**Actelion Pharmaceuticals Ltd**
**Janssen Research & Development**

# Statistical Analysis Plan

## Prospective, multi-center, double-blind, randomized, placebo-controlled, parallel-group study assessing the efficacy and safety of macitentan in Fontan-palliated adult and adolescent subjects

**Protocol AC-055H301; Phase 3**

**JNJ-67896062/ACT-064992 (Macitentan)**

| | |
|---|---|
| **Status:** | Approved |
| **Date:** | 23 August 2021 |
| **Prepared by:** | Actelion Pharmaceuticals Ltd., a Janssen Pharmaceutical Company of Johnson & Johnson |
| **Document No.:** | EDMS-RIM-264216, 3,0 |

**Compliance:** The study described in this report was performed according to the principles of Good Clinical Practice (GCP).

# TABLE OF CONTENTS

## VERSION HISTORY

**Table 1: SAP Version History Summary**

| SAP Version | Approval Date | Change | Rationale |
|---|---|---|---|
| Final Version 1 (D-16.427) | 24 Aug 2016 | Not applicable | Initial release for special protocol assessment to the Food and Drug Administration Agency. |
| Final version 1 (D-19.382) | 13 Dec 2019 | Janssen SAP template<br>Introduction of the Interim Analysis for sample size re-estimation | Special protocol assessment re-submission of protocol version 8 to the Food and Drug Administration Agency. |
| Final version 2 (D-20.077) | 20 Feb 2020 | Section 3.2.3<br>Appendix 1 | To incorporate comments from Food and Drug Administration Agency received by email on 13 Feb 2020 . |
| Final version 3 | 25 Feb 2021 | Development of all analyses as detailed in Section 10.1 of this document | Clinical Study Report. |
| Final version 4 | 12 July 2021 | Sections 4.4.1 and 4.4.2<br>Clarification of derivation for compliance | With the introduction of the COVID-19 protocol appendix, on-site visits could be performed remotely or delayed, and while study treatment may have been dispensed via direct-to-patient shipments, subjects were asked to return empty bottles/unused tablets only at their next on-site visit. The overall compliance and the compliance up to Week 16 will be calculated only for selected subjects as detailed in Sections 4.4.1 and 4.4.2. |
| | | Sections 4.5.2, 4.6, and 9 | Minor clarification for implementation. |
| | | Section 5.2.3.4 Removal of the estimator based on different model 1 (by adding treatment-by-geographical region interaction) | In order to assess the consistency of the treatment effect across regions, the subgroup analysis is already planned in section 5.2.3.5 where the treatment-by-region interaction is investigated. |
| | | Sections 5.2.3.4, 5.3.1.6, and 5.3.2.6 Addition of sensitivity analyses for primary and secondary efficacy endpoints | To explore the robustness of the main analysis treatment effect estimates to missing data, multiple imputation sensitivity analyses will be added. |
| | | Section 5.4.4<br>Vital Status Follow Up (VSFU) not assessed for subjects randomized in site 1101 | VSFU could not be assessed in site 1101. Therefore, randomized subjects in these sites are assumed with unknown vital status and censored at their EOS, unless they died prior to EOS. |
| Final version 5 | 23 August 2021 | Section 2.3 Addition of subgroups: BMI at baseline, $SpO_2$ at baseline, Peak respiratory exchange ratio at baseline | After consultation with the Study Results Committee constituted of external clinicians with expertise in subjects with Functional Single Ventricle Heart Disease on |

| SAP Version | Approval Date | Change | Rationale |
|---|---|---|---|
| | | | 13-Aug-2021, additional subgroups have been added based on clinical consideration and blinded aggregated baseline characteristics at baseline, to assess the consistency of the treatment effect on the primary and secondary efficacy endpoints. |
| | | Section 4.1.2 Addition of 'sildenafil citrate' as preferred term to qualify PAH specific therapies | The preferred term 'sildenafil citrate' was erroneously not included in the definition of PAH specific therapies. |
| | | Sections 5.2.3.4 and 5.3.1.6 | Minor clarification on the exclusion of invalid baseline peak $VO_2$ from the sensitivity multiple imputation analyses. This clarification complies with the general estimand language and does not impact the planned efficacy analyses provided in the previous SAP version and as agreed with FDA. |
| | | Sections 5.2.3.4, 5.3.1.6, and 5.3.2.6 Additional sensitivity analysis for primary and secondary efficacy enpdoints | After consultation with the Study Results Committee constituted of external clinicians with expertise in subjects with Functional Single Ventricle Heart Disease on 13-Aug-2021, an additional estimand excluding subjects with cardiac heterotaxy defects (isomerism) will be added to assess the robustness of results (for each primary and secondary efficacy endpoint) excluding subjects with extra-cardiac defects. |
| | | Section 5.4.4 Vital Status Follow up for those subjects with the End of Study within 4 weeks of the Study Closure announcement. | Minor clarification for implementation. |
| | | Sections 6.2.1 Additional analysis for total bilirubin, ALT, AST | After consultation with the Study Results Committee constituted of external clinicians with expertise in subjects with Functional Single Ventricle Heart Disease on 13-Aug-2021, additional analyses on total bilirubin, ALT and AST will be added to characterize the pattern over time of these parameters.<br><br>Subjects who have undergone Fontan surgery and have a Functional Single Ventricle will sometimes experience hepatic congestion as a consequence of volume overload. As a result of |

| SAP Version | Approval Date | Change | Rationale |
|---|---|---|---|
| | | | this, bilirubin may be raised with/without corresponding rises in AST and ALT reflective of hepatocellular damage. Given macitentan's potential to cause vasodilation, there is the possibility that a beneficial effect may be seen on both bilirubin and transaminases. These analyses will explore this clinical consideration. |

## ABBREVIATIONS

| | |
|---|---|
| AE | Adverse Event |
| AESI | Adverse Events of Special Interest |
| ALT | Alanine aminotransferase |
| ANCOVA | Analysis of covariance |
| AP | Alkaline phosphatase |
| AST | Aspartate aminotransferase |
| AV | Atrioventricular |
| BMI | Body Mass Index |
| BP | Blood Pressure |
| CGI-C | Clinician Global Impression of Change |
| CGI-S | Clinician Global Impression of Severity |
| CDISC | Clinical Data Interchange Standards Consortium |
| CL | Confidence Limit |
| CPET | Cardiopulmonary Exercise Testing |
| CRO | Contract Research Organization |
| CSR | Clinical Study Report |
| CTCAE | Common Terminology Criteria For Adverse Events |
| CV | Coefficient of Variation |
| CYP3A4 | Cytochrome P450 3A4 |
| DTS | Data Transfer Specifications |
| EC | Extra Cardiac |
| ECG | Electrocardiogram |
| eCRF | Electronic Case Report Form |
| EC-TCPC | Extra Cardiac Total Cavopulmonary Connection |
| ENR | Enrolled Analysis Set |
| EOP | End-of-Post-treatment observation period |
| EOS | End-of-Study |
| EOT | End-of-Treatment |
| EQ-5D-5L | Euro Quality of Life 5D (dimensions) |
| ERA | Endothelin Receptor Antagonist |
| FAS | Full Analysis Set |
| FC | Functional Class |
| FDA | Food and Drug Administration |
| FU | Follow-up |
| GDRA | Global Drug Regulatory Affairs |
| HR | Heart Rate |
| ICF | Informed Consent Form |
| IDMC | Independent Data Monitoring Committee |
| IRT | Interactive Response Technology |
| KM | Kaplan-Meier |
| LLOQ | Lower Limit of Quantification |
| LT-TCPC | Lateral Tunnel Total Cavopulmonary Connection |
| LV | Left Ventricle |
| MAR | Missing at Random |
| MedDRA™ | Medical Dictionary for Regulatory Activities |
| MI | Multiple Imputation |
| MMRM | Mixed Model Repeated Measures |
| NT-proBNP | N-terminal prohormone of Brain Natriuretic Peptide |
| NYHA FC | New York Heart Association Functional Class |
| OL | Open Label |
| OLE | Open Label Extension |
| PA-Ac | Physical Activity measured by Accelerometer |
| PGA-S | Patient Global Assessment of Disease Severity |
| PGI-C | Patient Global Impression of Change |
| PD | Protocol Deviation |
| PK | Pharmacokinetic |

| | |
|---|---|
| PKS | Pharmacokinetic Analysis Set |
| PLE | Protein-Losing Enteropathy |
| PP | Per-Protocol Analysis Set |
| PTOP | Post-Treatment Observation Period |
| Q-Q | Quantile-Quantile |
| QoL | Quality of life |
| RV | Right Ventricle |
| SAE | Serious Adverse Event |
| SAP | Statistical Analysis Plan |
| SCL | Study Closure |
| SD | Standard Deviation |
| SDTM | Study Data Tabulation Model |
| S-FU | Safety Follow-Up |
| SI | International System of Units |
| SMQ | Standardised MedDRA Query |
| SOC | System Organ Class |
| $SpO_2$ | Peripheral Oxygen Saturation |
| SS | Safety Analysis Set |
| SSRE | Sample Size Re-Estimation |
| TCPC | Total Cavo-Pulmonary Connection |
| TEMPO | Treatment with endothelin receptor antagonist in fontan patients, a randomized, placebo-controlled, double-blind study measuring peak oxygen consumption |
| ULN | Upper Limit of Normal |
| VAS | Visual Analogue Scale |
| $VCO_2$ | Carbon dioxide production<br>[$\dot{V}CO_2$ is a flow = a ratio of Volume of Carbon Dioxide by unit of time] |
| VE | Pulmonary ventilation.<br>[Respiratory minute volume, the total Volume of gas Expired per minute. ($\dot{V}E$ is a flow = a ratio of a volume by unit of time)] |
| $VO_2$ | Oxygen uptake/consumption<br>[$\dot{V}O_2$ is a flow = a ratio of Volume of Oxygen by unit of time] |
| VSFU | Vital Status Follow-Up |
| WHO | World Health Organization |
| WU | Wood Unit |

# 1. INTRODUCTION

This Statistical Analysis Plan (SAP) describes the planned statistical data analyses of the double-blind study AC-055H301 (RUBATO) for the purpose of the Clinical Study Report (CSR).

This SAP refers to the documents listed in Table 2.

**Table 2: Study Documents**

| Document | Version |
|---|---|
| Study Protocol AC-055H301 (RUBATO) | Final version 10 (D-20.425) |
| SAP for Special Protocol Assessment | Final version 2 (D-20.077) |
| SAP for Independent Data Monitoring Committee | Final version 2.1 (D-18.337) |
| eCRF specifications | Final version 008 (or latest implemented version) |
| IDMC Charter | Final version 5.0 |

Source data for the analyses are provided as Statistical Analysis Software (SAS®) data sets according to Clinical Data Interchange Standards Consortium Study Data Tabulation Model (SDTM). Source data are provided according to the Database Release Plan document.

All descriptive or formal statistical analyses will be performed using SAS statistical software (Version 9.4), unless otherwise specified.

## 1.1. Trial Objectives

The **primary objective** of the study is to assess the effect of macitentan on exercise capacity (measured by peak oxygen uptake $VO_2$) in comparison with placebo in Fontan-palliated subjects.

The **secondary objectives** are:

- To assess the effect of macitentan on long-term exercise capacity (measured by peak $VO_2$ over 52 weeks).
- To assess the effect of macitentan on daily physical activity measured by accelerometer (PA-Ac).
- To evaluate the safety and tolerability of macitentan.

**Other objectives** are:

- To assess the efficacy of macitentan on endpoints related to exercise capacity.
- To assess the effect of macitentan on NT-proBNP.
- To assess the effect of macitentan on clinical worsening.
- To assess the effect of macitentan on Quality of life (QoL).
- To assess the effect of macitentan on pharmacoeconomic endpoints.
- To assess the plasma concentration of macitentan and its metabolite (ACT-132577) at trough.

## 1.2. Trial Design

This is a prospective, multi-center, double-blind, randomized, placebo-controlled, parallel-group Phase 3 study with an adaptive sample size re-assessment.

At least 134 subjects will be randomized in a 1:1 ratio to either macitentan or placebo. Randomization will be stratified by geographical region (America, Europe, Asia, Oceania). The study will be conducted in approximately 31 investigational sites in 11 countries.

Given the uncertainty about both the true treatment effect and variability of the primary efficacy endpoint variable, the Sponsor made the decision on 24 January 2019 (based on an internal meeting) to implement an interim analysis (IA) for unblinded sample size re-estimation (introduced from protocol version 8). This IA is performed by an Independent Statistical Analysis Center (ISAC) / Statistical Support Group (SSG) for the Independent Data Monitoring Committee (IDMC). Only the IDMC and the ISAC/SSG are unblinded to the data. This sample size re-estimation allows the sample size to be potentially increased up to a total sample size of 268 subjects. Full details about the IA are presented in Section 3.

The study comprises the following consecutive periods:

**Screening period:** Starts with the signature of the informed consent (Visit 1) and ends with the subject's Randomization (Visit 2).

**Treatment period:** Starts with the administration of the first dose of study treatment (Visit 2 / Randomization) and ends on the day of the last dose of study treatment, expected to be at 52 weeks.

Subjects who prematurely and permanently discontinue study treatment before Week 52 (Visit 6) must have a premature End-of-Treatment (EOT) visit as soon as possible but not later than 7 days after last dose of study drug.

**Safety follow-up (S-FU) period:** After permanent study treatment discontinuation, subjects will be followed up for a minimum of 30 days (S-FU).

Subjects eligible to enter the AC 055H302 open-label (OL) extension study may do so as soon as their eligibility has been confirmed, even prior to completion of the full S-FU period. The period ends with administration of first study drug in the OL extension study.

**Post-treatment observation period (PTOP):**

Subjects who prematurely and permanently discontinue study treatment will enter the PTOP after their S-FU period is completed. The PTOP lasts until the planned study EOT (at most 52 weeks after Randomization) for each individual subject. This last visit of the PTOP is called End-of-PTOP (EOP). The PTOP consists of an abbreviated schedule of assessments at 90-day intervals.

**End-of-Study (EOS):**

EOS is reached by an individual subject:

- When the S-FU period has been completed (for those subjects who do not discontinue study treatment prematurely), or

- When the S-FU period and PTOP have been completed (for those subjects who discontinue study treatment prematurely), or
- If subjects enter the AC-055H302 RUBATO OL extension study, with first administration of OL study drug.

For all subjects, EOS corresponds to the last visit performed in this AC-055H301 RUBATO double-blind study.

Study closure (SCL) will be announced by the sponsor when all subjects have reached their EOS in AC-055H301 RUBATO.

**Vital status follow-up (VSFU):**

A VSFU will be performed within 8 weeks after the SCL announcement to determine vital status of all subjects randomized in AC-055H301 RUBATO, irrespective of whether the subject is included in the AC-055H302 OL extension study.

**SCL:**

The SCL applies to all randomized subjects and will occur:

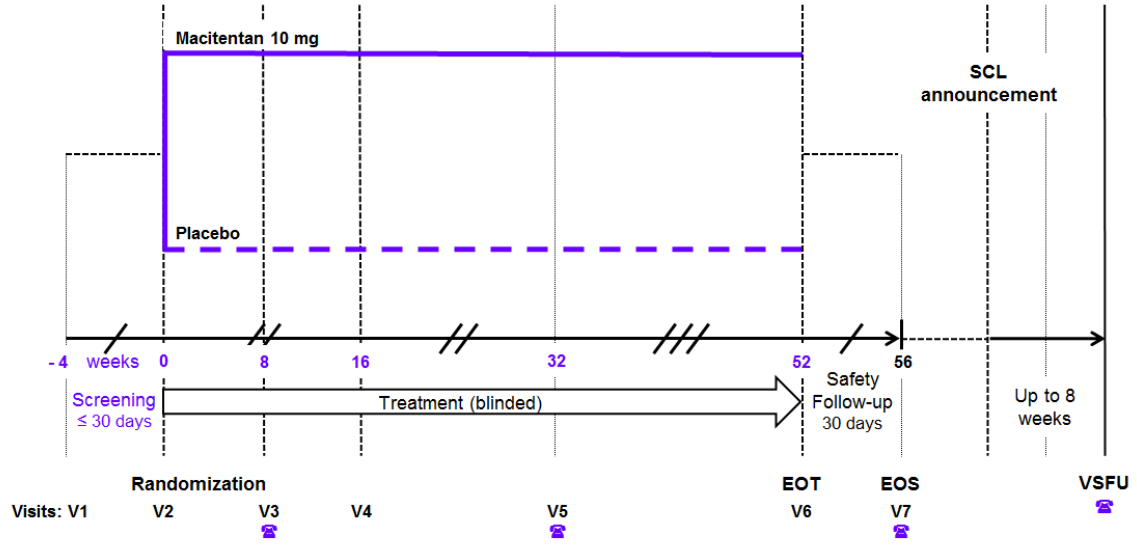
- When all subjects have had their individual VSFU, and
- No later than 8 weeks after SCL announcement.

The visit schedule and protocol-mandated procedures are performed according to the table of assessments in the protocol.

The overall study design is depicted in Figure 1. An example of premature discontinuation and PTOP is depicted in Figure 2.

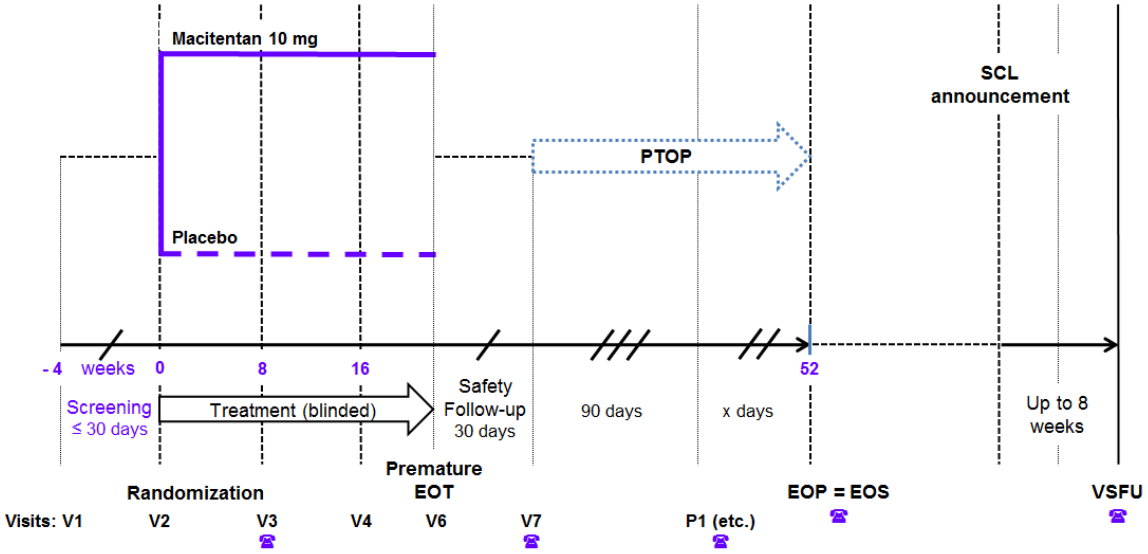
Figure 3 depicts the design for a subject who enters the AC-055H302 OL extension study at Week 52.

**Figure 1: Study design**



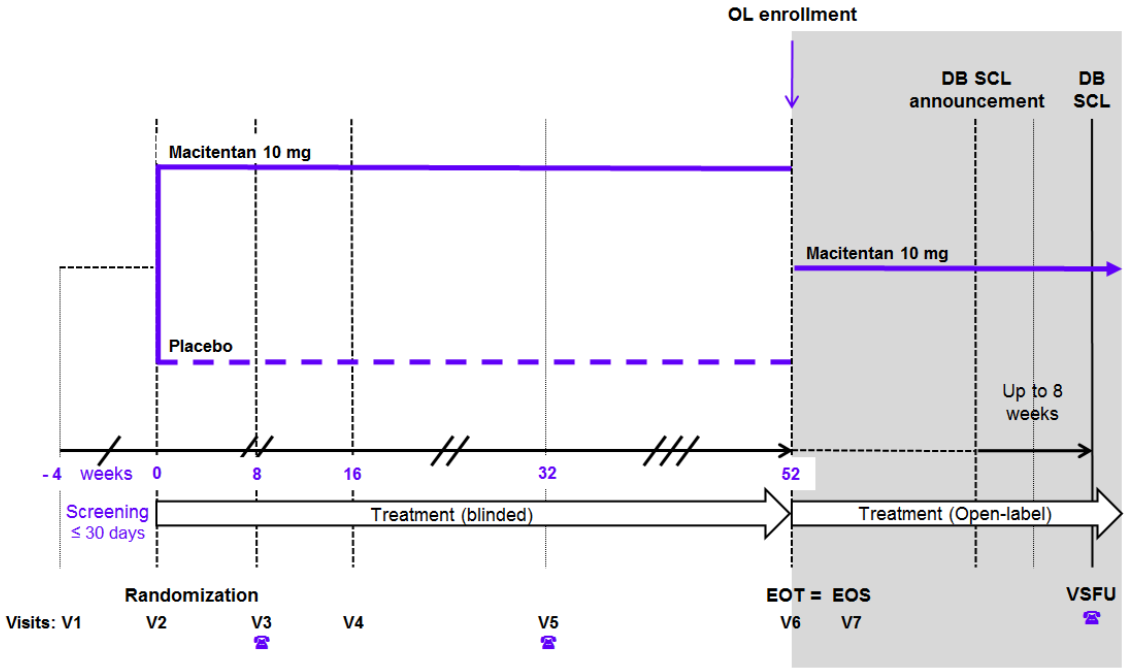

EOS = End-of-Study; EOT = End-of-Treatment; SCL = Study Closure; V = visit; VSFU = Vital Status Follow-Up; ☎= Telephone call (for visit).

**Figure 2: PTOP (example of discontinuation after Visit 4)**




EOS = End-of-Study; EOP = End-of-PTOP; EOT = End-of-Treatment; P = PTOP visit; PTOP = Post-treatment observation period; SCL = Study Closure; V = visit; VSFU = Vital Status Follow-Up; ☎= Telephone call (for visit).

**Figure 3: OL extension study (entry at Visit 6)**




DB = Double-blind; EOS = End-of-Study; EOT = End-of-Treatment; OL = Open-label (extension); SCL = Study Closure; V = visit; VSFU = Vital Status Follow-Up; ☎= Telephone call (for visit).

No S-FU will be performed in subjects entering the AC-055H302 RUBATO-OL extension study at individual subject's EOT/EOS for the AC-055H301 RUBATO study.

In instances where administration of AC-055H302 OL study drug does not immediately follow EOT in the AC-055H301 RUBATO study, the subject will enter the normal S-FU period until administration of the first dose of the OL study drug occurs.

## 1.3. Statistical Hypotheses for Trial Objectives

The primary efficacy endpoint is the change in peak $VO_2$ from baseline to Week 16. The primary null hypothesis to be tested is that macitentan 10 mg has no impact on exercise capacity, as assessed by the mean change in peak $VO_2$ from baseline to Week 16 compared to placebo. Rejection of the null hypothesis confirms the effects of macitentan 10 mg on exercise capacity; see Section 5.2.3.1 for further details.

The primary null hypothesis will be tested at a two-sided 1% alpha level, and if rejected the study will be declared to show 'conclusive' evidence of efficacy on the primary efficacy endpoint. If the primary null hypothesis can only be rejected at a two-sided 5% alpha level, the study will be declared 'positive'.

Conditional to conclusive or positive outcome of the primary efficacy endpoint at final analysis, the secondary efficacy endpoints will be formally evaluated at the same significance level as the primary efficacy endpoint according to the testing hierarchy following the order of the endpoints as listed in Section 5.3.

Statistical hypotheses for secondary endpoints are also addressed in Sections 5.3.1.3 and 5.3.2.3.

## 1.4. Sample Size Justification

The sample size is based on the primary efficacy endpoint, the 16-week change from baseline in peak $VO_2$. The aim of the trial is to reject the statistical hypothesis of no treatment difference.

Assuming that the macitentan versus placebo difference is not worse than the bosentan versus placebo difference in TEMPO study, a 2.4 mL/kg/min difference of macitentan to placebo on the change in peak $VO_2$ is hypothesized. The variability of the treatment effect of macitentan vs placebo is assumed to be greater than that observed in the TEMPO study of bosentan versus placebo in subjects with Fontan-palliated circulation (Hebert 2013, Hebert 2014), namely a standard deviation of 4.0 mL/kg/min yielding a standardized effect size of 0.6.

