

Descriptive statistical summaries will be provided for the SAF, FAS and PPS by treatment group for the following variables:

- treatment duration (calculated from the first dose of study medication to the date the last dose of study medication was ingested)

- number of tablets received during titration period (first 8 weeks), maintenance period and overall study
- average daily dose during titration period, maintenance period and overall study period
- treatment compliance (calculated as total number of administered doses / total number of scheduled doses*100) during titration period, maintenance period and overall study period
- number of subjects requiring dose up-titration, down-titration and without titration at each visit for riociguat group during titration period

6.4.2 Adverse events

All adverse events (AEs) will be coded using the most updated version of MedDRA.

AEs will be considered treatment-emergent if the events occurred or worsened on or after the date of the first dose of study drug and up to the 2 days after the end of treatment. Missing or partial data will be handled using the rules listed in section 4.3.2 in order to determine whether the AE is treatment-emergent.

AEs that occur before the first study dose will be considered AEs before first study dose and AEs that occur more than 2 days after the end of treatment will be considered AEs during the follow-up period.

All AE data including wash-out period for Adempas arm will be listed, including, onset relative to start of treatment, duration, intensity, relationship to study drug/protocol required procedure, action taken, outcome, and the coded terms. In addition, deaths, serious adverse events (SAEs), severe AEs, and AEs leading to study drug discontinuation or withdrawal from the study will be listed separately (if applicable).

An overview table will summarize by treatment and overall the number of events and the number and percentage of subjects with at least one of the following AEs:

- any AE before first study dose
- any AE during the follow-up period
- any TEAE
- any serious TEAE
- any severe TEAE
- TEAE leading to dose reduction
- TEAE leading to drug interruption
- TEAE leading to discontinuation of the study drug
- AE with the outcome of death

- any study drug-related TEAE
- any study drug-related serious TEAE
- any protocol-required procedure related TEAE
- TEAEs starting prior to week 8 visit
- TEAEs starting on or after week 8 visit

The number and percentage of subjects reporting each AE will be summarized by SOC and Preferred Term (PT). Tables will be sorted alphabetically by SOC. PTs will be sorted by descending overall total frequency within an SOC. The following summaries will be produced:

- TEAEs, by SOC and PT
- TEAEs starting prior to week 8 visit, by SOC and PT
- TEAEs starting on or after week 8 visit, by SOC and PT
- AEs during the pre-treatment period, by SOC and PT
- AEs during the follow-up period, by SOC and PT
- TEAEs that occurred in more than 5% of the subjects, by SOC and PT
- TEAEs, by SOC and PT and by intensity
- TEAEs, by SOC and PT and by relationship to the study drug
- serious TEAEs, by SOC and PT
- serious TEAEs that occurred in more than 5% of the subjects, by SOC and PT
- serious TEAEs, by SOC and PT and by intensity
- serious TEAEs, by SOC and PT and by relationship to the study drug
- TEAE leading to dose reduction by SOC and PT
- TEAE leading to dose interruption by SOC and PT
- TEAE leading to discontinuation of the study drug by SOC and PT
- AEs with the outcome of death, by SOC and PT

In the above summaries, subjects with more than one AE within a particular SOC are counted only once for that SOC. Similarly, subjects with more than one AE within a particular PT are counted only once for that PT. In case a subject has events with different intensities , the maximum reported intensity will be used. If intensity is missing, the event will be considered as severe. Similarly, if the same event is reported as both unrelated and related to the study drug in one subject, the event will be reported as related to study drug. If the drug relationship is missing, the event will be considered as being related to the study drug.

The exposure adjusted subject incidence per 100-person year will be produced for the following tables. The risk period is defined as from the first dose date to the end of treatment and all the events, including the recurrent ones will be counted for exposure adjusted subject incidence.