It is further assumed that all randomized subjects are evaluable. The family-wise two-sided Type I error is set to 1% (ie, the level of significance), and the targeted statistical power is 80% (ie, Type II error β of 20%).

The sample size is determined under the assumption of normal distribution of the primary efficacy endpoint (based on a two-sample t-test with equal unknown variances).

Based on the above assumptions and with a 1:1 randomization ratio, a total of 134 subjects, i.e. 67 per treatment group, are required to establish superiority of macitentan 10 mg over placebo at two sided α=1%.

Alternative sample size scenarios for the primary endpoint are given in Table 3, i.e., for different effect sizes from 0.4 to 0.8, and standard deviation from 3.0 to 6.0 mL/kg/min and treatment difference of 2.4 mL/kg/min.

**Table 3: Sample size with treatment effect of 2.4 mL/kg/min and statistical power 1-β of 80% in a 1:1 randomization ratio by level of significance α and effect size (with selected sample size in bold face)**

| α | Standardized Effect size | SD | n per arm | Total N |
|---|---|---|---|---|
| 1% | 0.4 | 6.0 | 148 | 296 |
| | 0.5 | 4.8 | 96 | 192 |
| | **0.6*** | **4.0** | **67** | **134** |
| | 0.8 | 3.0 | 39 | 78 |
| 5% | 0.4 | 6.0 | 100 | 200 |
| | 0.5 | 4.8 | 64 | 128 |
| | 0.6 | 4.0 | 45 | 90 |
| | 0.8 | 3.0 | 26 | 52 |

Standard deviation (SD) with a treatment difference of 2.4 mL/kg/min; n: sample size per treatment group (macitentan 10 mg or placebo) with a 1:1 ratio; N: total number of randomized subjects.
*Selected assumptions and sample size.

All calculations were performed using the East software version 6.3.

Initially (up to protocol version 7), the study contained a blinded sample size re-estimation, which would allow for an increase in the pre-planned sample size of 134 subjects in case of higher than expected variability for the primary efficacy endpoint. As described in Section 1.2, the IA for unblinded sample size re-estimation has since replaced the blinded sample size re-estimation, and was introduced in protocol version 8.

## 1.5. Randomization and Blinding

### 1.5.1. Randomization

After having verified that the subject meets all inclusion criteria and none of the exclusion criteria, the investigator/delegate contacts the interactive response technology (IRT) system at Visit 2 to randomize the subject. The IRT assigns a randomization number to the subject and assigns the treatment kit number, which matches the treatment group assigned by the randomization list to the randomization number.

Subjects are assigned to the two treatment groups in a 1:1 ratio based on a permutated-block randomization stratified by the geographical region (America, Europe, Asia, Oceania) in order to balance treatment groups throughout the conduct of the trial.

The randomization list is generated by an independent Contract Research Organization (CRO), Almac Clinical Technologies, using SAS® v9.3 and details regarding requirements for the generation of the Subject Randomization List(s) for this study are highlighted in the Biostatistics Addendum for the Subject Randomization List.

### 1.5.2. Blinding

This study will be performed in a double-blind fashion. The investigator and study personnel, the subjects, the clinical research associates (CRAs), Sponsor personnel, and CRO personnel involved in the conduct of the study will remain blinded to the study treatment until study completion. Sponsor personnel responsible for clinical study supply distribution will need to be unblinded to ensure adequate supply of study treatment. These persons will be clearly identified, their unblinding will be documented in the trial master file and they will not take part in any clinical trial team (CTT) meetings after study set-up has been completed.

Until the time of unblinding for final data analysis, the randomization list is kept strictly confidential and accessible only to authorized persons, e.g., the ISAC/SSG, who are not involved in the conduct of the study.

During the conduct of the study (in June 2020), the unblinding documents (i.e., randomization and material lists) were transferred from Actelion legacy Global Quality Management to the Secure Data Office according to the Janssen process.

# 2. GENERAL ANALYSIS DEFINITIONS

## 2.1. Visit Windows

As subjects do not always adhere to the protocol visit schedule, the following rules are applied to assign actual visits (including unscheduled ones) to analysis visits. Listed below are the visit windows and the target days for each visit. The reference day is Study Day 1 (see Section 2.4 for definition). If a subject has 2 or more actual visits with an available value in the same visit window the visit closest to the target day will be used as the protocol visit for that visit window. For peak $VO_2$, if a subject has 2 or more actual visits in the same visit window, the visit with the valid value closest to the target day will be used as the protocol visit for that visit window. The other additional visit(s) will not be used in the summaries or analyses by visit, but they can be used for determination of other endpoints (e.g., time to event endpoints, marked laboratory abnormalities). If 2 actual visits are equidistant from the target day within a visit window, the later visit is used. Except otherwise specified, if more than one assessment falls on the same day then the worst assessment (i.e., lowest for peak $VO_2$) is used.

All assignments will be made in chronological order. Once a visit date is assigned to a visit window, it will no longer be used for a later time point except for the endpoint. Listed below (Table 4) are the visit windows and the target days for each visit defined in the protocol.

Assessments which fall outside of any planned window will not be included in summary tables but will be included into individual subject listings.

**Table 4: Visit Windows**

| Parameter | Scheduled Visit Number | Time Interval (label on output) | Time Interval (Day)* | Target Time Point (Day) |
|---|---|---|---|---|
| Efficacy parameters | 2 | Baseline | <=1 | 1 |
| | 4 | Week 16 | 99 to 168 | 113 |
| | 5 | Week 32** | 169 to 293 | 225 |
| | 6 | Week 52 | 294 to 420 | 365 |
| Premature End-of-Treatment remapping for safety parameters and Quality of Life | 2 | Baseline*** | <=1 | 1 |
| | 3 | Week 8** | 2 to 98 | 57 |
| | 4 | Week 16 | 99 to 168 | 113 |
| | 5 | Week 32** | 169 to 293 | 225 |
| | 6 | Week 52 | 294 to open end | 365 |

*Relative to Study Day 1; ** on-site visits performed until protocol version 7, then converted to remote visits (phone call); *** Safety assessments (e.g., laboratory, vital signs) on Day 1 are considered baseline. Safety signals (i.e., adverse events) occurring/starting on Day 1 are considered treatment-emergent while safety events (i.e., laboratory marked abnormalities, vital signs abnormalities) occurring/starting on Day 1 are not considered treatment-emergent.

### 2.1.1. Efficacy

For the efficacy visit windows, all the assessment after the treatment start date will be considered, unless otherwise specified.

For subjects who prematurely discontinue study treatment, the End-of-Treatment visit (if available) will be re-assigned to a scheduled visit in the respective time window as defined above (including any assessment after treatment start date).

### 2.1.2. Safety

For safety, only assessments measured during the treatment period (i.e., from study treatment start up to 30 days after study treatment discontinuation, limits included) will be considered at the scheduled visit up to Visit 6/Week 52.

For subjects who prematurely discontinue study treatment, the End-of-Treatment visit (if available) will be re-assigned to a scheduled visit in the respective time window as defined above (including any assessment from study treatment start up to 30 days after study treatment discontinuation, limits included). In addition, the following assessments are derived:

- "Last on-treatment assessment" will be derived as last assessment prior to or on End-of-Treatment + 1 day (limits included)
- "Last follow-up assessment" will be derived as last assessment between End-of-Treatment + 1 day (limit excluded) and End-of-Treatment + 35 days (limit included).

### 2.1.3. Quality of life

For the quality of life, all the assessment after the treatment start date are considered at the scheduled visit up to Visit 6/Week 52. For subjects who prematurely discontinue study treatment, the End-of-Treatment visit (if available) will be re-assigned to a scheduled visit in the respective time window as defined above (including any assessment after treatment start date).

## 2.2. Analysis Sets

The following analysis sets are described in protocol Section 10.1.

### 2.2.1. Enrolled Analysis Set

The Enrolled Analysis Set (ENR) includes all subjects who are screened and have a subject identification number.

### 2.2.2. Full Analysis Set

The Full Analysis Set (FAS) includes all subjects assigned (i.e., randomized) to a study treatment. In order to adhere to the intention-to-treat principle as much as possible:

- Subjects are evaluated according to the study treatment they have been assigned to (which may be different from the study treatment they have received);
- All available data are included.

The FAS is part of the population definition for Estimand for the main analysis of primary and secondary efficacy endpoints.

### 2.2.3. Per-Protocol Analysis Set

The Per-Protocol Set (PP) is defined as part of the population definition for Estimand for the sensitivity analyses of primary and secondary efficacy endpoints.

- Definition of Population for Estimand - primary efficacy endpoint (PP1)

This estimand population (PP1) will comprise all subjects who received actual study treatment assigned by IRT and who complied with the protocol sufficiently up to and including Week 16 for the primary efficacy endpoint (refer to Table 5 to identify those protocol deviations that lead to the exclusion from PP1).

- Definition of Population for Estimand (PP2 and PP3) - secondary efficacy endpoints - (PP2 and PP3)

This estimand population (PP2 or PP3) will comprise all subjects who received actual study treatment assigned by IRT and who complied with the protocol sufficiently up to and including the timepoint of evaluation for the secondary efficacy endpoints (refer to Table 5 to identify those protocol deviations that lead to the exclusion from PP2 and PP3 respectively).

The identification of protocol deviations leading to exclusion of subjects/data from PP1, PP2, and PP3 has been conducted by an independent team before database lock as documented in ATTACHMENT 3.

**Table 5: Protocol Deviations leading to exclusion of subjects / data from PP1/PP2/PP3**

| PD Condition | PD Identifier | Categorization of PD | Primary endpoint (Peak $VO_2$ at Week 16 -PP1 | Secondary endpoint (Peak $VO_2$ over Week 52) - PP2 | Secondary endpoint (mean count per minute of daily PA-Ac at Week 16) - PP3 |
|---|---|---|---|---|---|
| Unable to understand and comply with the instructions, and to physically perform the CPET. | PD_MM.140 | Entered but did not satisfy criteria | X | X | |
| Known severely reduced single ventricle ejection fraction (< 30%) | PD_MM.144 | Entered but did not satisfy criteria | X | X | X |
| Known Fontan circulation stenosis, affected area > 50% of the diameter | PD_MM.145 | Entered but did not satisfy criteria | X | X | X |
| Known valvular defects (severe atrioventricular [AV] valve regurgitation, outflow obstruction) | PD_MM.146 | Entered but did not satisfy criteria | X | X | X |
| Known pulmonary-venous pathway obstruction | PD_MM.147 | Entered but did not satisfy criteria | X | X | X |
| Clinical worsening leading to medical interventions including reoperation of Fontan circulation within 3 months of screening | PD_MM.149 | Entered but did not satisfy criteria | X | X | X |
| Unscheduled hospitalization due to deterioration of the Fontan palliated condition within 3 months of screening | PD_MM.150 | Entered but did not satisfy criteria | X | X | X |
| Signs and symptoms of HF requiring change in diuretic therapy within 3 months of screening | PD_MM.151 | Entered but did not satisfy criteria | X | X | X |
| Ventricular tachyarrhythmia or supraventricular tachyarrhythmia within 3 months of screening | PD_MM.152 | Entered but did not satisfy criteria | X | X | X |
| Plastic bronchitis/chyloptysis within 3 months of screening | PD_MM.156 | Entered but did not satisfy criteria | X | X | X |
| Signs and symptoms requiring the addition of a new class of cardiovascular medication (e.g., nitrates, alpha-blockers, or ERAs) within 3 months of screening | PD_MM.157 | Entered but did not satisfy criteria | X | X | X |
| History of syncope during exercise | PD_MM.158 | Entered but did not satisfy criteria | X | X | X |
| Symptomatic coronary artery disease at Screening | PD_MM.159 | Entered but did not satisfy criteria | X | X | X |
| Exercise training program for cardiopulmonary rehabilitation in the 3-month period prior to Screening | PD_MM.161 | Entered but did not satisfy criteria | X | X | X |
| Treatment with nitrates, alpha-blockers, or ERAs in the 3-month period prior to Screening | PD_MM.171 | Entered but did not satisfy criteria | X | X | X |
| Introduction or change of dose of PH specific drugs (other than ERAs) in the 3-month period prior to Screening | PD_MM.172 | Entered but did not satisfy criteria | X | X | X |
| Male or female < 12 years old | PD_PM.351 | Entered but did not satisfy criteria | X | X | X |
| Not a Fontan-palliated subject with either LT-TCPC, or EC-TCPC surgery > 1 year before Screening | PD_PM.352 | Entered but did not satisfy criteria | X | X | X |
| Subject not in NYHA FC II or III (assessed by the investigator using the Specific Activity Scale | PD_PM.353 | Entered but did not satisfy criteria | X | X | X |
| Peripheral oxygen saturation ($SpO_2$) of < 88% at rest at Screening | PD_PM.354 | Entered but did not satisfy criteria | X | X | X |
| Respiratory limitation with a breathing reserve of < 10% | PD_PM.355 | Entered but did not satisfy criteria | X | X | X |
| Iron deficiency defined as ferritin < 10 µg/L at Screening | PD_PM.356 | Entered but did not satisfy criteria | X | X | X |
| Body mass index (BMI) for age reference (z scored) equivalent to BMI > 35 $kg/m^2$ (for adults) at Screening | PD_PM.357 | Entered but did not satisfy criteria | X | X | X |
| Peak $VO_2$ < 15 mL/kg/min at Baseline | PD_PM.362 | Entered but did not satisfy criteria | X | X | X |
| Initiation of forbidden therapy up to the week 16 visit | PD_PM.324 | Received a disallowed concomitant treatment | X | X | X |
| Initiation of forbidden therapy after the week 16 visit | PD_PM.341 | Received a disallowed concomitant treatment | | X | |
| Study medication not stored correctly but subject continued to use it. | PD_MM.124 | Received wrong treatment or incorrect dose | X | X | X |

| | | | | | |
|---|---|---|---|---|---|
| **Note:** If the PD date is available, exclude from PP1 and PP3 if the PD date is prior to the Visit 4 CPET assessment, and from PP2 if the PD date is prior to the Visit 6 PA-Ac assessment | | | | | |
| Administration of incorrect study treatment (i.e. Subject did not get medication bottle assigned by IRT).<br><br>**Note:** This PD leads to exclusion if, after unblinding, the allocated bottle did not contain the planned study treatment (i.e.,if the PD_PM.333 has been classified as "Important"). Thereafter, if the PD leads to exclusion and the PD date is available (either date of first intake from an incorrect bottle of - if this data is not available, use the date when wrong bottle was dispensed), only exclude if occurred prior to corresponding efficacy assessment date. | PD_PM.333 | Received wrong treatment or incorrect dose | X | X | X |
| Study treatment compliance for the whole treatment period < 80% or > 120% | PD_PM.335 | Received wrong treatment or incorrect dose | | X | |
| Study treatment compliance up to the Week 16 visit < 80% or > 120% | PD_PM.336 | Received wrong treatment or incorrect dose | X | | X |
| Treatment interruption of > 4 consecutive weeks up to Week 16 | PD_PM.367 | Received wrong treatment or incorrect dose | X | X | X |
| Treatment interruption of > 4 consecutive weeks after Week 16 | PD_PM.368 | Received wrong treatment or incorrect dose | | X | |
| Informed consent not personally signed and dated by subject /legal representative | PD_MM.101 | Other | X | X | X |
| Non-justified treatment code was break before unblinding of the study | PD_MM.111 | Other | X | X | X |
| Subject has developed a condition that interferes with CPET assessment | PD_MM.112 | Other | X<br><br>**Note:** if the PD is linked to Visit 4 | X<br><br>**Note**: if the PD is linked to Visit 6 | |
| CPET at Visit 4 conducted under different conditions than at baseline | PD_MM.114 | Other | X | | |
| CPET at Visit 4 done, but not valid/questionable on review by central reading facility | PD_MM.115 | Other | X | | |
| CPET at Visit 6 conducted under different conditions than at baseline | PD_MM.121 | Other | | X | |
| CPET at Visit 6 done, but not valid/questionable on review by central reading facility | PD_MM.318 | Other | | X | |
| Subject does not have sufficient PA-Ac baseline data | PD_PM.303 | Other | | | X |
| Randomization CPET done, but not valid/questionable on review by central reading facility | PD_PM.304 | Other | X | X | |
| Subject does not have sufficient PA-Ac data at Visit 4 | PD_PM.313 | Other | | | X |
| Subject not Randomized but received study drug | PD_PP.245 | Other | X | X | X |
| CPET at Visit 4 not done | PD_PP.251 | Other | X | | |
| CPET at Visit 6 not done | PD_PP.253 | Other | | X | |
| Peak $VO_2$ measurements at Week 16 (Visit 4) not available | PD_PP.260 | Other | X | | |

### 2.2.4. Safety Analysis Set

The Safety Set (SS) includes all subjects who received at least one dose of study treatment. Subjects will be analyzed according to the study treatment they actually received.

The treatment received will be considered different from the treatment assigned at randomization (randomized treatment) only in the case of a dispensing error sustained throughout the entire double-blind study. Other dispensing errors will not qualify for a change from the randomized treatment group.

### 2.2.5. Pharmacokinetics Analysis Set

The Pharmacokinetic Analysis Set (PKS) includes all randomized and treated subjects, for whom a PK blood sample at trough has been collected and who do not deviate from the protocol in a way that might affect the evaluation of the trough concentrations.

A subject will be excluded from the PKS if he/she meets the following criteria/events for all visits (Week 4/Week 8 and Week 16/EOT):

- PD_PP.250 Subject has no PK assessment
- PD_PM.364 PK sampling taken after the morning dose of study treatment

### 2.2.6. Usage of the analysis sets

The main analyses of the primary, secondary and other efficacy endpoints will be performed on the FAS, as per population definitions for estimands in the corresponding sections.

Safety and pharmacokinetic analyses will be performed on the SS and PKS, respectively.

Listings will be prepared on the FAS, unless otherwise specified.

## 2.3. Definition of Subgroups

In order to assess the consistency of the treatment effect across different subject subgroups, analyses will be performed on the primary efficacy endpoint and the secondary efficacy endpoints for the subgroups described in Table 6 and Table 7.

**Table 6: Subgroups definition**

| Subgroup | Definition |
|---|---|
| Age Group | • Adolescents: 12-<18 yrs<br>• Adults: >= 18 yrs |
| Geographical region | • America<br>• Europe<br>• Asia<br>• Oceania |
| Region | • US<br>• Non-US |
| Ventricular dominance | • Left ventricle<br>• Right ventricle / Mixed |

Countries will be assigned to geographical region based on the Standard Country or Area Codes for Statistical Use (M49) standard (https://unstats.un.org/unsd/methodology/m49/).

After consultation with the Study Results Committee constituted of external clinicians with expertise in subjects with Functional Single Ventricle Heart Disease on 13-Aug-2021, the following subgroups have been added based on clinical consideration and blinded aggregated baseline characteristics at baseline, to assess the consistency of the treatment effect on the primary and secondary efficacy endpoints:

**Table 7: Additional subgroups definition**

| Subgroup | Definition |
|---|---|
| BMI at baseline | • Underweight/Normal<br>• Overweight/Obese (including obese class I, II, III) |
| $SpO_2$ at baseline | • < 92 %<br>• >= 92% |
| Peak respiratory exchange ratio at baseline | • < 1.1<br>• >= 1.1 |

## 2.4. Study Day and Relative Day

Study Day 1 or Day 1 refers to the start date of the first study treatment administration (see Section 2.5.2 for definition). All efficacy and safety assessments at all visits will be assigned a day relative to this date.

Study day or relative day for a visit is defined as:

- Visit date − (date of Study Day 1) +1, if visit date is ≥ date of Day 1
- Visit date − date of Day 1, if visit date < date of Day 1

There is no 'Day 0'.

## 2.5. Baseline, Relevant Time-points and Analyses Periods

### 2.5.1. Baseline

Except where otherwise specified, the baseline is defined as the last non-missing assessment obtained before or on the day of the start of study treatment. If unscheduled/re-test visits are performed on the day of study treatment start, the non-missing assessment of the last unscheduled/re-test visit on study treatment start date is considered as baseline.

If more than one assessment falls on the same day then the worst assessment (i.e., lowest for peak $VO_2$) is used, unless otherwise specified.

Missing values for peak $VO_2$ at baseline are not allowed by protocol, but in case a subject has been randomized with a non-missing baseline value that has subsequently been evaluated as "invalid" by the independent central reading facility, it will be imputed for the main analysis of the primary efficacy endpoint as detailed in Section 5.2.3.2. In case of re-screening, the value from the first screening attempt will not be used for determining the baseline.

For the physical activity measured by accelerometer, the 9-day time period is selected by taking the first (in chronological order) interval of 9 consecutive days of daily daytime data with at least 4 complete daily time periods, starting from the day after or on the Screening visit (or Re-Screening as appropriate) up to treatment start date (i.e. the 9-day time period should fall entirely before or on treatment start date). To be considered evaluable, physical activity should have been measured for at least 4 complete daily daytime periods (out of 9 consecutive days) at baseline. A complete day is defined as a record of at least 7 hours of daily daytime data.

### 2.5.2. Study treatment start date

This is the "Treatment start date" (defined as the date of the first dose of study treatment intake) from the first interval, in chronological order, recorded in the "Study Drug Log" eCRF module. If missing, the randomization date will be used.

### 2.5.3. Study treatment end date / End of Treatment (EOT)

This is the "Treatment end date" (defined as the date of the last dose of study treatment intake) from the last interval, in chronological order, recorded in the "Study Drug Log" eCRF module. It is also called End of Treatment or permanent study treatment discontinuation.

EOT is the date of the last dose of double-blind study treatment in AC-055H301 RUBATO study.

### 2.5.4. End of Study (EOS)

The EOS corresponds to the last visit in this AC-055H301 RUBATO study as described in Section 1.2 and directly available in the SDTM. The EOS visit date does not include the Vital Status Follow up.

## 2.6. Imputation Rules for Missing/Incomplete dates and times fields

All dates and times used in the analyses are supposed to be complete, apart from the types included in Table 8.

In the following, 'lower limit' and 'upper limit' refer to the minimum or maximum, respectively, of a possible date. For example, if the day is missing, the lowest limit is the first day of the given month and the upper limit is the last day of the given month. If the day and month are missing, the lower limit refers to the first day of the given year and the upper limit to the last day of the given year. The earliest and the latest of different dates refer to the first or last date, respectively, when ordered in sequence.

**Table 8: Imputation Rules for Missing/Incomplete dates and times fields**

| Type of date | Date is incomplete | Date is missing |
|---|---|---|
| **Adverse Events** | | |
| AE resolution date | The upper limit. | No replacement, the AE is considered as ongoing in the analysis. |
| AE onset date | If the end date of the AE is not before the study treatment start date and if the study treatment start date falls in the range of possible dates, the study treatment start date is used. In all the other cases, the lower limit is used. | The earliest between the resolution date and the study treatment start date. |

| Type of date | Date is incomplete | Date is missing |
|---|---|---|
| **Fontan surgery completion and Medical History** | | |
| Date of last Fontan surgery completion | Day missing: 15th of the month<br>Day and month missing: 30th of June | No replacement. |
| Medical history end date | The upper limit. | No replacement. |
| Medical history start date | The lower limit. | No replacement. |
| **Medications** | | |
| Previous/Concomitant medication end date | The upper limit unless the medication started prior to study treatment start (or start date is missing) and 'Ongoing at start of treatment?' is ticked 'No', and the upper limit is after the study treatment start (DB) and the study treatment start falls in the range of possible dates then it is replaced with treatment start date - 1. | If the medication started prior to study treatment start and 'Ongoing at start of treatment?' is ticked 'No', then it is considered a previous medication and it is imputed with treatment start date -1. |
| Previous/Concomitant medication start date | The lower limit except when:<br>not tagged as ongoing at start of treatment<br>AND<br>the therapy end date is not collected or with the upper limit after the study treatment start date or the end date is after the study treatment start date<br>AND<br>the study treatment start date falls in the range of possible dates<br>in which case it is the study treatment start date. | No replacement, the medication is considered to have started before the date of consent. |
| **For determining treatment duration and emergent periods** | | |

| Type of date | Date is incomplete | Date is missing |
|---|---|---|
| Treatment start date | - Day is missing: replaced by the day of the randomization date. In the case the day of the randomization is in the previous month, replace the day by the lower limit.<br>- Day and month are missing: replace entirely by the randomization date. | Replace entirely by the randomization date. |
| Treatment end date (EOT) | Use the earliest date among:<br>upper limit<br>EOS<br>date of death. | Use the earliest date among:<br>EOS<br>date of last contact<br>date of death. |
| EOS | The upper limit. | Raw extraction date. |

## 3. INTERIM ANALYSIS AND DATA MONITORING COMMITTEE REVIEW

This study implemented one IA for unblinded sample size re-estimation. There is no early stopping for efficacy or early stopping for futility at the IA. The full type I error rate of 1% (two-sided) is spent only at final analysis.

The IA was conducted by an ISAC/SSG for the IDMC to allow unblinded sample size re-estimation. The IA took place when the first 95 randomized subjects had completed their Week 16 Cardiopulmonary Exercise Testing (CPET) assessment or withdrew from the study without a Week 16 CPET assessment (this corresponds to a 16-week information fraction for the primary efficacy endpoint of approximately 70% of the pre-planned sample size of 134 subjects). The data cut off applied for the IA was 11-Feb-2020.

Detailed guidelines for interim decision-making are described in the SAP, dated 25 February 2020 (D-20.077) for the purpose of the Special Protocol Assessment (SPA) submission of protocol Version 8 to the US Food and Drug Administration (FDA). The SAP for SPA outlines the statistical methods and analyses to be applied for the primary efficacy endpoint for the sample size re-estimation and it was used as reference by the ISAC/SSG.

The ISAC/SGG made unblinded results available to the IDMC. The IDMC reviewed the interim results and made corresponding recommendations in line with the IDMC charter. The recommendation related to the IA was provided to the Sponsor Committee only after all approvals (both IRBs/IECs and Health Authorities) of the relevant protocol for subjects contributing to the IA had been received. Based on the IDMC meeting of 7 April 2020, the IDMC recommended to keep the target sample size at 134 subjects, without sample size adjustment. When the IDMC recommendation related to the IA was received (i.e., 23 June 2020), a total of 141 subjects were

randomized into the study. The recruitment was closed with no possibility of new screening activities. Those subjects in screening were continued to be evaluated for eligibility into the study if all criteria were met. A total of 142 subjects were randomized into the study.

Furthermore, periodic IDMC meetings are scheduled to review safety and efficacy data according to the IDMC charter. Separate IDMC SAPs (D-18.337 9 March 2020 for AC-055H301/RUBATO and D-20.079 9 March 2020 for AC-055H302/RUBATO-OL) are provided for data monitoring committee review. The IDMC SAPs describe in detail the statistical analyses and presentation of the data for the open session of the IDMC meetings. These outputs provide support to the ISAC/SSG to ensure the studies are conducted to high scientific and ethical standards.

# 4. SUBJECT INFORMATION

The number of subjects in the ENR will be presented overall while the number of subjects in each analysis set (i.e., FAS, PP1, PP2, PP3, SS, and PKS) will be summarized and listed by treatment group and overall.

Listings will be provided displaying the following information:

- Subject membership in the different analysis sets (PP1, PP2, PP3, SS, PKS). Subjects in the ENR and not randomized will not be displayed. This listing will be run on FAS.
- Reasons for exclusion from the population definitions for estimand (i.e., PP1, PP2, PP3).

Reasons for exclusion from PP1, PP2, and PP3 will be summarized by treatment group on the FAS. Reasons for exclusion from PKS will be covered in Section 8.

In addition, the distribution of subjects by geographical region, country, and site ID by treatment group and overall will be presented.

## 4.1. Demographics and Baseline Characteristics

For re-screened subjects, the data assessed during the last screening attempt (i.e., re-screening visit) will be included in the analysis of demographic and baseline characteristics. The data coming from the first screening attempt will be included, if not available at re-screening.

### 4.1.1. Demographics

Table 9 presents a list of the demographic variables that will be summarized by treatment group and overall for the FAS. Demographics will also be summarized by subgroups as defined in Section 2.3, using the FAS.

**Table 9: Demographic Variables**

| **Continuous Variables:** | **Summary Type** |
|---|---|
| Age (years) | Descriptive statistics (N, mean, standard deviation [SD], median and range [minimum and maximum], and IQ range). |
| Weight (kg) | |
| Height (cm) | |
| Body Mass Index (BMI) (kg/m$^2$) | |
| **Categorical Variables** | |
| Age (12– <18, 18– <30, 30–<40, 40-<65, >= 65) | |
| Sex (male, female) | |

| | |
|---|---|
| Race (, American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or other Pacific Islander White, Other, Not Applicable) | Frequency distribution with the number and percentage of subjects in each category. |
| Ethnicity (Hispanic or Latino, not Hispanic or Latino, Unknown) | |
| Geographical region (America, Europe, Asia, Oceania)[a] | |
| Region (US, Non-US) | |
| BMI ([underweight <18.5 kg/m$^2$, normal 18.5-<25 kg/m$^2$, overweight 25-<30 kg/m$^2$, obese class I 30-<35 kg/m$^2$, obese class II 35-<40 kg/m$^2$, obese class III ≥40 kg/m$^2$]) [b] | |

[a] Countries will be assigned to geographical region based on the Standard Country or Area Codes for Statistical Use (M49) standard.

[b] For adolescents (boys or girls) between 12 and 18 years, BMI-for-age reference equivalent to BMI cutoffs for adults are calculated as provided in Attachment 1.

Individual subject's listing will be provided on the FAS.

### 4.1.2. Baseline Disease Characteristics

Table 10 presents a list of the baseline disease characteristics that will be summarized by treatment group and overall for the FAS. Baseline characteristics will also be summarized by subgroups as defined in Section 2.3, using the FAS.

**Table 10: Baseline Disease Characteristics Variables**

| **Continuous Variables:** | **Summary Type** |
|---|---|
| $SpO_2$ (%) at baseline | Descriptive statistics (N, mean, standard deviation [SD], median and range [minimum and maximum], and IQ range). |
| NT-proBNP (pmol/L) at baseline | |
| Peak $VO_2$ (mL/kg/min) at baseline | |
| Mean count per minute of daily Physical Activity at baseline | |
| Daily mean activity time in min (sedentary, light, moderate, vigorous) | |
| Time since Fontan palliation completion (years) | |
| **Categorical Variables** | |
| NYHA FC | |
| PAH specific therapies concomitant at start of study treatment | |
| % of the daily mean total activity time spent in sedentary, light, moderate, or vigorous (including very vigorous for adults) at baseline | |
| Congenital heart defect leading to Fontan-palliation | Frequency distribution with the number and percentage of subjects in each category. |
| Dominant ventricular morphology (left and right/mixed) | |
| Staged reconstructive surgeries performed before completion of the Fontan circulation | |
| Type of primary TCPC Fontan completion | |
| Type of secondary TCPC Fontan completion (TCPC=total cavopulmonary connection) | |
| Fenestration Status | |

Below the details for the derivation of the variables in the table above, as needed.

**Fontan History and Fontan completion / TCPC surgery**

- Congenital heart defect leading to Fontan-palliation (i.e. information collected in the eCRF "Fontan history" module, "Cardiac malformation" section) and associated dominant ventricular morphology (i.e., left and right/mixed),

- Staged reconstructive surgeries performed before completion of the Fontan circulation (Norwood, Glenn, etc.) [i.e. information collected in the eCRF module "Fontan history", "Single Ventricle: Surgery" section],
- Type of current total cavopulmonary connection TCPC (LT-TCPC, EC-TCPC) and whether TCPC is primary or secondary [i.e., information collected in the eCRF form for "Fontan completion / TCPC surgery"],
- Fenestration status (no never had; no initially yes, now closed; yes is open; yes initially no, now open) [i.e., information collected in the eCRF form for "Fontan completion / TCPC surgery"],
- Time since Fontan palliation completion (years), defined as the time from the last Fontan surgery to screening (or re-screening as appropriate). It is calculated by firstly taking the difference between the date of screening and the date of the last Fontan surgery completion and secondly by dividing the result by 365.25.

**PAH specific therapies concomitant at start of study treatment**

- PAH specific medications concomitant at start of study treatment (see Section 4.7.2 for definition). Pulmonary hypertension specific medications include: sildenafil, sildenafil citrate, tadalafil, vardenafil, udenafil, iloprost, beraprost, epoprostenol, treprostinil, selexipag, bosentan, ambrisentan, riociguat (as qualified by the corresponding preferred term).

**Physical Activity measured by Accelerometer (PA-Ac) at baseline**

- Mean count per minute of daily PA-Ac at baseline,
- Daily mean time spent in sedentary, light, moderate, or vigorous (including very vigorous for adults) at baseline.

% of the daily mean total activity time spent in sedentary, light, moderate, or vigorous (including very vigorous for adults) at baseline. The accelerometry central reading facility will read daily counts/min and will analyze duration of daily activity and time (minutes) spent in sedentary, light, moderate, and vigorous physical activity based on the (Freedson 1998) and (Evenson 2008) cut-points for adults and adolescents, respectively. These data will be transferred to the Sponsor.

Mean counts per minute of daily PA-Ac are calculated by dividing the sum of all activity counts collected during wear time for a complete day by the number of minutes of wear time in that day across all complete days (at least 4 days out of 9 days). A complete day is defined as a record of at least 7 hours of daily daytime data i.e. a complete day if "WearMinutes" >=7*60, (after excluding the periods when the device was apparently not worn). In summary, to compute the mean counts per minute of daily PA-Ac, it is necessary to identify all complete days; if less than 4 complete days out of 9 are identified then the baseline is missing, otherwise, if at least 4 complete days out of 9 are identified, then calculate the baseline on the complete days only.

Individual subject's listings will be provided on the FAS.

## 4.2. Medical history

Includes relevant medical history / current medical conditions/findings based on investigator's judgment present before and/or at the time of signing informed consent as collected during the screening visit on the "Medical history" eCRF module.

The original terms used by the investigators are assigned preferred terms for classification and tabulation using the latest implemented version of Medical Dictionary for Regulatory Activities (MedDRA).

Medical history is summarized, by treatment group and overall, displaying counts and percentages of subjects having been diagnosed with at least one disease. Counts and percentages of subjects having been diagnosed with at least one disease are presented by system organ class (SOC) and PT within each SOC as well as by PT. The summary tables are presented in descending order according to the incidence in the macitentan treatment group (e.g., SOC and PT within each SOC with the highest number of occurrences appears first). Equal frequency of different SOC/PTs is sorted in alphabetical order of the SOC/PT. Subjects with two or more occurrences of the same disease (as qualified by the same PT) are counted only once.

The summary tables will be provided on the FAS.

The medical history is also reported in a subject listing on the FAS.

## 4.3. Disposition Information

The following definitions are relevant to provide the disposition information:

### 4.3.1. Screened Subjects and Screening Failures

It is permitted to re-screen subjects once, if the reason for non-eligibility was transient (e.g., abnormal laboratory test, insufficient wash-out period of a forbidden medication). All screening assessments should then be repeated at the time of re-screening.

If a subject is screened twice and not randomized, the reason associated with the second screening attempt is considered in this analysis. If a subject is screened twice and then randomized, the first screening attempt is not considered in the evaluation of screening failures.

A subject is considered a screening failure if he/she signed the informed consent (i.e., available date of informed consent) but was eventually not randomized as identified from the answer to the eCRF question “Was the subject randomized?”. If the answer to this question is “No”, the reason provided in the eCRF (‘Subject withdrew consent’, ‘Not eligible as per inclusion/exclusion criteria’, ‘Lost to follow-up’ or ‘Other’) is used as the primary reason for screening failure.

The number and percentage of subjects in the following disposition categories will be summarized overall on the ENR, to display the following information:

- Screened subjects (i.e., ENR analysis set)
- Screen failures and reasons for discontinuation during screening.

A subject listing (including the inclusion criteria not met or the exclusion criteria met as collected in the “Eligibility Criteria” eCRF module) will be also provided for screening failures.

### 4.3.2. Study and Study Treatment Disposition

The number and percentage of subjects in the following disposition categories will be summarized by treatment group and overall on the FAS, to display the following information:

- Subjects randomized
- Subjects receiving study treatment
- Subjects completing the double-blind study treatment up to Week 52
- Subjects completing the study participation
- Subjects who terminated the study prematurely and reasons for premature termination of the study (see Section 4.3.2.2)
- Subjects who entered the PTOP (i.e., subjects with PTOP or End-of-post-treatment observation period [EOP] visit performed)

The number and percentage of subjects in the following disposition categories will be summarized by treatment group and overall on the SS, to display the following information:

- Subjects who discontinued the study treatment prematurely and reasons for premature discontinuation of study treatment (see Section 4.3.2.1).

#### 4.3.2.1. Study Treatment Discontinuation

Premature study treatment discontinuation (before Week 52) are collected in the "Study Drug Log" eCRF module and identified as those with a treatment end date and associated reason ('What was the reason for treatment end?') answered 'Premature Discontinuation'. Reasons for study treatment discontinuation are collected in the "Premature Discontinuation of Study Treatment' eCRF module. The full list of reasons is the following:

- Death
- Lost to follow-up
- Pre-specified study treatment discontinuation criteria
- Subject decision (AE, Lack of Efficacy, No reason provided, Other [medical/non-medical reasons])
- Physician decision (AE, Other [medical/non-medical reasons])
- Sponsor decision (Study termination, Other).

The distribution of the time to premature study treatment discontinuation will be displayed with Kaplan-Meier curves for the SS. Subjects who terminate treatment prematurely at any time will be considered an 'Event' and their date of study treatment discontinuation will be used in the time to event calculation. Subjects who complete the study treatment will be censored and the date of end of treatment will serve as the time of censoring.

#### 4.3.2.2. Study Discontinuation

Subjects who withdrew from the study are those with any withdrawal reason entered in the "Study Discontinuation" eCRF module. Possible reasons for study withdrawal are as follows:

- Death
- Lost to follow-up
- Subject decision/withdrawal of consent (AE, Lack of Efficacy, No reason provided, Other [medical/non-medical reasons])
- Physician decision (AE, Lack of efficacy, Other [medical/non-medical reasons])
- Sponsor decision (Study termination, Other).

A subject who completed the study participation would have no data entered in the eCRF form "Study Discontinuation" The EOS visit date is entered in the disposition SDTM domain as appropriate.

The distribution of the time to study discontinuation will be displayed with Kaplan-Meier curves for the FAS. Subjects who prematurely discontinue from the study at any time will be considered an 'Event' and their date of study discontinuation will be used in the time to event calculation. Subjects who complete the study as per protocol will be censored and the date of end of study will serve as the time of censoring.

Listings of subjects will be provided for the following categories:

- Randomization scheme and codes, including the stratification factor
- Subjects who discontinued study treatment prematurely and related reason(s)
- Subjects who terminated study prematurely and related reason(s)
- Subjects who were unblinded during the study.

## 4.4. Treatment Compliance

Study treatment compliance (%) is based on the study drug dispensing and accountability data recorded in the eCRF "Study Drug Dispensing & Accountability" module.

Compliance assessed at each visit by site personnel is calculated and entered in the eCRF and is not (re-) calculated. The reasons for non-compliance since last visit are also collected.

### 4.4.1. Overall Compliance (up to EOT)

Overall compliance (up to EOT) = [(total number of tablets dispensed at any time – total number of tablets returned )/ Total number of tablets that should have been taken during the treatment period] × 100.

With the introduction of the COVID-19 protocol appendix, on-site visits could be performed remotely or delayed, and while study treatment may have been dispensed via direct-to-patient shipments, subjects were asked to return empty bottles/unused tablets only at their next on-site visit. For the compliance up to EOT, irrespective of whether a bottle is returned before, at, or after

EOT, corresponding tablets are considered as having been taken prior to EOT. The overall compliance will not be calculated for those subjects with any dispensed bottles that were never returned.

### 4.4.2. Compliance (up to the Week 16)

Compliance (up to the week 16 Visit) = [(total number of tablets dispensed at any time from Visit 2 to the day before the Visit 4/Week 16 – total number of tablets returned after Visit 2 and up to the Visit 4/Week 16)/ Total number of tablets that should have been taken up to the Visit 4/Week 16] × 100.

For the same reason as above, the compliance up to Week 16 will be calculated if all bottles dispensed to a subject between Visit 2 and the day before Visit 4/Week 16 are returned up to Visit 4.

Analysis of study treatment compliance data will be performed on the SS, by treatment group. Compliance with study treatment will be displayed for two periods, from treatment start to Week 16 and from treatment start to EOT (i.e., for the whole treatment period).

Treatment compliance categorized as < 80%, 80–120%, and > 120% will be summarized on the SS for both periods.

Drug accountability data are provided in a subject listing on the SS.

## 4.5. Extent of Exposure

### 4.5.1. Study Treatment Interruptions

Analysis of study treatment exposure will be performed on the SS.

A subject is considered to have had a study treatment interruption if the reason for treatment end is either ‘Temporarily interrupted due to an AE’ or ‘Temporarily interrupted not due to an AE’ (“Study Drug Log” eCRF form).

### 4.5.2. Study Treatment Duration and Exposure

The study treatment duration (in weeks) is defined as the time interval between study treatment start date (see Section 2.5.2) and treatment end date (see Section 2.5.3) inclusive, regardless of treatment interruptions, calculated as:

$$\frac{treatment\ end\ date - treatment\ start\ date + 1}{7}$$

Study treatment exposure (in weeks) by excluding treatment interruptions is derived as the [study treatment duration (days) - the sum of all treatment interruptions (days)]/7.

For the calculation of study treatment exposure: if the treatment end day is missing in case of a study treatment interruption, it is replaced by day 15$^{th}$ of the respective month. In the case that day

is before the treatment start of the same interval as reported in the "Study Drug Log" eCRF form, it is replaced by the treatment start day of that interval.

Descriptive statistics for duration of study treatment and study treatment exposure (N, mean, standard deviation [SD], median and range [minimum and maximum], and IQ range) will be presented by treatment group.

In addition, the cumulative distribution of treatment duration by different class intervals (i.e., at least 4 weeks, at least 8 weeks, at least 16 weeks, at least 24 weeks, at least 32 weeks, and at least 52 weeks) will be summarized to show counts and percentages of subjects in each class interval.

Study treatment duration will be summarized in the following duration categories: <4 weeks, 4-<8 weeks, 8-<12 weeks, 12-<16 weeks, 16-<24 weeks, 24-<32 weeks , 32-<52 weeks, ≥52 weeks by treatment group and presented graphically in a histogram.

Subjects-year of treatment are calculated as [sum of study treatment duration (in days) for each subject/365.25]. "Total study treatment duration (subjects-year)" will be displayed by treatment group.

Exposure data will be provided in a subject listing on the SS.

## 4.6. Protocol Deviations

All PDs have been categorized into the following category levels prior to database lock by an independent team (as documented in ATTACHMENT 3):

- Developed withdrawal/interruption criteria but not withdrawn/interrupted
- Entered but did not satisfy criteria
- Received a disallowed concomitant treatment
- Received wrong treatment or incorrect dose
- Other

The analyses of PDs will be based on the FAS.

Subjects with important (i.e., major) PDs will be summarized by displaying the counts and percentages of subjects with at least an important PD by treatment group and overall. The summary table will be sorted by category. A separate similar summary of important PDs will be provided by geographical region and site.

A summary of subjects who did not meet inclusion criteria or who met exclusion criteria will be presented by treatment group and overall on the FAS.

All reported PDs will be described in a subject listing. Important PDs will be flagged accordingly.

A listing of subjects who received incorrect study treatment (if applicable) will be provided as well.

## 4.7. Prior and Concomitant Medications

All therapies as collected in the "Previous/Concomitant Medication" eCRF module.

The original terms used by the investigators to describe therapies are assigned preferred terms for classification and tabulation using the latest version of the World Health Organization (WHO) Drug code and Anatomic Therapeutic Chemical (ATC) class code dictionaries.

### 4.7.1. Prior Medications

Previous therapy is defined as any therapy with an end date on or prior to the date of informed consent.

If the end date is missing and the start date is prior to the date of informed consent, then the therapy is considered as previous if the checkbox for the questions 'Ongoing at start of treatment?' and "Ongoing at End of Study?" are ticked '**No**' by the investigator in the eCRF.

### 4.7.2. Study treatment concomitant medications

Study treatment concomitant therapy is any treatment that is either ongoing at the start of study treatment or is initiated during the double-blind treatment period, identified as:

- 'Ongoing at start of treatment? = **Yes**' ticked by the investigator in the eCRF,
  *or*
- Start date before treatment start date and end date on or after treatment start date,
  *or*
- Start date before treatment start date and end date missing with 'Ongoing at start of treatment?' ≠ **No**',
  *or*
- Start date on or after treatment start date and before study treatment end.

Concomitant therapies at start of treatment is any treatment that is either ongoing at the start of study treatment or is initiated at the start date of study treatment, identified as:

- 'Ongoing at start of treatment? = **Yes**' ticked by the investigator in the eCRF,
  *or*
- Start date before or on treatment start date and end date on or after treatment start date,
  *or*
- Start date before or on treatment start date and end date missing with 'Ongoing at start of treatment?' ≠ **No**'.

Counts and percentages of subjects having taken at least a study-treatment concomitant therapy will be presented, separately, by ATC class and PT within each ATC class as well as by PT. The summary table will present therapies in descending order according to the incidence in the macitentan treatment group (e.g., ATC and PT within each ATC with the highest number of occurrences appear first). Equal frequency of different ATC/PTs will be sorted in alphabetical

order of the ATC/PT. Subjects who took more than once the same therapy (as qualified by the same PT) will be counted only once.

Previous therapies will be summarized by ATC class and PT within each ATC class.

Individual subject's listing will be provided. Previous and study treatment concomitant therapies are flagged accordingly.

# 5. EFFICACY

The hierarchical testing will be applied to the primary and secondary efficacy endpoints as represented in Figure 4 below.

Conditional to conclusive or positive outcome of the primary efficacy endpoint results at final analysis (i.e., if the primary endpoint is statistically significant at a significance level α=1% or 5% two-sided), the secondary efficacy endpoints will be formally evaluated at the same significance level as the primary efficacy endpoint according to the testing hierarchy following the order of the endpoints as listed in Section 5.3. For example, if the test of the first secondary efficacy endpoint is statistically significant at 1% two-sided, then the test of the other secondary endpoint in the hierarchy can be formally performed at the same significance level. The formal testing process will stop when a test is not statistically significant at 1% two-sided level.

The stratification factor (i.e. geographical region) coming from the randomization assignment will be included in the analysis/model of the efficacy endpoints (when appropriate).

The number of missing data for primary and secondary efficacy endpoints will be monitored on a blinded fashion during the dry run review process. Additional sensitivity analyses may be added and justified, if deemed appropriate.

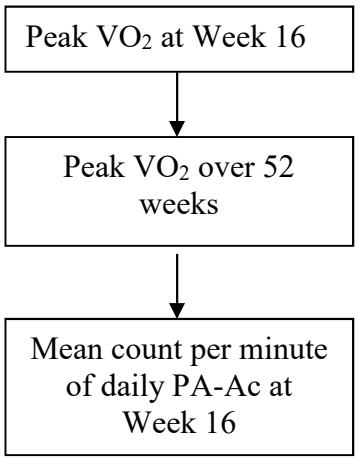
## 5.1. Analysis Specifications

### 5.1.1. Level of Significance

The overall type I error rate is 1% two-sided, fully spent at final analysis.

The study will be declared to show conclusive evidence of efficacy if the test of the primary efficacy endpoint is statistically significant at a significance level α = 1% two-sided. Otherwise, if the test is statistically significant at a significance level α = 5% two-sided, the study will be declared to show 'positive' evidence of efficacy. Secondary efficacy endpoints will be analyzed at the same significance level as the primary efficacy endpoint using a hierarchical testing procedure following the order of the endpoints as listed in Figure 4.

**Figure 4: Testing strategy for primary and secondary efficacy endpoint at final analysis (2-sided type I error rate, α=1% or 5%)**




To control the type I error against the unblinded sample size re-estimation at the interim timepoint, the final analysis of the primary and secondary efficacy endpoints will be conducted using the inverse normal combination method with pre-specified weights to combine first and second stage p-values (Lehmacher 1999).