- TEAEs, by SOC and PT
- serious TEAEs, by SOC and PT

6.4.3 Adverse events of special safety interest

Symptomatic hypotension and serious hemoptysis (which includes pulmonary hemorrhage) are considered AEs of special interest and will be tableted for each group. The following tables will be produced by treatment group:

- Treatment emergent hypotension, by SOC and PT
- Treatment emergent hypotension, by PT and by intensity
- Treatment emergent hypotension, by PT and by relationship to the study drug
- Treatment emergent hemoptysis, by SOC and PT
- Treatment emergent hemoptysis, by PT and by intensity
- Treatment emergent hemoptysis, by PT and by relationship to the study drug

The exposure adjusted subject incidence per 100-person year will be also produced for the symptomatic hypotension and serious hemoptysis/ pulmonary hemorrhage.

6.4.4 Safety laboratory assessments

The following lab parameters will be summarized by treatment group and visit. If data for any additional analytes are also recorded, these will be listed only.

• Hematology: white blood cell differential count, erythrocytes, hemoglobin, hematocrit, platelets

• Coagulation tests: prothrombin time, only for patients under anticoagulation therapy

• Clinical chemistry: AST, ALT, total bilirubin, serum albumin, creatinine, potassium

Biomarkers will be analyzed centrally according to the time points given in the protocol.

Out-of-reference-range laboratory values will be flagged as high (H) or low (L) in the listings.

Descriptive summaries will be provided by treatment group for laboratory test for both by-visit measures and change from baseline measures. Incidence rates of high lab abnormalities and low lab abnormalities will be reported. Shift tables presenting abnormal changes from baseline will be tabulated.

Additionally, the following pre-specified abnormalities/changes will be displayed in frequency table. It will also summarize a combined elevation of AST or ALT$\geqslant$ 3 times ULN and bilirubin $\geqslant$ 2 times ULN at the same visit.

ALT (U/L)	≥ 3 times ULN
AST (U/L)	≥ 3 times ULN
Total Bilirubin (mg/dl)	≥ 2 times ULN

6.4.5 Vital signs

Vital signs including systolic blood pressure (SBP), diastolic blood pressure (DBP), and heart rate (HR), measured at scheduled visits, will be summarized by treatment group and visit. Descriptive summaries will be provided for baseline values, post-baseline values and change from baseline.

6.4.6 12-lead electrocardiogram

The following quantitative electrocardiogram (ECG) measurements will be taken during the study:

- heart rate (beats/min)
- PR interval (msec)
- QRS interval (msec)
- QT& QTcB/QTcF interval (msec)

An overall Investigator assessment of ECG will be provided (categories “normal”, “abnormal, not clinically significant” and “abnormal, clinically significant”). Abnormal findings on 12-lead ECG will be summarized by treatment group by visit using descriptive statistics.

The ECG measurements in ECG will be summarized by treatment group using standard descriptive statistics. QT corrected for heart rate will additionally be described using Bazett’s formula (QTcB), Fridericia’s formula (QTcF) and Linear Correction Formula (QTcLC) .

6.4.7 Pregnancy test

Pregnancy tests with positive result for women at scheduled visits will be displayed in a listing.

6.4.8 Other procedures and variables

The descriptive analysis for pulse oximetry at baseline, including SaO2 (%),will be reported. The descriptive analysis for lung function testing at baseline including FEV1[L/s], FEV1 % of predicted, TLC [L] and TLC, % of predicted will also be reported.

7. Document history and changes in the planned statistical analysis

Changes from SAP Final v2 (26-Mar-2020) to SAP Final v3 (09-Apr-2020):

No change was made to the SAP. However, listing numbers were updated in the mocks.

Changes from SAP Final v1 (09-June-2017) to SAP Final v2 (26-Mar-2020):

Protocol deviations updated to reflect important deviations rather than major. Detail added regarding analysis sets being based on team review prior to database lock.

Subgroup analysis for Race has been removed.

Baseline characteristics updated to include subgroup analysis parameters.

Tipping point analysis updated to include an imputation with improvement range ranging from 1%-99% in control group whilst the treatment group is imputed at the observed rate.