Table 11 presents an overview of the set of subjects in each stage and the pre-defined weights for the inverse normal method at each stage used at the final analysis of primary and secondary efficacy endpoints. The pre-specified weights used for the main analysis are proportional to the pre-planned sample size of 134 subjects.

**Table 11: Significance level, set of subjects and weights for inverse normal combination method at final analysis**

| Endpoint | Analysis time point | Significance level | | Inverse normal combination method | | | |
|---|---|---|---|---|---|---|---|
| | | $Z_{WINC,final}$ | α | Stage 1 | | Stage 2 | |
| | | | | Subjs | Weight | Subjs | Weight |
| Primary endpoint: Peak $VO_2$ at Week 16 | Final | $c_{final}$=2.576* | 1% 2-sided* | $n_1 = 95$ | $\sqrt{t_1}$ with $t_1 := \frac{n_1}{n_p}$ $t_1 = \frac{95}{134}$ $\sqrt{t_1} = 0.841994789$ | Remaining $n_2 = n^* - n_1 = 142 - 95 = 47$ | $\sqrt{1-t_1} := \sqrt{1-\frac{95}{134}} = 0.539485659$ |
| Secondary endpoint: Peak $VO_2$ over 52 weeks | | | | | | | |
| Secondary endpoint: mean count per minute of daily PA-Ac at Week 16 | | | | | | | |

$Z_{WINC,final}$: Weighted inverse-normal combination (WINC) test statistic at final analysis. $c_{final}$= 2.575829

$n_1$:= 95 Actual number of subjects who have completed Week 16 CPET assessment (or withdrew from the study without a Week 16 CPET assessment) before the data cutoff (i.e. 11-Feb-2020) for the interim analysis (i.e., analysis time point 1).

$t_1 = \frac{n_1}{n_p} = \frac{95}{134}$: 16-week information fraction for the evaluation of the primary efficacy endpoint with $n_p = 134\ subjects$ (pre − planned total sample size)

$n_2$: $n^* - n_1$ = 142 - 95 = 47 subjects where $n^*$ is the actual total sample size after the IA results.

* The secondary efficacy endpoints will be formally evaluated at the same significance level as the primary efficacy endpoint according to the testing hierarchy, and conditional to conclusive/positive outcome of the primary efficacy endpoint results at final analysis.

### 5.1.2. Data Handling Rules

For both the primary and secondary endpoints, the corresponding estimand section describes data handling rules.

Baseline definitions are provided in Section 2.5.1.

## 5.2. Primary Efficacy Endpoint(s)

### 5.2.1. Definition

The primary efficacy endpoint is the change in peak $VO_2$ (mL/kg/min) from baseline to Week 16, defined as:

peak $VO_2$ (mL/kg/min) at Week 16 - peak $VO_2$ (mL/kg/min) at baseline

Peak $VO_2$ is measured during CPET. The values used for peak $VO_2$ at baseline and at Week 16 will be included based on a blinded data evaluation by the independent central reading facility, as reported in the eCRF.

For Week 16, the assessments (including the unscheduled ones) will be mapped to the Week 16 time window based on the Study Day (see Section 2.1). Any assessment in the Week 16 time window, irrespective of premature study treatment discontinuation before Week 16, will be included in the main analysis based on a 'treatment-policy' estimand strategy.

### 5.2.2. Estimand

The primary objective is to assess the effect of macitentan 10 mg on exercise capacity (as assessed by peak $VO_2$) in comparison with placebo in Fontan-palliated subjects.

Hence the primary estimand is defined as:

**A. Population:** All randomized Fontan-palliated adult and adolescent subjects defined through protocol inclusion/exclusion criteria (ie, FAS).

**B. Variable:** peak $VO_2$ as change from baseline to Week 16.

**C. Intercurrent events:**

- Death occurring before Week 16.

  A composite strategy will be applied: in case of death prior to Week 16, the change from baseline to Week 16 will be replaced with the lowest observed change in peak $VO_2$ from baseline to Week 16 across both treatment groups (i.e., worst-case imputation).

- Non-missing peak $VO_2$ value at Week 16 assessed as 'invalid' by the central reading facility.

  A composite strategy will be applied: in case an assessment at Week 16 has been evaluated as "invalid" by the independent central reading facility, it will not be included in the analysis and the change from baseline to Week 16 will be imputed with the 25$^{th}$ percentile of the observed

changes in peak $VO_2$ from baseline to Week 16 (across both treatment groups). If this leads to an improvement (a positive change), a change of 0 (no change) will be imputed.

- Premature study treatment discontinuation before Week 16

Premature study treatment discontinuation prior to Week 16 will be addressed by a "treatment-policy" strategy by including any assessment in the Week 16 time window, irrespective of premature study treatment discontinuation, in the main analysis.

**D. Population-level summary measure:** difference in mean change in peak $VO_2$ from baseline to Week 16 between macitentan 10 mg and placebo, as estimated from the ANCOVA model.

### 5.2.3. Analysis Methods

#### 5.2.3.1. Hypotheses and statistical model

The statistical hypotheses for the primary endpoint are formulated in terms of the arithmetic means (M) of change in peak $VO_2$ in subjects treated with either placebo or macitentan 10 mg from baseline to Week 16.

$H_0$: $M_{\text{macitentan 10 mg}} = M_{\text{placebo}}$ versus

$H_A$: $M_{\text{macitentan 10 mg}} \neq M_{\text{placebo}}$

The null hypothesis will be tested on the primary variable at a two-sided Type I error rate of 1% for conclusiveness based on the inverse normal combination method.

Assuming normal distribution of the primary variable, the ANCOVA model on the change in peak $VO_2$ will be applied to derive the first-stage and second-stage p-values for the final analysis (see Section 5.2.3.3). Model covariates will include randomized treatment group, stratification factor (ie, geographical region) and peak $VO_2$ at baseline.

It is expected to have a negligible incorrect stratification assigned during the randomization therefore the analysis is based on the stratification factors coming from the randomization assignment.

Assumptions of normality will be assessed with graphical diagnostics tools for the fit (i.e., fitted normal distribution function displayed on cumulative distribution function of the change from baseline to Week 16 [after applying imputation rules for handling of intercurrent events or missing data] by treatment group).

Assumptions of normality of residuals and homogeneity of variance (assessment of ANCOVA model assumptions) will be investigated and graphically presented (eg, Q-Q and residual plots).

#### 5.2.3.2. Handling of missing data for the main analysis

Missing values for peak $VO_2$ at baseline are not allowed by protocol, but in case a subject has been randomized with a non-missing baseline value that has subsequently been evaluated as "invalid" by the independent central reading facility, it will be imputed with the median peak $VO_2$ value observed at baseline from the remaining subjects (across treatment groups).

During the conduct of the study, all efforts will be made to avoid missing values in peak $VO_2$ at Week 16. If a subject discontinues treatment prematurely, a CPET will be conducted within 7 days of EOT. Furthermore, if treatment is discontinued prior to Visit 4/Week 16, and the subject enters the PTOP, it is requested that the subject returns for a CPET assessment at Week 16 (in addition to the EOT assessment).

Intercurrent events will be handled as defined in Section 5.2.2. Any missing values which may occur in peak $VO_2$ at Week 16 will be handled as in the table below.

| Reason | Strategy in handling the change from baseline to Week 16 |
|---|---|
| Other* | Imputed with the 25$^{th}$ percentile of the observed changes in peak $VO_2$ from baseline to Week 16 (across both treatment groups). If this leads to an improvement (a positive change), a change of 0 (no change) will be imputed. |

*Other could include: premature study treatment discontinuation because of an adverse event, lack of efficacy, or lost-to follow up and without a CPET assessment at Week 16.
CPET=Cardiopulmonary Exercise Testing, $VO_2$=Oxygen uptake/consumption.

#### 5.2.3.3. Main analysis

The main analysis of the primary efficacy endpoint will be conducted using the inverse normal combination method with pre-specified weights to combine first and second stage p-values (see details in Section 3.2.3 of the SAP for SPA).

The separate first-stage analysis will be based on the first-stage set of subjects contributing to the IA. The first-stage p-value from the analysis of covariance (ANCOVA) model including randomized treatment, stratification factor (i.e., geographical region), and peak $VO_2$ at baseline as covariates in the model will be used to construct the first stage p-value as used in the final analysis in the weighted inverse-normal combination (WINC) test statistic at final analysis.

The separate second-stage analysis will be based on the remaining number of subjects required to reach the total sample size based on the IA results. Due to the population-wise splitting, no subjects contributing to the separate first-stage test statistics are included into the separate second-stage analysis. The second-stage p-values will be calculated in the same way as the first-stage p-value.

The primary and secondary endpoints will be tested by combining p-values from the separate first and second stage analyses using the inverse-normal combination function with prefixed weights (i.e., see Table 11 in Section 5.1.1).

Final interpretation of the results will then be based on the final adjusted p-value, median unbiased estimator for the overall treatment effect and corresponding RCIs obtained by ADDPLAN™ 6.1 (ADDPLAN Inc. an Aptiv Solutions Co. 2014). Practically, the analysis will be conducted using ADDPLAN BASE module, “Adaptive Analysis” menu, “Test of means” functionality with the computation options specified in Section 3.2.3 of the SAP for SPA.

The number of subjects with intercurrent events or missing data will be presented by treatment group. Absolute values at baseline and at Week 16 as well as absolute change in peak $VO_2$ from

baseline to Week 16 (after handling of intercurrent events or applying imputation rules for missing data) will also be summarized using descriptive statistics by treatment group.

To graphically display changes in peak $VO_2$, results are visualized by means of a scatter plot of peak $VO_2$ values at baseline versus Week 16 (after handling of intercurrent events or applying imputation rules for missing data) for each treatment group. Subjects with imputed values for peak $VO_2$ are highlighted with different colors/symbols in the graph. The scatter plot is displayed with the reference line y = x.

#### 5.2.3.4. Supportive/sensitivity analyses

Absolute values at baseline and at Week 16 as well as absolute change in peak $VO_2$ from baseline to Week 16 will also be summarized using descriptive statistics, for each supportive and sensitivity analysis where applicable.

i. Additional estimands

Additional estimands will be assessed as defined in the tables below. The same ANCOVA model as the one described for the primary estimand (see Section 5.2.2) will be applied, based on the population wise splitting as applied for the main analysis.

**Table a Additional Estimand 1 (PP1)**

| | |
|---|---|
| **A.** | **Population:** All subjects who received study treatment as assigned by IRT and who complied with the protocol sufficiently up to and including the Week 16 for the primary efficacy endpoint (ie, PP1). Protocol deviations are defined in Section 2.2.3. |
| **B.** | **Variable: s**ame as for primary estimand (see Section 5.2.2) |
| **C.** | **Intercurrent events and strategies**: same as primary estimand (see Section 5.2.2) |
| **D.** | **Population-level summary measure:**. same as the ANCOVA model of the primary estimand (see Section 5.2.2) |

ANCOVA=Analysis of covariance, IRT=Interactive Response Technology, =PP1=Estimand Population for primary efficacy endpoint, $VO_2$ = Oxygen uptake/consumption.

**Table b Additional Estimand 2 (excluding subjects with invalid peak $VO_2$ at baseline)**

| | |
|---|---|
| **A.** | **Population:** same as primary estimand (FAS) but excluding subjects with invalid peak $VO_2$ at baseline (see Section 5.2.2). |
| **B.** | **Variable: s**ame as for primary estimand (see Section 5.2.2). |
| **C.** | **Intercurrent events and strategies**: same as primary estimand (see Section 5.2.2). |
| **D.** | **Population-level summary measure:**. same as the ANCOVA model of the primary estimand (see Section 5.2.2). |

ANCOVA=Analysis of covariance, FAS=Full Analysis Set, $VO_2$ = Oxygen uptake/consumption.

**Table c Additional Estimand 3 ("while on treatment")**

| | |
|---|---|
| **A.** | **Population:** same as primary estimand (FAS) |
| **B.** | **Variable: s**ame as for primary estimand (see Section 5.2.2) except that the assessments in the Week 16 time window will only be included in the analysis if they are performed while 'on-treatment', defined as up to 7 days after last dose of study treatment. |
| **C.** | **Intercurrent events and strategies**: same as primary estimand (see Section 5.2.2) except that death up to Week 16 will be considered if occurred up to 7 days after last dose of study treatment. In addition, premature study treatment discontinuation before Week 16 are captured through the variable definition. |
| **D.** | **Population-level summary measure:** same as the ANCOVA model of the primary estimand [see Section 5.2.2] |

ANCOVA=Analysis of covariance, FAS=Full Analysis Set, $VO_2$ = Oxygen uptake/consumption.

ii. Sensitivity analyses of the primary estimand

1. Estimator based on the inverse normal combination method with weights proportional to the final actual sample size (i.e., 142 subjects) - to assess the robustness of result when the weights are proportional to the actual sample size of 142 subjects.

   - Same population, same variable, same intercurrent events, and same handling of missing data as defined for the primary estimand (Section 5.2.2).

   - Same ANCOVA model as defined for the primary estimand (Section 5.2.2) with pre-specified weights proportional to the actual sample size of 142 subjects (i.e., $\sqrt{t_1} = \sqrt{\frac{95}{142}} = 0.817932812$ for the 1st stage and $\sqrt{1-t_1} = \sqrt{1-\frac{95}{142}} = 0.575313754$ for the 2nd stage.

For the primary estimand, the following sensitivity estimators (sensitivity analysis) or additional estimand will be evaluated, not controlling for the adaptive design and not based on the population wise splitting approach as applied for the main analysis.

Where applicable, the resulting LS means and corresponding 95% and 99% CLs obtained in each treatment group, and the LS means differences (95% and 99% CLs) for macitentan 10 mg vs placebo will be provided for the following analyses.

2. Estimator based on standard ANCOVA (pooled) – to test statistical hypothesis disregarding the adaptive design for the primary estimand on FAS and for Estimand 1, 2, and 3 as specified above in Table a, Table b, and Table c, respectively.

   - Same population, same variable, same intercurrent events, and same handling of missing data as defined for the primary estimand (Section 5.2.2).

   - Same ANCOVA model as defined for the primary estimand (Section 5.2.2) carried out on the total sample size (without the population wise splitting approach for the control of the adaptive design).

3. Additional Estimand based on observed cases – to assess the robustness of results without imputation for missing values

   - Same population, same variable as defined for the primary estimand (Section 5.2.2). Intercurrent events will not be considered. Missing (or invalid) peak $VO_2$ at baseline and at Week 16 will not be imputed.

   - Same ANCOVA model as defined for the primary estimand (Section 5.2.2).

4. Estimator based on ranks (a) – non-parametric analysis to address possible deviations from the normality assumptions

   - Same population, same variable, same handling of intercurrent events, same imputation rules for missing data as defined for the primary estimand (Section 5.2)

   - ANCOVA based on ranks. Ranking will be performed on the primary variable (change in peak $VO_2$ from baseline to Week 16) and the baseline peak $VO_2$ after addressing intercurrent events or applying rules for missing data as defined (Section 5.2)

The non-parametric ANCOVA procedure incorporates three steps:

a) Transformation of changes from baseline to Week 16 and baseline values for all subjects (regardless of treatment groups) to standardized ranks (i.e., ranks divided by the number of subjects ranked plus 1, mean ranks in case of ties).
b) Determination of residuals from the linear regression of the response variable standardized ranks on baseline variable standardized ranks.
c) Application of the one-sided Wilcoxon-Mann-Whitney test to these residuals. The standardized test statistic with a continuity correction of 0.5 is asymptotically standard normally distributed under the null hypothesis. Two-sided significance level of 0.05 is used.

5. Estimator based on ranks (b) non-parametric analysis to address possible deviations from the normality assumptions

- Same population, same variable as the primary estimand (as defined in Section 5.2.2)
- ANCOVA based on ranks. Ranking will be performed as follows:
  - Missing Week 16 values due to subject's death prior to Week 16 will receive a rank score corresponding to a value of the measurement that is worse than any actually observed value.
  - Missing Week 16 values for any other reason will receive a rank score corresponding to a value of the measurement that is worse than any actually observed value but better than subjects who have died prior to Week 16. Adjustments on ranking will be applied in order that:

$$\text{rank}_{\text{(subjects who die)}} < \text{rank}_{\text{(subjects with missing values)}} < \text{rank}_{\text{(subjects without missing values)}}.$$

To guarantee the above condition the ranking will be performed as follows:

- For subjects who died prior to their Week 16 assessment, missing Week 16 values will be imputed with an arbitrary value of -10000
- For invalid peak $VO_2$ at Week 16 or other conditions causing a missing Week 16 value (see Section 5.2.3.2), the Week 16 value will be imputed with an arbitrary value of -5000
- Changes from baseline will then be calculated using the observed Week 16 values if available or the imputed value (i.e., -10000 or -5000 as defined above)
- Ranking will be performed on the baseline peak $VO_2$ variable and on the changes from baseline to Week 16 in peak $VO_2$.

Same non-parametric ANCOVA procedure as applied above will be incorporated.

6. Estimator based on different model 2 (by adding NYHA functional class at baseline)

- Same population, same variable, same handling of intercurrent events, same imputation rules for missing data as defined for the primary estimand (Section 5.2.2)

- Same ANCOVA model as defined for the primary estimand (Section 5.2.2) by adding the NYHA functional class at baseline in the model.

7. Estimator based on the number of subjects with available Week 16 data pre- and post-Sponsor decision (ie, 24 January 2019) to implement the IA.

   To show that the Sponsor has not been influenced by the acquired data to introduce the IA, the estimator is based on 2 sub-populations: subjects randomized at least 16 weeks before 24 January 2019 (pre internal meeting) and all other remaining subjects (post internal meeting).

   - Same population divided in pre- and post-24-Jan-2019 sub-populations as described above, same variable, same intercurrent events, and same handling of missing data as defined for the primary estimand (Section 5.2.2).

To explore the robustness of the main analysis treatment effect estimate to missing data, the following multiple imputation (MI) sensitivity analyses will be performed on FAS (not controlling for the adaptive design) and applying the main model. Baseline peak $VO_2$ values considered invalid by the central reading facility (eg, improper gas calibration, lack of physiological response) are assumed missing completely at random and these subjects will be excluded from the following MI sensitivity analyses (#8, #9).

8. Sensitivity analysis assuming subjects with missing data, due to premature study treatment discontinuation, have worsened, otherwise missing data are assumed to be missing at random (MAR).

   The intercurrent events of death and premature study treatment discontinuation before week 16 are handled as for the main analysis (Section 5.2.2). Otherwise:

   - Missing/invalid peak $VO_2$ at Week 16 for subjects with premature study treatment discontinuation before Week 16 (if not addressed by a "treatment-policy" strategy): the change from baseline to Week 16 will be imputed with the $25^{th}$ percentile of the observed changes from baseline across both treatment groups (or with 0, if the $25^{th}$ percentile would lead to an improvement).
   - Missing/invalid peak $VO_2$ at Week 16 for subjects without premature study treatment discontinuation before Week 16 are assumed MAR. Changes from baseline to Week 16 will be multiply imputed based on the distribution in the group to which subjects were randomized.

Rationale: Per International Council for Harmonisation (ICH) E9 (R1) addendum, similar assumptions on missing data for participants with and without intercurrent events such as premature study treatment discontinuation may often be considered implausible. This sensitivity analysis explores the treatment effect under the assumption that following premature study treatment discontinuation, a different treatment effect would have been observed compared with participants on study treatment. Specifically, this analysis relaxes the MNAR assumption of worsening in subjects who have missing Week 16 data but have completed at least 16 weeks of DB treatment as compared with the main analysis.

9. Sensitivity analysis assuming missing data are MAR.

The intercurrent events of death and premature study treatment discontinuation before Week 16 are handled as for the main analysis (Section 5.2.2). Otherwise:

- Missing/invalid Week 16 data are assumed MAR and changes from baseline to Week 16 will be multiply imputed based on the distribution in the group to which subjects were randomized. This may include missing/invalid data following premature study treatment discontinuation before Week 16 not addressed by a "treatment-policy" strategy.

Rationale: Even if subjects have not completed 16 weeks of DB study treatment, they may still experience a treatment benefit. Under the MAR assumption, the treatment effect is assumed to be as large as in subjects with observed data. This sensitivity analysis relaxes the MNAR assumption of worsening in all subjects with missing data.

10. Sensitivity analysis adding a random error (normal distributed, variance equal to the residual variance estimated from the respective main model), to the single imputation values (worse case, 25$^{th}$ percentile, or 0 change) planned as main analysis.

Rationale: To evaluate the impact of a potential under-estimation of the variability with the single-value imputation approach.

After consultation with the Study Results Committee constituted of external clinicians with expertise in subjects with Functional Single Ventricle Heart Disease on 13-Aug-2021, the following additional estimand will be added.

11. Additional estimand excluding subjects with cardiac heterotaxy defects (isomerism) – to assess the robustness of results excluding subjects with extra-cardiac defects

- Population: All randomized Fontan-palliated adult and adolescent subjects defined through protocol inclusion/exclusion criteria (ie, FAS) excluding subjects with 'cardiac heterotaxy defects (isomerism)' as identified in Congenital heart defect leading to Fontan-palliation (ie, information collected in the eCRF "Fontan history" module, "Cardiac malformation" section).

- Same variable, same intercurrent events, and same handling of missing data as defined for the primary estimand (Section 5.2.2).

- Same ANCOVA model as defined for the primary estimand (Section 5.2.2) carried out on the total sample size (without the population wise splitting approach for the control of the adaptive design).

#### 5.2.3.5. Subgroups analyses

Analyses will be carried out the same way as on the entire population as described in Section 5.2.3.1, after addressing intercurrent events or applying the rules for imputing missing data described in Section 5.2.3.2, without controlling for the adaptive design and hence not based on the population wise split as applied for the main analysis. For consistency and model performance, a common model is applied for each subgroup as defined in Section 2.3. First, the model is run including randomized treatment, the subgroup variable, and baseline value as covariates, and the interaction treatment-by-subgroup variable. By adding the subgroup variable and treatment-by-subgroup variable interaction terms to the statistical model, interaction tests are performed to check consistency of results accounting for 'number of subgroup categories minus 1' degrees of freedom

at the 0.01 significance level. If the interaction is statistically significant (i.e., p-value < 0.01), least square means estimates for each treatment group as well as treatment effect estimates are derived as contrasts from the extended model (i.e., with the treatment-by-subgroup interaction). Otherwise, least square means for each treatment group as well as treatment effect estimates are taken from the model ran on each subgroup level without the subgroup variable and the interaction term. Least square (LS) means estimates, 99% confidence limits (CLs), and corresponding p-values for the different levels of each subgroup are tabulated by treatment group. Similarly, treatment effect LS means estimates are tabulated and also presented in a forest plot (Cuzick 2005) with a vertical reference line displayed at the level of the overall treatment effect macitentan versus placebo. An additional forest plot will be produced displaying 95% CLs instead of 99% CLs. In addition, to test heterogeneity of treatment effect across pre-defined subgroups, the p-value of the interaction test (treatment-by-subgroup) will be displayed on the plot.

## 5.3. Secondary Efficacy Endpoints

The secondary efficacy endpoints are listed here according to the hierarchical order in which they will be statistically tested at the same significance level as the primary efficacy endpoint:

- Change from baseline over 52 weeks in peak $VO_2$
- Change from baseline to Week 16 in mean count per minute of daily PA-Ac.

### 5.3.1. Change from baseline over 52 weeks in peak $VO_2$

#### 5.3.1.1. Definition

The first secondary efficacy endpoint is the change from baseline over Week 52 in peak $VO_2$ (mL/kg/min).

Peak $VO_2$ is measured during CPET. The values used for peak $VO_2$ at baseline, at Week 16, and at Week 52 (within the pre-defined corresponding time windows, see Section 2.1) will be based on a blinded data evaluation by the independent central reading facility, as reported in the eCRF.

For Week 16 and Week 52, the assessments (including the unscheduled ones) will be mapped to the corresponding time windows based on the Study Day (see Section 2.1), irrespective of premature study treatment discontinuation before the assessment.

#### 5.3.1.2. Estimand

The primary estimand for this secondary efficacy endpoint follow the same ‘treatment-policy’ strategy as the primary efficacy endpoint estimand.

The corresponding estimand defining the treatment effect on this endpoint has the following attributes:

**A. Population:** All randomized Fontan-palliated adult and adolescent subjects defined through protocol inclusion/exclusion criteria (ie, FAS).

**B. Variable:** peak $VO_2$ as average change from baseline to Week 16 and Week 52.

**C. Intercurrent events:**

- Death occurring before Week 52: composite strategy (worst-case).

In case of death prior to Week 52, all scheduled assessments following the date of death will be imputed by a worst-case imputation (ie, lowest observed change in peak $VO_2$ from baseline value across both treatment groups and over all assessments).

- Non-missing peak $VO_2$ value at Week 16 or Week 52 assessed as 'invalid' by the central reading facility: composite strategy (minimum of 25$^{th}$ percentile and no change)

In case an assessment at Week 16 or Week 52 has been evaluated as "invalid" by the independent central reading facility, it will be ignored. This will result in imputing the change from baseline to Week 16 or Week 52 with the 25$^{th}$ percentile of the observed changes in peak $VO_2$ from baseline value (across both treatment groups and over all assessments up to study day 420). If this leads to an improvement (change >0), a change of 0 (no change) will be imputed.

- Premature study treatment discontinuation before Week 52.

According to the protocol, if treatment is discontinued prior to Week 16, and the subject enters the PTOP, it is requested that the subject returns for a CPET assessment at Week 16 but not at Week 52. Therefore, the premature study treatment discontinuation has been added as an intercurrent event for the evaluation of peak $VO_2$ at Week 52. Any assessment in the Week 52 time window (including the unscheduled ones), irrespective of premature study treatment discontinuation, will be included in the main analysis if available. In this way, there is consistency of strategy for the evaluation of peak $VO_2$ in the short-term (i.e., at Week 16) and in the long-term (i.e, over 52 weeks).

If the premature study treatment discontinuation occurs prior to Week 16:

  - If a valid assessment is available within the time window of Week 16, this will result in imputing only the change from baseline to Week 52 with the 25$^{th}$ percentile of the observed changes in peak $VO_2$ from baseline value (across both treatment groups and over all assessments up to study day 420). If this leads to an improvement, a change of 0 (no change) will be imputed.
  - If no valid assessment is available within the time window of Week 16, all scheduled assessments following the date of event will be imputed by applying the 25$^{th}$ percentile of the observed changes in peak $VO_2$ from baseline value (across both treatment groups and over all assessments up to study day 420). If this leads to an improvement (change>0), a change of 0 (no change) will be imputed. This is consistent with the primary estimand for the primary efficacy endpoint.

From protocol version 7 onwards, Visit 5/Week 32 was converted from a site visit to a telephone call, therefore the CPET at Week 32 is not anymore conducted. It is expected that approximately 40 subjects will have a CPET assessment at Week 32. In order to limit the amount of missing values, in case of missing data at Week 52 following premature study treatment discontinuation between Week 16 and Week 52, peak $VO_2$ at Week 52 will be replaced with a valid value within the Week 32 time window (see Section 2.1), if available.

If the premature study treatment discontinuation occurs after Week 16 and before Week 52 (and no valid assessment is available within the time window of Week 32 or Week 52), this will result in imputing the change from baseline to Week 52 with the 25$^{th}$ percentile of the observed changes in peak $VO_2$ from baseline value (across both treatment groups and over all assessments up to study day 420). If this leads to an improvement (change>0), a change of 0 (no change) will be imputed.

**D. Population-level summary measure**: difference in mean change in peak $VO_2$ from baseline over Week 52 (i.e., average of Week 16 and Week 52) between macitentan 10 mg and placebo, as estimated from the MMRM.

#### 5.3.1.3. Hypotheses and statistical model

The statistical hypotheses for this secondary efficacy endpoint are formulated in terms of the M of change from baseline over 52 weeks (i.e., average of Week 16 and Week 52) in peak $VO_2$ in subjects treated with either placebo or macitentan 10 mg.

$H_0$: $M_{\text{macitentan 10 mg}} = M_{\text{placebo}}$ versus

$H_A$: $M_{\text{macitentan 10 mg}} \neq M_{\text{placebo}}$

The null hypothesis will be tested at the same significance level of the primary efficacy endpoint ie, 1% two-sided) based on the inverse normal combination method.

The null hypothesis will be tested by means of a MMRM utilizing the post-baseline assessments at Week 16 and Week 52.

#### 5.3.1.4. Handling of missing data for the main analysis

Missing values for peak VO2 at baseline are not allowed by protocol, but in case a subject has been randomized with a non-missing baseline value that has subsequently been evaluated as "invalid" by the independent central reading facility, it will be imputed with the median peak $VO_2$ value observed at baseline from the remaining subjects (across treatment groups).

Handling of intercurrent events of death, invalid peak $VO_2$, and premature study treatment discontinuation are described under the attribute 'C. Intercurrent events" of the estimand in Section 5.3.1.2.

Missing data at Week 52 arising from any other reason (e.g., administrative reason) are assumed to be missing at random, and as such, will require no further imputation, since the MMRM approach will implicitly impute these values under a missing at random assumption.

Any assessment in the Week 16 time window (including the unscheduled ones), irrespective of premature study treatment discontinuation before Week 16, will be included in the main analysis. Missing values at Week 16 will be implicitly imputed by the MMRM approach, if not addressed under the attribute 'C. Intercurrent events" of the estimand in Section 5.3.1.2.

#### 5.3.1.5. Main analysis

The final analysis of this secondary efficacy endpoint will be conducted using the inverse normal combination method with pre-specified weights to combine first and second stage p-values (see Section 3.2.3 of the SAP for SPA).

The difference between macitentan 10 mg and placebo over Week 52 (i.e., average of Week 16 and Week 52) and corresponding p-value will be estimated at each stage separately (first-stage and second-stage) from the MMRM, using a population-wise splitting of data approach. The model will include randomized treatment group, time (via a categorical variable for visit), treatment-by-time interaction, baseline-by-time interaction, and peak $VO_2$ at baseline as fixed effects. An unstructured variance-covariance matrix will be specified.

Similarly to the primary efficacy endpoint, the first-stage p-value from the MMRM will be based on the actual number of subjects who contributed to the IA and have performed the Week 52 assessment. The separate second-stage analysis will be based on the remaining number of subjects required to reach the total sample size based on the IA results. Due to the population-wise splitting, no subjects contributing to the separate first-stage test statistics are included into the separate second-stage analysis. The second-stage p-values will be calculated in the same way as the first-stage p-values.

The separate first and second p-values will be combined using the inverse-normal combination function with prefixed weights (i.e., see Table 11 in Section 5.1.1). Final interpretation of the results will then be based on the final adjusted p-value, median unbiased estimator for the overall treatment effect and corresponding RCIs obtained by ADDPLAN™ 6.1 (ADDPLAN Inc. an Aptiv Solutions Co. 2014).

The number of subjects with intercurrent events or missing data will be presented by treatment group and the change in peak $VO_2$ from baseline overWeek 52 (i.e., average change from baseline to Week 16 and Week 52 after handling of intercurrent events or applying imputation rules for missing data) will also be summarized using descriptive statistics by treatment group.

To explore the treatment course over time, a plot of the LS mean changes (±SE) in peak $VO_2$ from baseline over time will be displayed by treatment group based on a MMRM analysis disregarding the adaptive design. The baseline is added to the time axis to visualize an initial change from 0.

The normality and equal variance assumptions underlying the MMRM model will be assessed graphically for the change from baseline over 52 weeks in peak $VO_2$.

#### 5.3.1.6. Supportive/sensitivity analyses

Absolute values at baseline and at Week 52 as well as absolute change in peak $VO_2$ from baseline to Week 52 will also be summarized using descriptive statistics, for each supportive and sensitivity analysis where applicable.

i. Additional estimands

Additional estimands will be assessed as defined in the tables below. The same MMRM model as the one described for the primary estimand (see Section 5.3.1.2) will be applied, based on the population wise splitting as applied for the main analysis.

**Table d Additional Estimand 1 (PP2)**

| | |
|---|---|
| **A.** | **Population:** All subjects who received study treatment as assigned by IRT and who complied with the protocol sufficiently up to and including the Week 52 for this secondary efficacy endpoint (ie, PP2). Protocol deviations are defined in Section 2.2.3. |
| **B.** | **Variable: s**ame as for primary estimand (see Section 5.3.1.2) |
| **C.** | **Intercurrent events and strategies**: same as primary estimand (see Section 5.3.1.2) |
| **D.** | **Population-level summary measure:**. Same as the MMRM model of the primary estimand (see Section 5.3.1.2) |

IRT=Interactive Response Technology, MMRM=Mixed Model Repeated Measures, PP2= Estimand Population for secondary efficacy endpoint.

**Table e Additional Estimand 2 (excluding subjects with invalid peak $VO_2$ at baseline)**

| | |
|---|---|
| **A.** | **Population:** same as primary estimand (FAS) but excluding subjects with invalid peak $VO_2$ at baseline (see Section 5.3.1.2) |
| **B.** | **Variable: s**ame as for primary estimand (see Section 5.3.1.2) |
| **C.** | **Intercurrent events and strategies**: same as primary estimand (see Section 5.3.1.2) |
| **D.** | **Population-level summary measure:**. same as the MMRM model of the primary estimand (see Section 5.3.1.2) |

FAS=Full Analysis Set, MMRM= Mixed Model Repeated Measures, $VO_2$= Oxygen uptake/consumption.

**Table f Additional Estimand 3 ("while on treatment")**

| | |
|---|---|
| **A.** | **Population:** same as primary estimand (FAS) |
| **B.** | **Variable: s**ame as for primary estimand (see Section 5.3.1.2) except that the assessments in the Week 16 or Week 52 time windows will be included in the analysis if 'on-treatment', defined as up to 7 days after last dose of study treatment. |
| **C.** | **Intercurrent events and strategies**: same as primary estimand (see Section 5.3.1.2) except that death will be considered up to 7 days after last dose of study treatment. Premature study treatment discontinuation are captured through the variable definition. |
| **D.** | **Population-level summary measure:** same as the MMRM model of the primary estimand (see Section 5.3.1.2) |

FAS=Full Analysis Set, MMRM= Mixed Model Repeated Measures.

ii. Sensitivity analyses of the primary estimand

1. Estimator based on the inverse normal combination method with weights proportional to the final actual sample size (i.e., 142 subjects) - to assess the robustness of result when the weights are proportional to the actual sample size of 142 subjects.

   - Same population, same variable, same intercurrent events, and same handling of missing data as defined for the primary estimand (Section 5.3.1.2).

   - Same MMRM model as defined for the primary estimand (Section 5.3.1.2) with pre-specified weights proportional to the actual sample size of 142 subjects (i.e., $\sqrt{t_1} = \sqrt{\frac{95}{142}} = 0.817932812$ for the 1st stage and $\sqrt{1-t_1} = \sqrt{1-\frac{95}{142}} = 0.575313754$ for the 2nd stage.

For the primary estimand, the following sensitivity estimator (sensitivity analysis) will be evaluated. Sensitivity estimators will not control for the adaptive design and will not be based on the population wise splitting as applied for the main analysis.

2. Estimator based on standard MMRM (pooled) to test statistical hypothesis disregarding the adaptive design for the primary estimand on FAS and for Estimand 1, 2, and 3 as specified in Table d, Table e, and Table f above respectively

   - Same population, same variable, same intercurrent events, and same handling of missing data as defined for the primary estimand (Section 5.3.1.2).
   - Same MMRM model as defined for the primary estimand (Section 5.3.1.2) carried out on the total sample size (without the population wise splitting approach for the control of the adaptive design).

At Week 52, the resulting LS means and corresponding 95% and 99% CLs obtained in each treatment group, and the LS means differences (95% and 99% CLs) for macitentan 10 mg vs placebo will be provided.

3. Estimator based on ANCOVA

**A. Population:** same as primary estimand (FAS).
**B. Variable:** change from baseline to Week 52.
**C. Intercurrent events and strategies**:
- Death before Week 52: worst-case imputation (ie, lowest observed change in peak $VO_2$ from baseline to Week 52 (across both treatment groups)
- Non-missing peak $VO_2$ at Week 52 assessed as 'invalid' by the central reading facility: the change from baseline to Week 52 will be imputed with the 25$^{th}$ percentile of the observed changes in peak $VO_2$ from baseline to Week 52 (across both treatment groups). If this leads to an improvement, a change of 0 (no change) will be imputed.
- Premature study treatment discontinuation before Week 52: peak $VO_2$ at Week 52 (if not available within the corresponding time window) will be replaced with a valid value within the Week 32 time window (see Section 2.1), if available. Otherwise, it will be imputed with the 25$^{th}$ percentile of the observed changes in peak $VO_2$ from baseline to Week 52 (across both treatment groups). If this leads to an improvement, a change of 0 (no change) will be imputed.

**D. Population-level summary measure:** difference in change in peak $VO_2$ from baseline to Week 52 between macitentan 10 mg and placebo, as estimated from the ANCOVA model including randomized treatment group, geographical region and baseline peak $VO_2$ as covariates in the model.

Any missing values which may occur in peak $VO_2$ at Week 52 will be handled as follows: the change from baseline to Week 52 will be imputed with the 25$^{th}$ percentile of the observed changes in peak $VO_2$ from baseline to Week 52 (across both treatment groups). If this leads to an improvement, a change of 0 (no change) will be imputed.

ANCOVA=Analysis of covariance, FAS=Full Analysis Set, $VO_2$= Oxygen uptake/consumption.

Similarly to the primary efficacy endpoint, to explore the robustness of the main analysis treatment effect estimate to missing data, the following MI sensitivity analyses will be performed on FAS (not controlling for the adaptive design) and applying the main model. Baseline peak $VO_2$ values considered invalid by the central reading facility (eg, improper gas calibration, lack of physiological response) are assumed missing completely at random and these subjects will be excluded from the following MI sensitivity analyses (#4, #5, #6):

4. Sensitivity analysis assuming subjects with missing Week 52 data, due to premature study treatment discontinuation, have worsened (irrespective of Week 32 results being available). Otherwise, missing data are assumed to be missing at random (MAR).

   The intercurrent events of death are handled as for the main analysis (Section 5.3.1.2). The intercurrent events of premature study treatment discontinuation are handled as detailed below.

Missing/invalid peak $VO_2$ for subjects with premature study treatment discontinuation:

If the premature study treatment discontinuation occurs prior to Week 16:

- If a valid assessment is available within the time window of Week 16, this will result in imputing only the change from baseline to Week 52 with the 25$^{th}$ percentile of the observed changes in peak $VO_2$ from baseline value (across both treatment groups and over all assessments up to study day 420). If this leads to an improvement, a change of 0 (no change) will be imputed.
- If no valid assessment is available within the time window of Week 16, all scheduled assessments following the date of event will be imputed by applying the 25$^{th}$ percentile of the observed changes in peak $VO_2$ from baseline value (across both treatment groups and over all assessments up to study day 420). If this leads to an improvement (change>0), a change of 0 (no change) will be imputed.

If the premature study treatment discontinuation occurs after Week 16:

- If no valid assessment is available within the time window of Week 52, this will result in imputing the change from baseline to Week 52 with the 25$^{th}$ percentile of the observed changes in peak $VO_2$ from baseline value (across both treatment groups and over all assessments up to study day 420). If this leads to an improvement (change>0), a change of 0 (no change) will be imputed

Missing/invalid Week 16 or Week 52 peak $VO_2$ for subjects without premature study treatment discontinuation are assumed MAR. Changes from baseline to Week 16 or Week 52 will be multiply imputed based on the distribution in the group to which subjects were randomized.

5. Sensitivity analysis assuming missing data are MAR:

   The intercurrent events of death are handled as for the main analysis (Section 5.3.1.2). Otherwise:

   - Missing/invalid Week 16 or Week 52 peak $VO_2$ are assumed MAR and changes from baseline to Week 16 or Week 52 will be multiply imputed based on the distribution in the group to which subjects were randomized.

6. Sensitivity analysis assuming subjects with missing data would have worsened and experienced no treatment effect, irrespective of completing treatment.

   The intercurrent events of death are handled as for the main analysis (Section 5.3.1.2). Otherwise:

   - If no valid assessment is available within the time window of Week 16 or Week 52, the corresponding change from baseline to Week 16 or Week 52 will be imputed by applying the 25$^{th}$ percentile of the observed changes in peak $VO_2$ from baseline value

(across both treatment groups and over all assessments up to study day 420). If this leads to an improvement (change>0), a change of 0 (no change) will be imputed.

7. Sensitivity analysis adding a random error (normal distributed, variance equal to the residual variance estimated from the respective main model), to the single imputation values (worse case, 25$^{th}$ percentile, or 0 change) planned as main analysis.

After consultation with the Study Results Committee constituted of external clinicians with expertise in subjects with Functional Single Ventricle Heart Disease on 13-Aug-2021, the following additional estimand will be added.

8. Additional Estimand excluding subjects with cardiac heterotaxy defects (isomerism) – to assess the robustness of results excluding subjects with extra-cardiac defects

   - Population: All randomized Fontan-palliated adult and adolescent subjects defined through protocol inclusion/exclusion criteria (ie, FAS) excluding subjects with 'cardiac heterotaxy defects (isomerism)' as identified in Congenital heart defect leading to Fontan-palliation (ie, information collected in the eCRF "Fontan history" module, "Cardiac malformation" section).

   - Same variable, same intercurrent events, and same handling of missing data as defined for the primary estimand (Section 5.3.1.2).

   - Same MMRM model as defined for the primary estimand (Section 5.3.1.2) carried out on the total sample size (without the population wise splitting approach for the control of the adaptive design).

#### 5.3.1.7. Subgroups analyses

Analyses will be carried out the same way as on the entire population as described in Section 5.3.1.3, applying the rules for imputing missing data described in Section 5.3.1.4. Subgroup analyses will be displayed as described for the primary efficacy endpoint in Section 5.2.3.5.

### 5.3.2. Change from baseline to Week 16 in mean count per minute of daily PA-Ac

The daily physical activity (counts/min) of the subject is assessed via accelerometer during the daytime. The accelerometer is given to the subject at Visit 1 (Screening visit), and data are collected for 9 consecutive daily daytime periods after Visit 1, Visit 4 (Week 16) and before Visit 6 (Week 52).

The accelerometry central reading facility will read daily counts/min and will analyze duration of daily activity and time (minutes) which will be transferred to the Sponsor.

For Week 16 and Week 52, the 9-day time period is selected by taking the first (in chronological order) interval of 9 consecutive days of daily daytime data with at least 4 complete daily time periods within the corresponding time window (i.e. the 9-day time period should not necessarily fall entirely in the corresponding time window, but it should overlap with the time window for at least 1 day). All assessments (including the unscheduled ones) will be mapped based on the Study Day (see Section 2.1), irrespective of premature study treatment discontinuation before the

assessment. To be considered evaluable, physical activity should have been measured for at least 4 complete daily daytime periods (out of 9 consecutive days) at a specific time point of assessment.

#### 5.3.2.1. Definition

Mean counts per minute of daily PA-Ac are calculated by dividing the sum of all activity counts (Y axis) collected during wear time for a complete day by the number of minutes of wear time in that day across all valid days (at least 4 days out of 9 days).

The activity counts collected during wear time is identifiable with "WearFilteredAxisYcounts" in Data Transfer Specifications (DTS); the number of minutes of wear time corresponds to the "WearMinutes " variable in DTS.

A complete day is defined as a record of at least 7 hours of daily daytime data, ie, a complete day if "WearMinutes" >=7*60, (after excluding the periods when the device was apparently not worn (Troiano 2008).

In summary, to compute the mean counts per minute of daily PA-Ac, it is necessary to identify all complete days; if less than 4 complete days out of 9 are identified then value is missing, otherwise, if at least 4 complete days out of 9 are identified, then the value is calculated on the complete days only.

The secondary efficacy endpoint is the change from baseline to Week 16 in mean count per minute of daily PA-Ac, defined as:

mean count per minute of daily PA-Ac during 9 consecutive daily daytime periods after Visit 4/Week 16 − baseline value.

#### 5.3.2.2. Estimand

The primary estimand defining the treatment effect on this secondary efficacy endpoint has the following attributes:

**A. Population:** All randomized Fontan-palliated adult and adolescent subjects defined through protocol inclusion/exclusion criteria (ie,. FAS) with evaluable 9-day time period for the daily physical activity at baseline.

**B. Variable:** Mean count per minute of daily PA-Ac as change from baseline to Week 16.

**C. Intercurrent event:**

- Death occurring before Week 16.

A composite strategy will be applied: in case of death prior to Week 16, the change from baseline to Week 16 will be replaced with the lowest observed change from baseline to Week 16 across both treatment groups (i.e., worst-case imputation).

- Premature study treatment discontinuation before Week 16

Premature study treatment discontinuation prior to Week 16 will be addressed by a "treatment-policy" strategy by including any assessment in the Week 16 time window, irrespective of premature study treatment discontinuation, in the main analysis.

**D. Population-level summary measure**: Difference in change from baseline to Week 16 in mean count per minute of daily PA-Ac between macitentan 10 mg and placebo, as estimated from the ANCOVA model including randomized treatment group and the mean count per minute of daily PA-AC at baseline as covariates in the model.

#### 5.3.2.3. Hypotheses and statistical model

The statistical hypotheses for this secondary efficacy endpoint are formulated in terms of the M of change in mean count per minute of daily PA-Ac in subjects treated with either placebo or macitentan 10 mg from baseline to Week 16.

$H_0$: $M_{\text{macitentan 10 mg}} = M_{\text{placebo}}$ versus

$H_A$: $M_{\text{macitentan 10 mg}} \neq M_{\text{placebo}}$

The null hypothesis will be tested at the same significance level of the primary efficacy endpoint ie, 1% two-sided) based on the inverse normal combination method.

Assuming normal distribution of the primary variable, the null hypothesis will be tested by means of an ANCOVA model on the change in mean count per minute of daily PA-Ac. Model covariates will include randomized treatment group and the mean count per minute of daily PA-AC at baseline as covariates.

#### 5.3.2.4. Handling of missing data for the main analysis

Intercurrent events will be handled as defined in Section 5.3.2.2.

If the baseline is available, in case of missing values in mean count per minute of daily PA-Ac at Week 16, the following approach for imputation will be applied:

| Reason | Strategy in handling the missing change from baseline to Week 16 |
|---|---|
| Other | A change of 0 (no change) will be imputed. |

#### 5.3.2.5. Main analysis

The final analysis of this secondary efficacy endpoint will be conducted using the inverse normal combination method with pre-specified weights to combine first and second stage p-values (see Section 3.2.3 of the SAP for SPA).

The difference between macitentan 10 mg and placebo and corresponding p-value will be estimated at each stage separately (first-stage and second-stage) from the ANCOVA model, using a population-wise splitting of data approach. Model covariates will include randomized treatment group and mean count per minute of daily PA-Ac at baseline.

The separate first-stage analysis will be based on the first-stage set of subjects contributing to the IA. The separate second-stage analysis will be based on the remaining number of subjects required to reach the total sample size based on the IA results. Due to the population-wise splitting, no subjects contributing to the separate first-stage test statistics are included into the separate second-stage analysis. The second-stage p-values will be calculated in the same way as the first-stage p-values.

The separate first and second p-values will be combined using the inverse-normal combination function with prefixed weights (i.e., see Table 11 in Section 5.1.1). Final interpretation of the results will then be based on the final adjusted p-value, median unbiased estimator for the overall treatment effect and corresponding RCIs obtained by ADDPLAN™ 6.1 (ADDPLAN Inc. an Aptiv Solutions Co. 2014).

Absolute values at baseline and at Week 16 as well as absolute change from baseline to Week 16 in mean count per minute of daily PA-Ac after handling of intercurrent events or applying imputation rules for missing data will also be summarized using descriptive statistics by treatment group.

#### 5.3.2.6. Supportive/sensitivity analyses

Absolute values at baseline and at Week 16 as well as absolute change in mean count per minute of daily PA-Ac from baseline to Week 16 will also be summarized using descriptive statistics, for each supportive and sensitivity analysis where applicable.

i. Additional estimands

Additional estimands will be assessed as defined in the tables below. The same ANCOVA model as the one described for the primary estimand (see Section 5.3.2.2) will be applied, based on the population wise splitting as applied for the main analysis.