Analysis on AEs updated to include TEAEs prior to week 8 visit and post week 8 visit. AEs leading to discontinuation of study by SOC and PT has been removed.

8. References

9. Appendix

9.1 Study schedule of procedures

	Mandatory Screening period (2 weeks)	Rando-mization (baseline)	Overall treatment period (24 weeks)								Follow-up period
Visit	**V 0**	**V 1 [a]**	**H 1 [b]**	**H 2 [b]**	**H 3 [b]**	**V 2 [c]**	**Tele-phone contact [d]**	**V 3**	**Tele-phone contact [d]**	**EOT**	**Safety FU [e]**
Week/Days	**W -2 (± 2 D)**	**W 0 (± 2 D)**	**W 2 (± 2 D)**	**W 4 (± 2 D)**	**W 6 (± 2 D)**	**W 8 (± 2 D)**	**W 12 (± 2 D)**	**W 16 (± 2 D)**	**W 20 (± 2 D)**	**W 24 (± 5 D)**	**30 D (± 5 D)**
		Rio Start									
Informed consent	X										
Check of inclusion/ exclusion criteria	X	X									
Lung function test	X										
Pulse oximetry	X	X									
Echocardiography (LVEF)	X										
Medical history	X										
Smoking status	X	X		X		X	X	X	X	X	X
Specific PAH medication	X	X									
Demographics	X										
Pregnancy test [f]	X	X		X		X	X	X	X	X	X
Physical examination	X									X	
Weight and height	X										
Dispense of riociguat			X	X	X	X	X	X			

	Mandatory Screening period (2 weeks)	**Rando-mization (baseline)**	**Overall treatment period (24 weeks)**								**Follow-up period**
Visit	**V 0**	**V 1** [a]	**H 1** [b]	**H 2** [b]	**H 3** [b]	**V 2** [c]	**Tele-phone contact** [d]	**V 3**	**Tele-phone contact** [d]	**EOT**	**Safety FU** [e]
Week/Days	**W -2 (± 2 D)**	**W 0 (± 2 D) Rio Start**	**W 2 (± 2 D)**	**W 4 (± 2 D)**	**W 6 (± 2 D)**	**W 8 (± 2 D)**	**W 12 (± 2 D)**	**W 16 (± 2 D)**	**W 20 (± 2 D)**	**W 24 (± 5 D)**	**30 D (± 5 D)**
medication [g]											
Drug accountability			X	X	X	X		X		X	
6MWD (blinded) [h]	X [i]	X				X		X		X	
Biomarkers (central lab)		X				X		X		X	
WHO FC (blinded) [h]	X [i]	X				X		X		X	
mRRS		X						X		X	
LPH (QoL) [l]		X								X	
Safety laboratory (local laboratory)	X							X		X	
Vital signs	X	X	X	X	X	X		X		X	X
ECG [j]	X										
Change in concomitant medication [h]	X	X	X	X	X	X	X	X	X	X	
AE assessment and reporting [h]	X	X	X	X	X	X	X	X	X	X	X
cMRI (exploratory, reported separately) [k]		X								X	

Abbreviations: 6MWD = 6-minute walking distance; AE = adverse event; cMRI = cardiac magnetic resonance imaging; D = Days; ECG = electrocardiogram; EOS = end of study; FU = Follow-up; LVEF = left ventricular ejection fraction; mRRS = modified Reveal Risk Score; NT-proBNP = N-terminal pro-brain natriuretic peptide; LPH =Living with Pulmonary Hypertension Questionnaire; PAH = Pulmonary Arterial Hypertension; QoL = Quality of Life; W = Week; WHO FC = World Health Organization Functional Class.

[a] Riociguat wash-out, first dose and dose titration:

Patients in the riociguat arm will have a **wash-out period between discontinuation of PDE-5i and the first dose of riociguat**: 24-hour wash-out following last dose of sildenafil and 48-hour wash-out following last dose of tadalafil.