**Table g Additional Estimand 1 (PP3)**

| | |
|---|---|
| **A.** | **Population:** All subjects who received study treatment as assigned by IRT and who complied with the protocol sufficiently up to and including the Week 16 for this secondary efficacy endpoint (ie, PP3). Protocol deviations are defined in Section 2.2.3. |
| **B.** | **Variable:** Same as for primary estimand (see Section 5.3.2.2) |
| **C.** | **Intercurrent events and strategies**: same as primary estimand (see Section 5.3.2.2) |
| **D.** | **Population-level summary measure:** Same as the ANCOVA model of the primary estimand (see Section 5.3.2.2) |

ANCOVA=analysis of Covariance, IRT=Interactive Response Technology, PP3= Estimand Population for secondary efficacy endpoint.

ii. Sensitivity analyses of the primary estimand

1. Estimator based on the inverse normal combination method with weights proportional to the final actual sample size (i.e., 142 subjects) - to assess the robustness of result when the weights are proportional to the actual sample size of 142 subjects.

    - Same population, same variable, same intercurrent events, and same handling of missing data as defined for the primary estimand (Section 5.3.2.2).

- Same ANCOVA model as defined for the primary estimand (Section 5.3.2.2) with pre-specified weights proportional to the actual sample size of 142 subjects (i.e., $\sqrt{t_1} = \sqrt{\frac{95}{142}} = 0.817932812$ for the 1st stage and $\sqrt{1-t_1} = \sqrt{1-\frac{95}{142}} = 0.575313754$ for the 2nd stage.

For the primary estimand, the following sensitivity estimator (sensitivity analysis) or additional estimand will be evaluated, not controlling for the adaptive design and not based on the population wise splitting approach as applied for the main analysis.

Where applicable, the resulting LS means and corresponding 95% and 99% CLs obtained in each treatment group, and the LS means differences (95% and 99% CLs) for macitentan 10 mg vs placebo will be provided for the following analyses.

2. Estimator based on standard ANCOVA (pooled) - to test statistical hypothesis disregarding the adaptive design for the primary estimand on FAS and on PP (i.e., Estimand 1 in **Table g** above)

   - Same population, same variable, same intercurrent events, and same handling of missing data as defined for the primary estimand (Section 5.3.2.2).
   - Same ANCOVA model as defined for the primary estimand (Section 5.3.2.2) carried out on the total sample size (without the population wise splitting approach for the control of the adaptive design).

3. Additional Estimand based on observed cases

   - Same population, same variable as defined for the primary estimand (Section 5.3.2.2). Intercurrent events will not be considered. Missing mean count per minute of daily PA-Ac at Week 16 will not be imputed.
   - Same ANCOVA model as defined for the primary estimand (Section 5.3.2.2).

4. Estimator based on the number of subjects with available Week 16 data pre- and post-Sponsor decision (ie, 24 January 2019) to implement the IA.

As for the primary efficacy endpoint, the estimator is based on 2 sub-populations: subjects randomized at least 16 weeks before 24 January 2019 (pre internal meeting) and all other remaining subjects (post internal meeting).

   - Same population divided in pre- and post-24-Jan-2019 sub-populations as described above, same variable, same intercurrent events, and same handling of missing data as defined for the primary estimand (Section 5.3.2.2).

Similarly to the primary efficacy endpoint, to explore the robustness of the main analysis treatment effect estimate to missing data, the following multiple imputation sensitivity analyses will be performed on FAS (not controlling for the adaptive design) and applying the main model. Missing baseline data are assumed missing completely at random and therefore are not included in the following analyses. To investigate the plausibility of this assumption, peak $VO_2$ at baseline and other key baseline characteristics will be summarized by PA-Ac data available at baseline, to

explore if subjects with missing baseline for PA-Ac do not have systematically lower or higher maximum exercise capacity.

5. Sensitivity analysis assuming subjects with premature study treatment discontinuation before Week 16 have experienced no treatment effect, otherwise missing data are assumed to be MAR.

   The intercurrent event of death and premature study treatment discontinuation before Week 16 are handled as for the main analysis (Section 5.3.2.2).

   - Missing mean count per minute of daily PA-Ac at Week 16 for subjects with premature study treatment discontinuation before Week 16 (in case not addressed by treatment-policy strategy): the change from baseline to Week 16 will be imputed with a change of 0.
   - Missing mean count per minute of daily PA-Ac at Week 16 for subjects without premature study treatment discontinuation before Week 16 are assumed MAR. Changes from baseline to Week 16 will be multiply imputed based on the distribution in the group to which subjects were randomized.

6. Sensitivity analysis assuming missing data are MAR:

   The intercurrent events of death and premature study treatment discontinuation before Week 16 are handled as for the main analysis (Section 5.3.2.2). Otherwise:

   - Missing mean count per minute of daily PA-Ac data at Week 16 are assumed MAR and changes from baseline to Week 16 will be multiply imputed based on the distribution in the group to which subjects were randomized. This may include missing/invalid data following premature study treatment discontinuation before Week 16 not addressed by a "treatment-policy" strategy.

7. Sensitivity analysis adding a random error (normal distributed, variance equal to the residual variance estimated from the respective main model), to the single imputation values (worse case or 0 change) planned as main analysis.

After consultation with the Study Results Committee constituted of external clinicians with expertise in subjects with Functional Single Ventricle Heart Disease on 13-Aug-2021, the following additional estimand will be added.

8. Additional Estimand excluding subjects with cardiac heterotaxy defects (isomerism) – to assess the robustness of results excluding subjects with extra-cardiac defects

   - Population: All randomized Fontan-palliated adult and adolescent subjects defined through protocol inclusion/exclusion criteria (ie, FAS) excluding subjects with 'cardiac heterotaxy defects (isomerism)' as identified in Congenital heart defect leading to Fontan-palliation (ie, information collected in the eCRF "Fontan history" module, "Cardiac malformation" section).
   - Same variable, same intercurrent events, and same handling of missing data as defined for the primary estimand (Section 5.3.2.2).
   - Same ANCOVA model as defined for the primary estimand (Section 5.3.2.2) carried out on the total sample size (without the population wise splitting approach for the control of the adaptive design).

#### 5.3.2.7. Subgroups analyses

Analyses will be carried out the same way as on the entire population as described in 5.3.2.3, applying the rules for imputing missing data described in Section 5.3.2.4. Subgroup analyses will be displayed as described for the primary efficacy endpoint in Section 5.2.3.5.

### 5.4. Other Efficacy Endpoint(s)

All other efficacy endpoints are evaluated in an exploratory and descriptive manner on the FAS. The estimators will not control for the adaptive design and will not be based on the population wise splitting as applied for the main analysis of the primary and secondary efficacy endpoints.

For Week 16 and Week 52, the assessments (including the unscheduled ones) will be mapped to the corresponding time windows based on the Study Day (see Section 2.1), irrespective of premature study treatment discontinuation before the assessment.

Missing values will not be imputed, unless otherwise specified.

#### 5.4.1. Endpoints related to exercise capacity

Endpoints related to exercise capacity are collected at baseline, Week 16 (Visit 4) and Week 52 (Visit 6) with CPET.

#### 5.4.1.1. Definition

Other efficacy variables related to exercise capacity are:

- change in peak $VO_2$ (mL/kg/min) from baseline to Week 52,

This variable is different from the secondary efficacy endpoint *'Change from baseline over 52 weeks in peak $VO_2$'* as it does not account for the intermediate timepoint at Week 16.

- change from baseline to Week 16 in: $VO_2$ at ventilatory anaerobic threshold (VAT) in mL/kg/min, oxygen pulse expressed as $VO_2$/HR (mL $O_2$/beat), oxygen uptake efficiency slope (OUES), % predicted peak $VO_2$, $VE/VCO_2$ slope
- change from baseline over 52 weeks in: $VO_2$ at ventilatory anaerobic threshold (VAT) in mL/kg/min, oxygen pulse expressed as $VO_2$/HR (mL $O_2$/beat), oxygen uptake efficiency slope (OUES), % predicted peak $VO_2$, $VE/VCO_2$ slope.

#### 5.4.1.2. Statistical methods

- The change in peak $VO_2$ from baseline to Week 52 will be analyzed by means of an ANCOVA, including randomized treatment group and baseline peak $VO_2$ as covariates in the model. This analysis is different from the sensitivity analysis of the secondary efficacy endpoint (i.e., Estimator based on ANCOVA) which includes randomized treatment group, geographical region, and baseline peak $VO_2$ as covariates in the ANCOVA model. The same handling of intercurrent events and missing data will be applied as the Estimator based on ANCOVA for the secondary efficacy endpoint.

- The change from baseline to Week 16 in $VO_2$ at VAT, OUES, and % predicted peak $VO_2$ will be analyzed similarly with an ANCOVA, including randomized treatment group and corresponding baseline value as covariates in the model.
- The change from baseline over 52 weeks (i.e., average of Week 16 and Week 52) in $VO_2$ at VAT, OUES, and % predicted peak $VO_2$ will be analyzed by means of MMRM model. The model will include randomized treatment group, time (via a categorical variable for visit), treatment-by-time interaction, baseline-by-time interaction and the value at baseline as fixed effects. An unstructured variance-covariance matrix will be specified.
- The change from baseline to Week 16 and Week 52 in $VO_2$/HR and $VE/VCO_2$ slope will be analyzed using descriptive statistics.
- $VE/VCO_2$ slope and peak $VO_2$ at baseline will be analyzed as categorical variables displaying for each treatment group the number and proportion of subjects with any clinical event up to End-of-Study as defined in Section 5.4.3 in each category.

The categories for $VE/VCO_2$ slope and peak $VO_2$ at baseline:

- $VE/VCO_2$: VC I (≤29.9), VC II (30.0-35.9), VC III (36.0-44.9), VC IV (≥45.0);
- peak $VO_2$ (<25, 25-<40, 40-<55, ≥55 mL/kg/min).

Receiver operating characteristic (ROC) curve analysis will be performed in order to confirm the categorizations for $VE/VCO_2$ slope and peak $VO_2$ variables for predicting the first occurrence of any of the clinical events.

Clinical Events are considered as binary (whether or not a composite clinical event has occurred), predicted by $VE/VCO_2$ and $VO_2$ at baseline which will be included in the model as continuous variables to identify the best cut-offs. From this analysis, optimal cut-offs for $VE/VCO_2$ and peak $VO_2$ variables will be derived based on the sensitivity and specificity values. The optimal cut-off is the point closest in distance to the upper left hand corner of the ROC curve where sensitivity and specificity are equal to 1, it is easily identifiable taking the cut-off that minimizes the square root of the sum of squares of 1-sensitivity and 1-specificity, i.e. $\min\{\text{sqrt}[(1\text{-sensitivity})^2+(1\text{-specificity})^2]\}$. In addition, the Area Under the Curve (AUC) will be provided for a better characterization of the accuracy of the prediction.

Absolute values (at baseline and at each post-baseline visit) as well as the absolute change from baseline to each visit for the efficacy variables related to exercise capacity (i.e., peak $VO_2$, $VO_2$ at VAT, oxygen pulse expressed as $VO_2$/HR, OUES, % predicted peak $VO_2$, and $VE/VCO_2$ slope) will be summarized using descriptive statistics by treatment group on the FAS.

Individual subject's listings will be provided on the FAS.

### 5.4.2. NT-proBNP

NT-pro-BNP is measured at Visits 2 (Randomization), 4 (Week 16), and 6 (Week 52), according to central laboratory working procedures. NT-pro-BNP (pmol/L) samples are processed through

the central laboratory and the results will be sent electronically to the Sponsor after database lock and unblinding of the study treatment code. The results will be available in SDTM for the analysis.

#### 5.4.2.1. Definition

The percent of baseline in NT-proBNP at Week 16 (defined as 100 times the post-baseline value divided by the baseline value, see formula below) will be evaluated.

$$\left(\frac{NT-proBNP\ at\ Week\ 16}{NT-proBNP\ at\ baseline}\right)\times 100$$

Similarly, the percent of baseline in NT-proBNP at Week 52 (defined as 100 times the post-baseline value divided by the baseline value, see formula below) will be evaluated.

$$\left(\frac{NT-proBNP\ at\ Week\ 52}{NT-proBNP\ at\ baseline}\right)\times 100$$

#### 5.4.2.2. Statistical methods

NT-proBNP is assumed to follow a log-normal distribution. Therefore, data are log transformed before analysis.

An ANCOVA is carried out on log transformed data (i.e., natural log scale) first for Week 16 and then for Week 52. The model for log transformed ratio Week 16 (or Week 52)/baseline includes randomized study treatment, the stratification factor (i.e. geographical region) and baseline NT-proBNP (in natural log scale).

For Week 16 and Week 52 separately, after fitting the ANCOVA model, the resulting adjusted least square means and 2-sided 95% CLs are inversely transformed using the exponential function and multiplied by 100 to display the geometric mean and corresponding 2-sided 95% CLs, expressed in percent, for each treatment group. Similarly the placebo-corrected treatment effect, presented as ratio of geometric means (macitentan over placebo) together with its 2-sided 95% CLs is provided.

Values at baseline and at each post-baseline visit (Week 16 and Week 52) as well as the absolute change from baseline to each visit will be summarized displaying, for each treatment group, descriptive statistics for continuous variables. The percent of baseline will also be summarized at each post-baseline visit using geometric means and two-sided 95% confidence intervals of the geometric means.

To graphically display changes in NT-proBNP over time, a plot of the geometric mean (and 95% CLs) for the percent of baseline at each post-baseline visits will be displayed by treatment group. A reference line is included at 100% (zero change) and baseline is added to the time axis to visualize an initial change from 100%.

Individual subject's listing will be provided on the FAS.

### 5.4.3. Time to clinical events

#### 5.4.3.1. Definition

The individual components of the following two clinical events endpoints are defined below:

- Composite endpoint of events related to Fontan-palliated clinical worsening and,
- Composite endpoint of events related to Fontan-palliated morbidity.

The dates of occurrence of each component of the above two clinical events endpoints are reported in the eCRF.

For each composite endpoint, a subject is considered as having an event if the investigator reports the occurrence of any of the individual component in the eCRF. Subjects who did not experience an event will be right-censored at EOS.

The times to clinical events are expressed in weeks and calculated as:

- For subjects with an event: the earliest date of occurrence of any of the individual components minus date of randomization plus 1 divided by 7,
- For censored subjects: the EOS date minus date of randomization plus 1 divided by 7.

**Time to event endpoint definition**

- **Composite endpoint of events related to Fontan-palliated clinical worsening**

The components of the composite endpoint of events related to Fontan-palliated clinical worsening are the following (these are all assessed by the investigator and recorded in the eCRF):

- Unscheduled hospitalization for Fontan-palliated morbidity event.
- Signs and symptoms of heart failure, requiring change in diuretic therapy.
- Clinical worsening leading to interventions related to the Fontan-palliated condition.
- Worsening to NYHA FC III, investigator assessed using the Specific Activity Scale [Appendix 4 of protocol].
- Signs and symptoms requiring the addition of a new class of cardiovascular medication (e.g., nitrates, alpha-blockers, or Endothelin receptor antagonists (ERAs)), or insertion of a pacemaker.
- Failing-Fontan defined as one or more of the following:
    - Enlisted on the active list for heart transplantation or effective heart transplantation,
    - Reoperation (e.g., mechanical circulatory support, Fontan take down, Fontan revision / conversion, AV valve repair/replacement),
    - Worsening to NYHA FC IV, investigator assessed using the Specific Activity Scale [Appendix 4 of protocol],
    - Protein-losing enteropathy (PLE),

- Plastic bronchitis/chyloptysis,
- Peritoneal, pleural, mediastinal, or pericardial effusions,
- Severe hepatic impairment (as described in exclusion criterion 6.3),
- Severe renal impairment (as described in exclusion criterion 6.5),
- Death related to Failing-Fontan.

Clinical events are reported in "Clinical worsening events" section of the eCRF.

- **Composite endpoint of events related to Fontan-palliated morbidity.**

The components of the composite endpoint of events related to Fontan-palliated morbidity are the following:

- Ventricular tachyarrhythmia or supraventricular tachyarrhythmia,
- Thrombotic or hemorrhagic complications (including thromboembolism and hemoptysis).

Clinical events are reported in "Other Fontan-palliated morbidity events" section of the eCRF.

#### 5.4.3.2. Statistical methods

For each of the two composite endpoints listed above the time to the clinical event will be analyzed, separately, using the Kaplan-Meier (KM) product limit method, providing estimates for each treatment group and corresponding 95% two-sided CLs at weeks 8, 16, 32, 52, and EOS. The CLs are constructed using Greenwood's formula (Collett 2003) for the standard error of the Kaplan-Meier estimate and are added to the plot, as well as the p-vale from the logrank test. The number of subjects at risk, censored, and with events will be computed and displayed at each time point for each group.

Hazard ratios and corresponding 95% confidence intervals will be estimated by means of a Cox regression model.

Data will be summarized in tables or figures, including: number of events, number of censored observations, number of subjects at risk, and KM estimates of the survival function for time-to-event variables.

The graph of the estimated survival function of the time to first composite clinical endpoint for each treatment group obtained from the Kaplan-Meier product-limit method will be displayed up to the time at which at least 10% of all subjects remain at risk of an event. The graphical presentation follows the recommendations from (Pocock 2002).

Individual subject's listings will be provided for both composite endpoints separately.

### 5.4.4. Overall Survival (OS)

This endpoint will not be analyzed if no death has occurred. The following estimates will be provided if at least one death per treatment group has occurred.

#### 5.4.4.1. Definition

The vital status (alive, dead, or unknown) will be reported in the eCRF within 8 weeks of study closure announcement at the VSFU visit ("Vital Status Follow-up" eCRF form), or in the death form if the subject died prior to EOS.

Subjects who are known to be alive at the time of VSFU will be censored at the date of VSFU (date of contact). Subjects with 'unknown' vital status at VSFU Visit as collected in the "Vital Status Follow-up" form (or VSFU Visit not performed) will be censored at EOS.

For those subjects with the EOS within 4 weeks of the SCL announcement (ie, 30-Jun-2021), the Vital Status Follow up does not need to be performed. Therefore, these subjects are assumed to be alive at VSFU and will be censored at their EOS, which is their VSFU.

The VSFU could not be assessed for those subjects randomized in site 1101 due to rejection from Human Genetics Resources Administration of China to continue the study at this site. Therefore, these subjects are assumed with 'unknown' vital status and will be censored at their EOS unless the subject died prior to EOS.

Time to death is expressed in weeks and calculated as the date of death minus date of randomization plus 1, divided by 7 or, for censored subjects, as censoring date minus date of randomization plus 1, divided by 7.

#### 5.4.4.2. Statistical methods

Time to death will be analyzed using the Kaplan-Meier (KM) product limit method, providing estimates for each treatment group and corresponding 95% two-sided CLs at weeks 8, 16, 32, 52, and EOS. The CLs are constructed using Greenwood's formula (Collett 2003) for the standard error of the Kaplan-Meier estimate and are added to the plot, as well as the p-vale from the logrank test. The number of subjects at risk, censored, and with events will be computed and displayed at each time point for each group.

Hazard ratios and corresponding 95% confidence intervals will be estimated by means of a Cox regression model.

Data are listed and summarized in tables or figures, including: number of events, number of censored observations, number of subjects at risk, and KM estimates of the survival function for time-to-event variables.

The graph of the estimated survival function of the time to first composite clinical endpoint for each treatment group obtained from the Kaplan-Meier product-limit method will be displayed up to the time at which at least 10% of all subjects remain at risk of an event. The graphical presentation follows the recommendations from (Pocock 2002).

Hazard ratios and corresponding 95% confidence intervals obtained using Cox regression model will be summarized.

### 5.4.5. Physical activity

#### 5.4.5.1. Definition

Physical activity efficacy variables include:

- change from baseline over 52 weeks in mean count per minute of daily activity.
- change from baseline to Week 16 and Week 52 in daily mean time in minutes spent in sedentary, light, moderate, or vigorous (including very vigorous for adults).

#### 5.4.5.2. Statistical methods

The change from baseline over 52 weeks (i.e., average of Week 16 and Week 52) in mean count per minute will be analyzed by means of a MMRM model. The model will include randomized treatment group, time (via a categorical variable for visit), treatment-by-time interaction, baseline-by-time interaction and mean count per minute at baseline as fixed effects. An unstructured variance-covariance matrix will be specified. Absolute values (at baseline and at each post-baseline visit) as well as absolute changes from baseline will also be summarized using descriptive statistics.

To explore the treatment course over time, a plot of the LS mean changes (±SE) in mean count per minute from baseline over time will be displayed by treatment group based on the MMRM analysis. The baseline is added to the time axis to visualize an initial change from 0

The daily mean time in minutes spent in:

- sedentary physical activity,
- light physical activity,
- moderate physical activity,
- vigorous (including very vigorous for adults) physical activity

will be summarized as a continuous variable by displaying, for each treatment group, descriptive statistics for the absolute values at baseline, at each post-baseline visit as well as for the absolute change from baseline. Furthermore, the percent of daily mean total activity time spent in each category will be summarized similarly (as continuous variable) including the graphical display (i.e. bar plot) at each time point of assessment.

In addition, the number of complete days of wearing the accelerometer at baseline, Week 16 and Week 52 will also be summarized descriptively.

## 6. SAFETY

The following safety endpoints are defined in the protocol (see protocol Section 6.2):

- Treatment-emergent adverse events (AEs) and serious AEs up to 30 days after study treatment discontinuation, or until initiation of study drug in the AC-055H302 OL extension study, whichever occurs first
- AEs leading to premature discontinuation of study treatment

- Change in vital signs (systolic and diastolic arterial BP and pulse rate), including $SpO_2$ and body weight from baseline to all assessed time points during the study.
- Treatment-emergent marked laboratory abnormalities up to 30 days after study treatment discontinuation or until initiation of study drug in the AC-055H302 OL extension study, whichever occurs first
- Change in laboratory parameters from baseline to all assessed time points during the study.

All safety analyses will be based on the safety analysis set (SS) based on actual treatment received, unless otherwise specified. Listings will be provided on the FAS.

Unless otherwise specified, the treatment-emergent period is defined from study treatment start up to 30 days after study treatment discontinuation, limits included.

Safety assessments (e.g., laboratory, vital signs) on Day 1 are considered baseline. Safety events (i.e., adverse events) occurring/starting on Day 1 are considered treatment-emergent while safety signals (i.e., laboratory marked abnormalities, vital signs abnormalities) occurring/starting on Day 1 are not considered treatment-emergent.

For those subjects who entered the AC-055H302 / RUBATO OL extension study immediately after the last dose of study treatment in this AC-055H301 RUBATO study, AEs/SAEs or marked laboratory abnormalities after initiation of OL study treatment are reported and followed up in the OL extension study, and are not included in the analysis of the DB study.

For all continuous safety variables, descriptive statistics will include the N, mean, standard deviation [SD], median and range [minimum, and maximum], and IQ range. Categorical variables will be summarized using frequency counts and percentages.

## 6.1. Adverse Events

The original terms used by the investigators to describe AEs are assigned preferred terms (PT) and system organ class (SOC) for classification and tabulation using the latest implemented MedDRA version dictionary.

The following definitions are relevant for the analysis of the adverse events.

- Frequency of the adverse events

AEs reported more than once for a subject (as qualified by the same preferred term(s)) are counted in the frequency table once. If the reported AE is assigned to several preferred terms, subjects are counted for each individual preferred term.

- Intensity of the adverse events

For AEs reported more than once for a subject (as qualified by the same preferred term(s)), the worst outcome is considered. The categories of intensity are defined as follows: mild, moderate, severe, If the intensity is missing, the event is considered severe.

- Relationship of the adverse events

Relationship to study treatment is defined as 'related' or 'not related'. For AEs reported more than once for the same subject (as qualified by the same preferred term(s)), the worst relationship is considered. Adverse events with missing relationship are considered in any analysis as related.

- Serious adverse events

For the analysis, an AE is considered serious if the tick box 'Yes' for 'Serious?' is checked on the AE eCRF module. If the information on seriousness is missing, the adverse event is considered serious.

- Adverse Events of Special Interest

AEs of special interest (AESIs) are defined using a selection of preferred terms (Internal MedDRA Query). The full definition is included in Attachment 2:

The following groups of AEs of special interest will be summarized:

- Hepatic adverse events of special interest,
- Edema and fluid retention,
- Anemia / hemoglobin decrease,
- Hypotension.

Overall summary tables of AEs will be provided, containing number and percentages of subjects having experienced at least 1 occurrence for the following categories of AEs:

- Treatment-emergent AEs
- Severe treatment-emergent AEs
- Treatment-emergent AEs related to study treatment
- AEs with fatal outcome
- Treatment-emergent serious AEs (SAEs)
- Treatment-emergent SAEs related to study treatment
- AEs leading to premature discontinuation of study treatment

For each category of AEs of special interest, a similar overall summary table will be provided, containing number and percentages of subjects having experienced at least 1 occurrence for the above categories. In order to account for differences in the time of observation among subjects, this table will also present the subject-years of observation (SYO) and the incidence rate per 100 subject-years.

The subject-observation time will be calculated, for each subject, as follows:

i. For subjects without AESI: by considering the study treatment duration as (study treatment end date – study treatment start date + 1);

ii. For subjects with AESI: by considering the treatment duration up to the start date of first event (min [date of first event, study treatment end date] – study treatment start date + 1)

SYO will be calculated by first summing the subject-observation time for all subjects and then dividing the results by 365.25 days.

The incidence rate for an AESI category, per 100 subject-years. will be calculated by dividing the number of subjects with treatment-emergent AESI by the SYO and multiplying by 100.

$$\text{Adjusted Incidence Rate} = 100 \times (\text{Number of subjects with at least one AE/SYO})$$

Separate summary tables will be provided by SOC and PT within each SOC for the categories of AEs specified above (except for severe treatment-emergent AEs). The summary tables are presented in descending order according to the incidence in macitentan 10 mg treatment group (e.g., SOC and PT within each SOC with the highest number of occurrences appears first). Equal frequency of different SOC/PTs is sorted in alphabetical order of the SOC/PT.

Similarly, separate summary tables will be provided by PT, in descending order of incidence in the macitentan 10 mg treatment group for the categories of AEs specified above (except for severe treatment-emergent AEs) with the addition of:

- Treatment-emergent AEs by maximum intensity
- AEs of special interest.

A summary table will be provided, containing number and percentages of subjects who died on the SS. Frequencies for the following parameters will be included in the summary table:

- Number of subjects who died
- Primary cause of death

The summary will be based on the “Death” eCRF module.

The summary will be provided if at least 3 deaths have occurred overall. The summary table is presented by descending order of PT according to the incidence of deaths in the macitentan 10 mg treatment group.

In addition to the summary tables, separate listings will be provided for subjects who:

- Had AEs
- Had SAEs
- Had AEs leading to discontinuation of study treatment.
- Had AEs of special interest.
- Had died.

Treatment-emergent AEs (or deaths) are flagged accordingly.

## 6.2. Clinical Laboratory Tests

The analysis of the changes from baseline will be based on the laboratory test results provided by central laboratory at scheduled visits. For each post-baseline scheduled visit, only assessments performed after study treatment start date up to 30 days after study treatment discontinuation will be included. End-Of-Treatment visit is mapped to a scheduled visit (see Section 2.1.2). Values from unscheduled visits are not included in the summaries by visits but are included in the analysis for marked laboratory abnormalities (including additional liver and hemoglobin abnormalities, estimated Glomerular Filtration Rate and Protein Losing Enteropathy). Data are evaluated in Standard International System of Units (SI). In case of local laboratory, values (incl. reference ranges) are converted in SI units according to the Sponsor's internal guidelines. The tests converted in SI units are available in SDTM for the analysis. Local laboratory values will be included, together with central laboratory values, in the analyses of marked abnormalities only.

All values reported as below or above the limit of quantification (e.g., '< 3', '> 100') are substituted with the limit of quantification (e.g., '< 3', is substituted by '3') for the purpose of the analysis. The values are listed including the < or > sign.

If more than one value for a laboratory parameter is assessed on the same day, from central and local laboratory, the value from the central laboratory is considered for the analysis. If more than one value falls on the same time point (date and time) then the one with the last sequential number in SDTM will be used.

### 6.2.1. Laboratory parameters

Below the list of laboratory parameters with corresponding SI units.

**Hematology**

- Hemoglobin (g/L)
- Hematocrit (L/L)
- Erythrocyte count (reticulocyte count) ($10^{12}$/L)
- Leukocyte count with differential (neutrophils, lymphocytes, monocytes, eosinophils, basophils) counts ($10^9$/L).
- Platelet count ($10^9$/L).

**Clinical chemistry**

- Alanine aminotransferase (ALT) (U/L)
- Aspartate aminotransferase (AST) (U/L)
- Alkaline phosphatase (U/L)
- Total and direct bilirubin (µmol/L)
- Gamma glutamyl transpeptidase (U/L)
- Creatinine (µmol/L)
- Blood urea nitrogen (mmol/L)

- Uric acid (µmol/L)
- Glucose (mmol/L)
- Cholesterol (mmol/L)
- Triglycerides (mmol/L)
- Sodium, potassium, chloride, calcium (mmol/L)
- Serum albumin (g/L)
- Total protein (g/L)
- Alpha-fetoprotein (ug/L)
- Cystatin-C (mg/L).

**Coagulation tests**

- Prothrombin time and/or international normalized ratio.

All laboratory parameters (defined above) provided by the central and local laboratory will be listed together, including those from unscheduled visits. Marked laboratory abnormalities (MLAs) are flagged accordingly.

Changes from baseline to all scheduled visits will be summarized by treatment group, displaying descriptive statistics for hematology (including coagulation tests) and clinical chemistry. Similarly, changes from baseline to the 'last on-treatment assessment" and to the "last follow-up assessment" (see Section 2.1.2) will be summarized as well.

For selected laboratory tests (see list below), results will be visualized by means of scatter plots of values at baseline versus last-on treatment assessment. The scatter plot is displayed with the reference line y = x.

- Hematology: hemoglobin, hematocrit, leukocytes, platelets
- Clinical chemistry: ALT, AST, Alkaline phosphatase, total bilirubin, creatinine.

After consultation with the Study Results Committee constituted of external clinicians with expertise in subjects with Functional Single Ventricle Heart Disease on 13-Aug-2021, the following analyses of change from baseline in total bilirubin, ALT, and AST will be performed considering all observed values (from scheduled or unscheduled assessments) and the actual assessment day to assess and characterize the pattern of total bilirubin, ALT, and AST (separately) over time.

Gilbert's syndrome is a common inherited disease caused by a gene mutation whereby the liver is unable to properly process bilirubin, a byproduct of the breakdown of red blood cells. It results in elevated bilirubin levels. Given that subjects with Gilbert's Syndrome will have bilirubin elevations that are not necessarily caused by hepatic congestion, these subjects (as qualified by the preferred term 'Gilbert's syndrome' in the "Medical history" eCRF module) will be removed from the analyses of change in bilirubin, ALT and AST over time.

- Spaghetti plots will be presented by treatment group to explore the longitudinal pattern of observed values.
- A parametric longitudinal mixed-effects model, with a (fixed-effects) mean structure that is quadratic over time (with time corresponding to the actual assessment day) and with a random intercept and a random slope will be considered for changes from baseline to each post-baseline assessment. The model will include time, time-by-treatment interaction term, time-by-time-by- treatment quadratic interaction term, an indicator variable and the baseline value. Random intercept and slope at the level of the subject will be considered. Estimates by treatment groups at analysis visits Week 16 and Week 52, will be presented along with their 95% confidence intervals from the model. Graphical presentation of model estimates along with 95% confidence regions will be overlayed to the observed values of the change from baseline.

#### 6.2.1.1. Marked Laboratory Abnormalities

A laboratory test abnormality is defined as any value outside the normal range as provided by the central and local laboratory. The direction of the abnormality (below or above the normal range) is indicated using ‘H’ and ‘L’.

A marked laboratory abnormality (MLA) is defined as any value that fulfills the applicable condition for LL / HH, as provided in Table 12. More severe marked abnormalities are indicated by LLL / HHH, where applicable.

The definitions of marked abnormal values are based mainly on the Common Terminology Criteria for Adverse Events (CTCAE) (CTCAE 2010) grading system and, in specific cases (e.g., lymphocyte levels), are adjusted based on the known pharmacodynamic effect of the study drugs (e.g., LLL threshold for lymphocytes).

Treatment-emergent MLAs, including liver function test, hemoglobin, estimated glomerular filtration rate and protein losing enteropathy abnormalities, are those occurring after the study treatment start and up to 30 days after study treatment discontinuation, that were not present at baseline in the same or worse category (considering the direction of worsening). In other words, the MLAs are evaluated independently by direction of worsening (e.g., a post-baseline MLA of “HH” is considered treatment-emergent if the baseline is “L” or “LL” or “LLL” (where applicable) or “H”, or within normal limits or missing. On the other hand, it is not considered as treatment-emergent if the baseline is “HH” or “HHH” when applicable).

**Table 12: Threshold for Marked Laboratory Abnormalities**

| Parameter (SI unit) | LL | LLL | HH | HHH |
|---|---|---|---|---|
| Hemoglobin (g/L) | < 100<br>**ALERT:**<br>**< 100 (re-test)** | < 80<br>**ALERT:**<br>**< 80 (re-test)** | Increase in > 20 g/L above ULN or above baseline (if baseline is above ULN) | Increase in > 40 g/L above ULN or above baseline (if baseline is above ULN) |
| Hematocrit (L/L) | < 0.28 (female)<br>< 0.32 (male) | < 0.20 | > 0.55 (female)<br>> 0.60 (male) | > 0.65 |
| Platelet count ($10^9$ /L) | < 75 | < 50 | > 600 | > 999 |
| Eythrocyte count ($10^{12}$ /L) | ND | ND | ND | ND |
| Leukocyte count ($10^9$ /L) | NA | < 1.9 | > 20.0 | > 100.0 |
| Lymphocyte ($10^9$ /L) | ND | < 0.2 | > 4.0 | >20 |
| Neutrophils ($10^9$ /L) | < 1.5 | < 1.0 | ND | ND |
| Eosinophils ($10^9$ /L) | ND | ND | > 5.0 | ND |
| AST (U/L)* | ND | ND | ≥ 3 ULN<br>**ALERT:**<br>**≥ 3 ULN** (exclusion at baseline or re-test) | ≥ 5 ULN<br>**ALERT:**<br>**≥ 5 ULN** (re-test)<br>**≥ 8 ULN** (discontinuation) ** |
| ALT (U/L)* | ND | ND | ≥ 3 ULN<br>**ALERT:**<br>**≥ 3 ULN** (exclusion at baseline or re-test) | ≥ 5 ULN<br>**ALERT:**<br>**≥ 5 ULN** (re-test)<br>**≥ 8 ULN** (discontinuation) ** |
| Total bilirubin (umol/L) | ND | ND | ≥ 2 ULN<br>**ALERT:**<br>**≥ 2 ULN combined with ALT or AST ≥ 3 ULN** (discontinuation) | ≥ 5 ULN |
| Alkaline Phosphatase (U/L) | ND | ND | > 2.5 ULN | > 5 ULN |
| INR* | ND | ND | ≥ 1.5 ULN or<br>≥ 1.5 × above baseline if on anticoagulation | ≥ 2.5 ULN or<br>≥ 2.5 × above baseline if on anticoagulation |
| Creatinine (umol/L)* | ND | ND | >1.5 ULN or<br>>1.5 baseline (if baseline > ULN) | > 3 ULN or<br>>3 × baseline (if baseline > ULN) |

| Parameter (SI unit) | LL | LLL | HH | HHH |
|---|---|---|---|---|
| eGFR (mL/min/1.73 $m^2$) | < 60 | < 30 **ALERT: <30** (exclusion at baseline) | ND | ND |
| Blood Urea Nitrogen (mmol/L) | ND | ND | > 2.5 ULN | > 5 ULN |
| Albumin (g/L) | < 30 **ALERT: < 30 combined with total protein < 50** (exclusion at baseline) | < 20 | ND | ND |
| Glucose (mmol/L) | < 3.0 | < 2.2 | > 8.9 | >13.9 |
| Potassium (mmol/L) | < 3.2 | < 3.0 | > 5.5 | > 6.0 |
| Sodium (mmol/L) | ND | < 130 | > 150 | > 155 |
| Calcium (mmol/L) | < 2.0 | < 1.75 | > 2.9 | > 3.1 |
| Triglyceride (mmol/L) | ND | ND | > 3.42 | > 11.4 |
| Cholesterol (mmol/L)* | ND | ND | > 7.75 | > 12.92 |
| Serum pregnancy test | ND | ND | ND | Positive **ALERT: Positive** |

The above values come from the protocol Table 5 in Appendix 1 of the protocol.
* HH and HHH based on CTCAE 2010 v4.03 (CTCAE 2010)
**** Also HHHH as > 8 x ULN
For INR, a subject is considered to be on anticoagulation if the subject took antithrombotic agent (ATC codes B01A) and date of INR assessment is on or after start of anticoagulation therapy.
ALERT = study-specific alerts that trigger specific actions by the investigator [see Protocol Section 5.1.11; ALT = alanine aminotransferase; AST = aspartate aminotransferase; INR = International Normalized Ratio; MCH = mean corpuscular hemoglobin; MCV = mean corpuscular volume; may be complemented by definitions provided by the central laboratory (see central laboratory manual); NA = not applicable; ND = not defined; SI = international system of units; ULN = upper limit of normal.

For each category (i.e., LL, LLL, HH, HHH, HHHH), treatment-emergent MLAs (separately for hematology and clinical chemistry) will be summarized by treatment group, displaying counts and percentages of subjects with at least a treatment-emergent marked laboratory abnormality for each parameter for which the marked abnormality is defined (except for serum pregnancy test). MLAs related to coagulation tests will be included in the summary of hematology MLAs.

In addition, for each parameter for which the abnormalities are defined, the distribution of subjects at baseline versus the following categories will be displayed:

- worst low value up to 30 days after study treatment discontinuation,
- worst high value up to 30 days after study treatment discontinuation,
- "last on-treatment assessment" (see definition in Section 2.1.2).

If LL/LLL or HH/HHH are not defined for a parameter "NA" will appear in the table for the corresponding parameter/category.

All laboratory tests for subjects with at least one treatment-emergent MLAs are provided in a separate subject listing.

#### 6.2.1.2. Additional Liver Function Test and Hemoglobin Abnormalities

The following elevated liver test abnormalities are considered:

- ALT or AST ≥ 3 × ULN;
- ALT or AST ≥ 3 × ULN and < 5 × ULN;
- ALT or AST ≥ 5 × ULN and < 8 × ULN;
- ALT or AST ≥ 8 × ULN.
- ALT or AST ≥ 3 × ULN in combination (i.e., at the same visit) with total bilirubin ≥ 2 × ULN

The following hemoglobin abnormalities are considered:

- Hemoglobin < 80 g/L
- Hemoglobin ≥ 80 g/L and < 100 g/L
- Hemoglobin < 100 g/L

The highest ALT or AST value at any post-baseline time point of assessment up to 30 days after study treatment discontinuation is considered for classification in the categories above.

The lowest hemoglobin value at any post-baseline time point of assessment up to 30 days after study treatment discontinuation is considered, as defined above.

Treatment-emergent liver/hemoglobin abnormalities are summarized by treatment group, displaying counts and percentages of subjects with at least a treatment-emergent abnormality for each predefined abnormality.

Using the "evaluation of drug-induced serious hepatotoxicity" (eDISH) plots, graphical representations of maximum treatment-emergent ALT (in multiples of ULN) by maximum treatment-emergent total bilirubin (in multiples of ULN) will be produced, to identify potential Hy's Law cases for each subject. The graph will be on a log10 scale for all subjects having assessments > 0.0625 × ULN for both peaks. Two reference lines will be plotted identifying the 2 × ULN for total bilirubin and 3 × ULN for ALT. The normal subjects are on the left lower quadrant, the possible Hy's Law cases appear on the right upper quadrant. The peak is defined as the maximum value from study treatment start up to 30 days after study treatment discontinuation (for the same subjects, not necessarily the peak of ALT occurs at the same time of the peak of total bilirubin).

#### 6.2.1.3. Estimated glomerular filtration rate (eGFR) and Protein Losing Enteropathy (PLE)

The abnormality associated to the estimated creatinine clearance for determining glomerular filtration rate reflecting severe renal function is defined as eGFR < 30 mL/min/1.73m$^2$.

If creatinine clearance is estimated, the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation (Levey 2009) will be used for adults and the Bedside Schwartz equation (Schwartz 2009) for adolescents, for determining glomerular filtration rate (GFR) according to the formula in the protocol Section 7.2.4.2. The derived GFR is available in the SDTM.

The distribution of subjects in the following categories will be displayed by treatment group:

- G1 (Normal or high): ≥ 90 mL/min/1.73m$^2$
- G2 (Mildly decreased): 60-89 mL/min/1.73m$^2$
- G3a (Mildly to moderately decreased): 45-59 mL/min/1.73m$^2$
- G3b (Moderately to severely decreased): 30-44 mL/min/1.73m$^2$
- G4 (Severely decreased): 15-29 mL/min/1.73m$^2$
- G5 (Kidney failure): < 15 mL/min/1.73m$^2$

The lowest eGFR value at any post-baseline time point of assessment up to 30 days after study treatment discontinuation (that was not present at baseline) is considered.

The protein losing enteropathy is defined as:

- Albumin < 30 g/L in combination (i.e., at the same visit) with total protein < 50 g/L.

Treatment-emergent eGFR or PLE abnormalities are summarized by treatment group, displaying counts and percentages of subjects with at least one abnormality.

## 6.3. Vital Signs and Physical Examination Findings

Vital signs parameters include: systolic and diastolic blood pressure (mmHg), pulse rate (bpm), $SpO_2$ (%), height (cm), and body weight (kg).

The analysis of the changes from baseline will be based on results at scheduled visits. For each post-baseline scheduled visit, only assessments performed after study treatment start date up to 30 days after study treatment discontinuation will be included. End-of-Treatment (if any) visit is mapped to a scheduled visit (see Section 2.1.2). Values from unscheduled visits are not included in the summaries by visits but are included in the analysis for abnormality.

Changes from baseline to all scheduled visits will be summarized by treatment group, displaying descriptive statistics for systolic and diastolic blood pressure (mmHg), pulse rate (bpm), $SpO_2$ (%), and body weight (kg). Similarly, changes from baseline to the 'last on-treatment assessment" and to the "last follow-up assessment" (see Section 2.1.2) will be summarized as well. Height will be summarized only at baseline together with the demographic characteristics.

Box plots will be presented to graphically display the distribution of the change from baseline in vital signs (BP, pulse rate, $SPO_2$, and body weight) over time (i.e., at each scheduled visit from Week 1 to Week 52).

The number and percentages of subjects with the following abnormality for the systolic blood pressure: < 90 mmHg for adults (or < 85 mmHg for adolescents (<18 years old) who are < 150 cm

in height) at any post-baseline assessment (scheduled or unscheduled) after the study treatment start up to 30 days after study treatment discontinuation (that were not present at baseline), will be presented by treatment group.

A listing of all vital sign measurements will be provided. Position (supine or sitting) and location (left or right arm) for vital signs will be included as well as the flag for abnormal systolic blood pressure.

Physical examinations performed during the course of the study are reported in a subject listing, with associated general assessment. If an abnormality is found, further details describing the signs and symptoms related to the abnormality are specified.

## 6.4. Electrocardiogram

No direct reporting of ECG results is planned.

An ECG result is reported as an AE in the case of an abnormal assessment representing a clinically significant finding that was either not present at baseline, or worsened after the start of study treatment.

## 6.5. Other Safety Parameters

### 6.5.1. Echocardiography

Single ventricle ejection fraction (%), Fontan circulation stenosis value (%), ventricular outflow to aorta doppler velocity (m/s) and valvular defects [vena contracta diameter (mm) and regurgitant orifice ($cm^2$)] are measured at screening and End of Treatment (EOT) visit and reported in the eCRF.

Severe AV valve regurgitation is present where the vena contracta is ≥ 7 mm or the effective regurgitant orifice is ≥ 0.4 $cm^2$. In the eCRF, up to two valve regurgitation values may be collected. If both are present in the eCRF, the largest will be considered for the analysis.

Presence of left ventricular outflow obstruction is confirmed by a pulsed-wave Doppler velocity > 2.5 m/s.

The analysis of echocardiography will be based on results at scheduled visits.

Changes from baseline (i.e.,screening) to Week 52/EOT in Fontan circulation stenosis (%) and ejection fraction (%) will be summarized by treatment group, displaying descriptive statistics for continuous variables.

Valvular defects (identified by severe AV valve regurgitation and presence of left ventricular outflow obstruction) will be summarized at baseline (i.e.screening) and Week 52/EOT by treatment group using frequency counts and percentages.

Individual listing of echocardiographic parameters will be provided.

### 6.5.2. Pregnancy Test

For females of childbearing potential, a serum pregnancy test is performed at Screening (Visit 1), and urine pregnancy tests at Randomization (Visit 2), and every 4 weeks (+/−7 days) thereafter up to EOS. In addition, a serum pregnancy test is performed at any time if pregnancy is suspected during the study.

Pregnancies must be entered in the AE page of the eCRF. Methods of contraception are collected in the Contraceptive Methods Log page of the eCRF. Change in childbearing potential status is collected in the Childbearing Potential Status Change page of the eCRF.

Childbearing potential status, contraception and pregnancy test data are reported in subject's listings.

## 7. QUALITY OF LIFE

Quality of life (QoL) parameters will be assessed for all subjects by 2 sets of QoL questionnaires including the EQ-5D-5L and the patient/clinician assessment of severity and assessment of change (PGA-S, PGI-C and CGI-S, CGI-C, respectively). These will be completed on paper, and site staff will enter the corresponding values in the eCRF.

The PGA-S and PGI-C are completed by subjects, while the CGI-S and CGI-C are completed by clinicians (a physician, either an investigator or sub-investigator). Completion should occur at the beginning of the visit.

Disease Severity (PGA-S and CGI-S) is assessed at Randomization (Visit 2), Week 16 (Visit 4), and Week 52 (Visit 6), or EOT in case it occurs prior to Week 52.

Impression of Change (PGI-C and CGI-C) is assessed at Week 16 (Visit 4), and Week 52 (Visit 6), or EOT in case it occurs prior to Week 52.

The EQ-5D-5L questionnaires are completed by subjects at Randomization (Visit 2), Week 16 (Visit 4), Week 52 (Visit 6), or EOT in case it occurs prior to Week 52. If subjects enter the PTOP, the questionnaire will be also be completed using designated scripts.

The following QoL endpoints are considered:

- Patient Global Assessment of Disease Severity (PGA-S) and Clinician Global Impression of Severity (CGI-S) at each scheduled visit (baseline, Visit 4/Week 16, and Visit 6/Week 52),
- Patient Global Impression of Change (PGI-C) and Clinician Global Impression of Change (CGI-C) questionnaire at each post-baseline scheduled visit (Visit 4/Week 16, and Visit 6/Week 52) ,
- Proportion of subjects within each response category for each dimension (i.e., mobility, self-care, usual activities, pain/discomfort, anxiety/depression) of the Euro QoL 5D (EQ-5D-5L) questionnaire at baseline and post-baseline scheduled visit (Visit 4/Week 16 and Visit 6/Week 52),

- Change from baseline to each post-baseline scheduled visit (Visit 4/Week 16 and Visit 6/Week 52) in self-rated health status score assessed by the EQ-5D-5L questionnaire.

The QoL endpoints will be analyzed descriptively on the FAS by randomized treatment group .

The analysis of the quality of life endpoints will be based on results at scheduled visits. For each post-baseline scheduled visit, only assessments performed after study treatment start date will be included. End-of-Treatment (if any) visit is mapped to a scheduled visit (see Section 2.1.2). Values from unscheduled visits are not included in the summaries by visits.

For PGA-S, CGI-S, and each dimension of EQ-5D-5L the proportion of subjects within each response category will be summarized at baseline and at each post-baseline visit.

Each dimension of EQ-5D-5L questionnaire is divided into 5 levels of perceived problems: Level 1 indicating no problems, Level 2 indicating slight problems, Level 3 indicating moderate problems, Level 4 indicating severe problems, and Level 5 indicating extreme problems. The frequency distribution of subject's responses over time will be provided in a graphical representation (i.e. bar plot) for ach treatment group. The percentages of subjects reporting problems will be summarized by grouping responses from Level 2 to Level 5 for each dimension of EQ-5D-5L questionnaire.

For PGI-C and CGI-C, the proportion of subjects within each response category will be summarized at each post-baseline visit.

The self-rated health status score assessed by the EQ-5D-5L questionnaire will be summarized at baseline, at each post-baseline visit as well as for the absolute change from baseline to each post-baseline visit.