Based on the time of the last PDE-5i dose, date and time of the first dose of riociguat after the wash-out period will be determined by the investigator and communicated to the patient. Before the first dose of riociguat a blood pressure measurement needs to be performed to ensure that systolic blood pressure is ≥ 95 mmHg. If the blood pressure is below this value, it can be measured again within 24 hours and if still below 95 mmHg, riociguat must not be administered. Measurement of blood pressure and application of first riociguat dose may be performed at the study center, or by an experienced nurse in the outpatient setting. Date and time of intake of the first dose of riociguat need to be documented by the investigator. The nurse will contact the center for this purpose via phone.

Patients in the control arm will continue their current specific PAH treatment (PDE-5i +/-ERA).

[b] Home Visits (H1, H2, H3):

These visits will be conducted by an experienced nurse at the patient's (both treatment arms) home. At the discretion of the investigators and if applicable (e.g., short distance to study center) these visits may also be performed at the study center.

For the riociguat arm these visit are part of the titration period that will be performed according to the titration scheme (until Week 8, Visit 2). While in the titration period dose increase can only occur at the scheduled visits, a dose decrease (-0.5 mg) can be performed at any time based on the patient's systolic blood pressure and well-being. In both arms blood pressure will be measured; in the riociguat arm the next riociguat dose will be decided in phone contact with the center according to the titration scheme.

[c] Visit 2, Week 8: This is the first regular visit at the center after randomization. For the riociguat arm it will also be the last visit of the titration period. Following the titration period, patients will continue to receive riociguat TID at the optimal dose achieved at Week 8 (start of maintenance period).

[d] Patients will be contacted via telephone every 4 weeks, if there is no scheduled visit to the study center.

[e] Safety Follow-up visit after last dose of study drug.

[f] Only women of childbearing potential will have a urine or serum test, depending on local routine at each study site. If test is performed > 48 hours before randomization, the test will need to be repeated.

[g] For the riociguat arm IXRS consultation will be performed to supply drug.

[h] Forms part of the clinical worsening assessment.

[i] Blinded assessment starts at randomization

[j] An ECG can be done at any time during the study, at the discretion of the investigator. It should be reported in the unscheduled procedure page.

[k] This will be performed by core laboratory and reported separately.

l To be performed if not available within last 6 months.

9.2 Sample SAS code

General Estimation Equation:

```
PROC GENMOD data=dataset descend;
    CLASS subjectid treatment  PAH ;
    MODEL outcome=treatment PAH visit visit*treatment / dist=bin;
    REPEATED subject=subjectid / logor=exch;
RUN;
```

Tipping point analysis

Step 1: Impute data with specified value:

```
PROC MI data=dataset seed=12345 nimpute=10 out=outmi;
          CLASS treatment PAH improvement;
           MONOTNE logistic;
          MNAR adjust (improvement(event='1')/shift=<log odds (shift1 + completer response rate) – log odds of completer response rate>      adjustobs(TRTP='Riociguat');
              adjust (improvement(event='1')/shift=<log odds (shift2 + completer response rate) – log odds of completer response rate>  adjustobs(TRTP='PDE-5i');
              VAR treatment PAH improvement;
RUN;
```

Step 2: Analyze the data with imputed datasets: (sort the data first before BY statement)

```
PROC LOGISTIC data=outmi descending;
         BY <_imputation_>;
          CLASS treatment PAH;
          MODEL improvement=treatment PAH / dist=binomial link=logit;
          ESTIMATE 'Week 24 Improvement rate difference for riociguat vs. PDE5i' TRTP 1 -1 /exp alpha=0.05;
```

```
        ODS OUTPUT ParameterEstimates=lgsparms CovB=lgscovb;
RUN;
```

Step 3: Combine results for inference:

```
 PROC MIANALYZE parms=lgsparms  covb=lgscovb;
        MODEL EFFECTS TREATMENT;
 RUN;
```