Individual subject listings will be provided for each QoL questionnaire.

# 8. PHARMACOKINETICS/PHARMACODYNAMICS

## 8.1. Pharmacokinetics

Trough (pre-dose) plasma concentrations of macitentan and its metabolite ACT-132577, are obtained via blood samples taken at Week 4 (Visit 2, during the regular corresponding monthly blood draw) and Week 16 (Visit 4), or at EOT in the case of premature study treatment discontinuation prior to Week 16.

PK analyses will be performed on the PKS. Reasons for exclusion from PKS will be summarized by treatment group on the FAS.

Trough macitentan and ACT-132577 analyte concentrations will be summarized by visit (i.e., Week 4, Week 16/EOT) by descriptive statistics (i.e., N, mean, standard deviation [SD], range [minimum and maximum], coefficient of variation (%), and geometric mean). A subject will be included in the analysis at each visit if the corresponding plasma concentration value is flagged

at "steady-state", as reported in the SDTM, and if the corresponding sample is not taken after the morning dose of study treatment (i.e., without PD_PM.364 for the visit under consideration).

Up to protocol version 6, PK sampling were taken at Visit 3/Week 8. With protocol version 7, Week 8 was converted to a telephone call therefore PK sampling is taken during Visit 2/Week 4. For the evaluation at Week 4, the value at Week 8 (if available) will replace the value at Week 4 (if not available).

In case of premature treatment discontinuation prior to Visit 4/Week 16, the value at EOT will replace the value at Week 16 (if not available).

Similarly, trough macitentan and ACT-132577 analyte concentrations will be summarized by age group (i.e., adolescents: 12-18 yrs vs adults: ≥ 18 yrs), by sex (i.e., male vs female), by body weight at baseline (< 50 kg, >=50kg) as well as by age group * body weight at baseline.

Individual subject listing will be provided.

## 9. HEALTH ECONOMICS

The following pharmacoeconomic endpoints are considered:

- Number per year of all-cause and Fontan-related hospitalizations from baseline up to EOT.
- Number per year of in-patient hospital days for all causes and Fontan-related causes from baseline up to EOT.

The number of all-cause and Fontan-related hospitalizations will be calculated during the double-blind treatment period (defined as the period between the treatment start date, see Section 2.5.2, and the EOT, see Section 2.5.3, limits included). The number per year of hospitalizations will be calculated as follows:

- at group-level: total number of hospitalizations across subjects divided by the total study treatment duration (subjects-year)
- at subject-level: (number of hospitalizations for each subject / individual treatment duration in days) * 365.25.

The number per year of hospital days (all-cause and Fontan-related) will be calculated (at subject-level) as:

(total number of hospital days for each subject / individual treatment duration in days) * 365.25.

For each hospitalization, the number of hospital days is calculated as discharge date minus admission date plus 1. For each subject, the number of days spent in hospital is the sum of the number of days in hospital for all his/her hospitalizations up to EOT.

Number per year of all-cause and Fontan-related hospitalizations and hospital days will be summarized descriptively on the FAS by randomized treatment group.

Individual subject listing will be provided.

# 10. CHANGES OR CLARIFICATIONS TO ANALYSES PLANNED IN THE STUDY PROTOCOL

## 10.1. Changes to the Statistical Analysis Plan for Special Protocol Submission

SAP version 3 finalized on 25-Feb-2021 has been developed based on the document that was finalized and sent to FDA Agency for the SPA re-submission of protocol version 8 (SAP for SPA Final version 2 (D-20.077)).

The analyses for the primary and secondary efficacy endpoints were already described in Section 4 of the SAP for SPA, which are currently under Section 5 of this document.

All the analyses which will be performed for the CSR purpose have been added in this document and outlined below. Any change, addition or clarification compared to the SAP for SPA for the primary and secondary efficacy endpoints are detailed below.

- Section 4: Subject Information (i.e., demographics and baseline disease characteristics, medical history, screening failures, study and study treatment disposition, treatment compliance, extent of exposure, protocol deviations, and prior and concomitant medications)

- Section 5: Analyses of primary and secondary efficacy endpoints as in the SAP for SPA (see changes below) with the addition of the analysis of other efficacy variables

    - Primary and secondary efficacy endpoints:

        - the estimators based on the inverse normal combination method with weights proportional to the final actual sample size (i.e., 142 subjects) have been added in this document (Section 5.2.3.4, Section 5.3.1.6, and 5.3.2.6 for primary and secondary endpoint respectively) to assess the robustness of result when the weights are proportional to the actual sample size of 142 subjects.

        - the estimator based on different model 1 (Section 5.2.3.4), by adding treatment-by-geographical region interaction to the main model, has been deleted as the treatment-by-geographical region interaction is already investigated in the subgroup analysis as described in Section 5.2.3.5.

    - Secondary efficacy endpoint: *Change from baseline over 52 weeks in peak $VO_2$*

      The estimator based on ANCOVA has been clarified for the handling of intercurrent events (i.e., premature study treatment discontinuation was not spelled out clearly as an event in the SAP for SPA) and the rules for missing data were added in line with the primary estimand of this endpoint.

      The estimator based on the number of subjects with available Week 52 data pre- and post- Sponsor decision (i.e., 24 January 2019) to implement the IA as present in the SAP for SPA (Section 4.3.1.6) has been deleted as almost no subject would have Week 52 data before January 2019, being the first subject randomized in December 2017.

- Section 6: Safety
- Section 7: Quality of life
- Section 8: Pharmacokinetic
- Section 9: Health Economics

## 10.2. Changes to the analyses planned in the study protocol

✓ Efficacy

- The estimator based on different model 1 (Section 5.2.3.4), by adding treatment-by-geographical region interaction to the main model, has been deleted as the treatment-by-geographical region interaction is already investigated in the subgroup analysis as described in Section 5.2.3.5.

✓ Safety

- Adverse events: The analysis of treatment-emergent non-serious AEs (protocol Section 10.3.4.1) will not be part of this document. The list of results to be posted to the clinical registry will be detailed in the Data Presentation Specification Part 1 (Section 1.3).
- Treatment-emergent marked laboratory abnormalities (protocol Section 10.3.4.2): shift tables of categories of MLAs at baseline versus post-baseline visits, worst low, worst high and last post-baseline have been clarified to be with shift tables of categories of MLAs at baseline vs worst low value (up to 30 days after study treatment discontinuation), worst high value (up to 30 days after study treatment discontinuation), "last on-treatment assessment" value.

  Reason: In order to have a consistent overview of LMAs, the shift tables are aligned to the evaluation of treatment-emergent LMAs during the treatment period up to 30 days after study treatment discontinuation for the worst low and worst high value, without the need to look at each intermediate post-baseline visits. The evaluation on last on treatment assessment has been added.

## 10.3. Clarifications concerning endpoint definitions and related variables or statistical methods

✓ Analysis sets

The protocol Section 10.1 namely calls "Screened Analysis set (SCR)", "Per-Protocol Set (PPS)". These analysis sets are renamed "All Enrolled Set (ENR)", "Per-Protocol Set (PP), respectively [refer to Section 2.2] in accordance with current terminology/standards in Janssen.

The definition of SCR remains unchanged.

The definition of the PP has been detailed for each primary and secondary efficacy endpoint in line with the estimand language of each endpoint. This will allow to assess the robustness of results on a different population (therefore additional estimand) compared to the main

estimand. Therefore the PP has been replaced with PP1, PP2, PP3 for the primary and two secondary efficacy endpoints, respectively.

✓ Other efficacy endpoints

**Physical activity**
The analysis of the mean time spent in minutes in each of the category of physical activity have been clarified as compared to protocol Section 10.3.5.3. The mean time will be summarized as a continuous variable by displaying, for each treatment group, descriptive statistics for the absolute values at baseline, Week 16, and Week 52 as well as for the absolute change from baseline. Furthermore, the percentage of daily time spent in each category will be summarized similarly including the graphical display (i.e. bar plot) at each time point of assessment.

The number of subjects who spent time in physical activity categories will not be summarized in shift tables (categories at baseline vs categories at post-baseline visits) as indeed it is stated in protocol Section 10.3.5.3. Each subject spends the daily time in one or more categories therefore the shift tables should make assumption for the selection of which category should be compared at baseline vs post-baseline. It is indeed of interest to summarize the percentage of daily time spent in each category as continuous variable at baseline and at each timepoint of assessment, including the changes from baseline.

✓ Safety

To monitor liver function test and hemoglobin abnormalities, specific categories of abnormalities have been better defined in Section 6.2.1.2 of this document as well as in Section 6.2.1.3 for eGFR and PLE.

## REFERENCES

ADDPLAN Inc. an Aptiv Solutions Co. (2014). ADDPLAN BASE version 6.1 User Manual. Aptiv Solutions, Cologne, Germany. 2014.

Collett D (2003) Modelling Survival Data in Medical Research, 2nd Edition. Chapman & Hall; 2003.

CTCAE (2010) Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0 Published: May 28, 2009 (v4.03: June 14, 2010) U.S.Department Of Health And Human Services. National Institutes of main Health National Cancer Institute.

Cuzick J (2005). Forest plots and the interpretation of subgroups. Lancet 2005;365:1308.

Evenson KR (2008), Catellier DJ, Gill K, Ondrak KS, McMurray RG. Calibration of two objective measures of physical activity for children, Journal of Sports Sciences, 2008;26:14,1557-1565.

Freedson PS (1998), Melanson E, Sirard J. Calibration of the Computer Science and Pallications, Inc. accelerometer. Official Journal of the American College of Sports Medicine 1998;777-781.

Hebert A (2013), Jensen AS, Idorn A et al. The effect of bosentan on exercise capacity in fontan patients; rationale and design for the TEMPO study. BMC Cardiovascular Disorders 2013; 13: 36-42.

Hebert A (2014), Mikkelsen UR, Thilen U et al. Bosentan Improves Exercise Capacity in Adolescents and Adults After Fontan Operation - The TEMPO (Treatment With Endothelin Receptor Antagonist in Fontan Patients, a Randomized, Placebo-Controlled, Double-Blind Study Measuring Peak Oxygen Consumption Study. Circulation 2014; 130: 2021-30.

Lehmacher W (1999), Wassmer G. 1999. Adaptive sample size calculations in group sequential trials. Biometrics 55(4):1286–90.

Levey AS (2009), Stevens LA, Schmid CH, et al. CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration). A new equation to estimate glomerular filtration rate. Ann Intern Med. 2009 May 5;150(9):604-12.

Pocock SJ (2002), Clayton TC, Altman DG. Survival plots of time-to-event outcomes in clinical trials: good practice and pitfalls, Lancet. 2002;359:1686 89.

Schwartz GJ (2009), Work DF. Measurement and estimation of GFR in children and adolescents. Clin J Am Soc Nephrol. 2009 Nov;4(11):1832-43.

Troiano RP (2008), Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Phyisical activity in the United States measured by accelerometer. Med Sci Sports Exerc. 2008 Jan;40 (1):181-8.

# ATTACHMENTS

## ATTACHMENT 1 - BMI-FOR-AGE REFERENCE

BMI-for-age cutoffs for adolescents were calculated based on the WHO Child Growth Standards (2007) developed using data collected in the WHO Multicentre Growth Reference study.

According to WHO classification an adult is classified underweight, normal weight, overweight, obese class I, obese class II, obese class III according to BMI cutoffs below:

| BMI ($kg/m^2$) | Classification |
|---|---|
| < 18.5 | Underweight |
| 18.5 - 25 | Normal weight |
| 25 - 30 | Overweight |
| 30 - 35 | Class I obese |
| 35 - 40 | Class II obese |
| ≥ 40 | Class III obese |

The z-score value associated with a BMI of 18.5, 25, 30, 35, and 40 for a 18 year subject was used as the reference for the calculation of the equivalent BMI in adolescents.

Therefore an adolescent (boy or girl) between 12 and 18 years is classified underweight, normal weight, overweight, obese class I, obese class II, obese class III according to BMI cutoffs below:

| Sex | Age (years) | Classification | BMI ($kg/m^2$) |
|---|---|---|---|
| Boy | 12 | Underweight | <15.41 |
| | 13 | Underweight | <15.98 |
| | 14 | Underweight | <16.58 |
| | 15 | Underweight | <17.16 |
| | 16 | Underweight | <17.68 |
| | 17 | Underweight | <18.13 |
| | 18 | Underweight | <18.50 |
| Girl | 12 | Underweight | <16.18 |
| | 13 | Underweight | <16.83 |
| | 14 | Underweight | <17.40 |
| | 15 | Underweight | <17.86 |
| | 16 | Underweight | <18.17 |
| | 17 | Underweight | <18.37 |
| | 18 | Underweight | <18.50 |

| Sex | Age (years) | Classification | BMI ($kg/m^2$) |
|---|---|---|---|
| Boy | 12 | Normal weight | ≥ 15.41 and < 20.22 |
| | 13 | Normal weight | ≥ 15.98 and < 21.13 |
| | 14 | Normal weight | ≥ 16.58 and < 22.06 |
| | 15 | Normal weight | ≥ 17.16 and < 22.93 |
| | 16 | Normal weight | ≥ 17.68 and < 23.72 |
| | 17 | Normal weight | ≥ 18.13 and < 24.41 |
| | 18 | Normal weight | ≥ 18.50 and < 25 |
| Girl | 12 | Normal weight | ≥ 16.18 and < 21.41 |

| | | | |
|---|---|---|---|
| | 13 | Normal weight | ≥ 16.83 and < 22.39 |
| | 14 | Normal weight | ≥ 17.40 and < 23.27 |
| | 15 | Normal weight | ≥ 17.86 and < 23.96 |
| | 16 | Normal weight | ≥ 18.17 and < 24.45 |
| | 17 | Normal weight | ≥ 18.37 and < 24.78 |
| | 18 | Normal weight | ≥ 18.50 and < 25 |

| Sex | Age (years) | Classification | BMI (kg/m$^2$) |
|---|---|---|---|
| Boy | 12 | Overweight | ≥ 20.22 and < 24.65 |
| | 13 | Overweight | ≥ 21.13 and < 25.87 |
| | 14 | Overweight | ≥ 22.06 and < 27.0 |
| | 15 | Overweight | ≥ 22.93 and < 27.98 |
| | 16 | Overweight | ≥ 23.72 and < 28.81 |
| | 17 | Overweight | ≥ 24.41 and < 29.48 |
| | 18 | Overweight | ≥ 25 and < 30 |
| Girl | 12 | Overweight | ≥ 21.41 and < 25.98 |
| | 13 | Overweight | ≥ 22.39 and < 27.19 |
| | 14 | Overweight | ≥ 23.27 and < 28.22 |
| | 15 | Overweight | ≥ 23.96 and < 28.99 |
| | 16 | Overweight | ≥ 24.45 and < 29.52 |
| | 17 | Overweight | ≥ 24.78 and < 29.82 |
| | 18 | Overweight | ≥ 25 and < 30 |

| Sex | Age (years) | Classification | BMI (kg/m$^2$) |
|---|---|---|---|
| Boy | 12 | Obese class I | ≥ 24.65 and < 30.22 |
| | 13 | Obese class I | ≥ 25.87 and < 31.77 |
| | 14 | Obese class I | ≥ 27.0 and < 32.97 |
| | 15 | Obese class I | ≥ 27.98 and < 33.82 |
| | 16 | Obese class I | ≥ 28.81 and < 34.41 |
| | 17 | Obese class I | ≥ 29.48 and < 34.79 |
| | 18 | Obese class I | ≥ 30 and < 35 |
| Girl | 12 | Obese class I | ≥ 25.98 and < 31.26 |
| | 13 | Obese class I | ≥ 27.19 and < 32.62 |
| | 14 | Obese class I | ≥ 28.22 and < 33.69 |
| | 15 | Obese class I | ≥ 28.99 and < 34.43 |
| | 16 | Obese class I | ≥ 29.52 and < 34.85 |
| | 17 | Obese class I | ≥ 29.82 and < 35 |
| | 18 | Obese class I | ≥ 30 and < 35 |

| Sex | Age (years) | Classification | BMI (kg/m$^2$) |
|---|---|---|---|
| Boy | 12 | Obese class II | ≥ 30.22 and < 37.79 |
| | 13 | Obese class II | ≥ 31.77 and < 39.67 |
| | 14 | Obese class II | ≥ 32.97 and < 40.58 |
| | 15 | Obese class II | ≥ 33.82 and < 40.81 |
| | 16 | Obese class II | ≥ 34.41 and < 40.70 |
| | 17 | Obese class II | ≥ 34.79 and < 40.41 |
| | 18 | Obese class II | ≥ 35 and < 40 |
| Girl | 12 | Obese class II | ≥ 31.26 and < 37.56 |

| | | | |
|---|---|---|---|
| | 13 | Obese class II | ≥ 32.62 and < 38.94 |
| | 14 | Obese class II | ≥ 33.69 and < 39.85 |
| | 15 | Obese class II | ≥ 34.43 and < 40.35 |
| | 16 | Obese class II | ≥ 34.85 and < 40.48 |
| | 17 | Obese class II | ≥ 35 and < 40.32 |
| | 18 | Obese class II | ≥ 35 and < 40 |

| Sex | Age (years) | Classification | BMI (kg/m$^2$) |
|---|---|---|---|
| Boy | 12 | Obese class III | ≥ 37.79 |
| | 13 | Obese class III | ≥ 39.67 |
| | 14 | Obese class III | ≥ 40.58 |
| | 15 | Obese class III | ≥ 40.81 |
| | 16 | Obese class III | ≥ 40.70 |
| | 17 | Obese class III | ≥ 40.41 |
| | 18 | Obese class III | ≥ 40 |
| Girl | 12 | Obese class III | ≥ 37.56 |
| | 13 | Obese class III | ≥ 38.94 |
| | 14 | Obese class III | ≥ 39.85 |
| | 15 | Obese class III | ≥ 40.35 |
| | 16 | Obese class III | ≥ 40.48 |
| | 17 | Obese class III | ≥ 40.32 |
| | 18 | Obese class III | ≥ 40 |

# ATTACHMENT 2 - ADVERSE EVENTS OF SPECIAL INTEREST

Adverse events of special interest are defined as follows:

## I. Hepatic events of special interest

AEs are included in this grouping if their coded PTs are included in the "Hepatic disorders" SQM with including all its sub-SMQ with the exception of "Liver-related coagulation and bleeding disturbances (SMQ)".

## II. Edema and fluid retention

AEs are included in this subgroup if their coded PTs is "Pulmonary congestion" or if within the SMQ "Haemodynamic oedema, effusions and fluid overload (SMQ)" with the exception of PTs containing "site".

## III. Anemia / hemoglobin decrease

AEs are included in this grouping if their coded PTs are included in the SMQs "Haematopoietic erythropenia" OR "Haematopoietic cytopenias affecting more than one type of blood cell (SMQ)" (with the exception of two unspecific PTs: "blood disorder", "blood count abnormal") OR an event with any MedDRA PT containing the text "anaemia".

## IV. Hypotension

AEs are included in this grouping if their coded PTs are: Blood pressure ambulatory decreased, Blood pressure decreased, Blood pressure diastolic decreased, Blood pressure immeasurable, Blood pressure orthostatic decreased, Blood pressure systolic decreased, Blood pressure systolic inspiratory decreased, CT hypotension complex, Diastolic hypotension, Hypotension, Mean arterial pressure decreased, Neonatal hypotension, Orthostatic hypotension, Postoperative hypotension, Procedural hypotension.

# ATTACHMENT 3 - PROTOCOL DEVIATIONS CLASSIFICATION FROM THE INDEPENDENT TEAM

memo

To PPD
From
Date 5 February 2021

An independent team with representatives from clinical sciences PPD and biostatistics PPD PPD reviewed the Protocol Deviation List (Version 6, dated 30 Nov 2020) for Study AC-055H301 (RUBATO DB). The aim of the review was to:

**1) Identify protocol deviations (PDs) that lead to the exclusion of subjects/data from the per-protocol analysis set**

**2) Identify major PDs that need to be summarized and discussed in the CSR as per ICH E3**

The RUBATO DB study follows a hybrid of Actelion/J&J processes for the collection, reporting and analysis of PDs. At initiation of the study, a comprehensive list of PDs list was developed in accordance with Actelion SOPs. This list captured ALL possible deviations. Following this approach, ALL protocol deviations are collected and categorized as follows

- Important protocol deviations (Yes/No)
- Timing (at screening, at study entry, during treatment period and follow-up)

For the reporting of PDs in the CSR, the team wanted to align with the J&J approach and **summarize and discuss only major protocol deviations in the CSR, and group them into the following categories:**

- Developed withdrawal criteria but not withdrawn
- Entered but did not satisfy criteria
- Received a disallowed concomitant treatment
- Received wrong treatment or incorrect dose
- Other

Several TCs were conducted between 18-Jan-2021 and 04-Feb-2021 to undertake this review. In addition, we also had a TC with BRQC representatives PPD on 25-Jan-2021 to discuss the possibility to reconsider PDs categorized as "Important" in the PD list as "Non-Major". After internal consultation, BRQC's recommendation was to remain consistent with how the PDs were captured during the trial and reported at the site level and communicated to EC/IRBs per local requirements, i.e. to consider all "Important" PDs as "Major" PDs for the CSR. We agreed to follow this advice i.e. for the reporting of PDs in the CSR, the term "Important PD" will be considered equivalent or synonymous to "Major PD".

PDs selected for exclusion from per-protocol analyses were identified in joint discussion between Sanjeev and Verena and are documented in the attached spreadsheet Rubato Independent PD Classification_Final.xlsx. In addition, we categorized all PDs into one of the 5 categories listed above, but suggested to change the wording of the first category from "Developed withdrawal criteria but not withdrawn" to "Developed withdrawal/interruption criteria but not withdrawn/interrupted". This is reflected in the same spreadsheet

PPD

**Actelion Pharmaceuticals Ltd**
Gewerbestrasse 16 | CH-4123 Allschwil | Switzerland | phone +41 61 565 65 65 |
www.actelion.com
Address for visitors: Hegenheimermattweg 95

**Protocol deviations**

Classification based on RUBATO PD list V6

| Condition | Identifier | Important | Categorization of PD (joint independent review) | Exclude from PP (joint independent review) | PP details (joint independent review) |
|---|---|---|---|---|---|
| *Informed consent not personally signed and dated by subject /legal representative* | *PD_MM.101* | *Yes* | Other | Yes | all PP |
| *Study assessments performed but no informed consent signed* | *PD_MM.102* | *Yes* | Other | No | |
| *Informed consent process (including re-consenting) not followed* | *PD_MM.103* | *Yes* | Other | No | |
| *ECG not assessed by cardiologist* | *PD_MM.104* | *Yes* | Other | No | |
| *ECHOC not performed according to guidelines* | *PD_MM.105* | *No* | Other | No | |
| *SAE not reported within 24 hours of knowledge, or not reported as per protocol* | *PD_MM.106* | *Yes* | Other | No | |
| *Other violation at screening not listed above* | *PD_MM.107* | *No* | Other | No | |
| *SAE not reported within 24 hours of knowledge, or not reported as per protocol* | *PD_MM.109* | *Yes* | Other | No | |
| *Other violation at study entry not listed above* | *PD_MM.110* | *No* | Other | No | |
| *Non-justified treatment code was break before unblinding of the study* | *PD_MM.111* | *Yes* | Other | Yes | all PP |
| *Subject has developed a condition that interferes with CPET assessment* | *PD_MM.112* | *Yes* | Other | Yes | PP1 (if PD linked to visit 4) and PP2 (if PD linked to visit 6). As per info from DM, a visit should always be linked to this PD. |
| *Visit 3 vital signs assessed under different conditions than at baseline* | *PD_MM.113* | *No* | Other | No | |
| *CPET at Visit 4 conducted under different conditions than at baseline* | *PD_MM.114* | *Yes* | Other | Yes | PP1 |
| *CPET at Visit 4 done, but not valid/questionable on review by central reading facility* | *PD_MM.115* | *Yes* | Other | Yes | PP1 |
| *CPET at Visit 4 not done according to guidelines and central reader manual* | *PD_MM.116* | *Yes* | Other | No | |
| *Visit 4 vital signs assessed under different conditions than at baseline* | *PD_MM.117* | *No* | Other | No | |
| *CPET at Visit 5 conducted under different conditions than at baseline* | *PD_MM.118* | *Yes* | Other | No | |
| *CPET at Visit 5 not done according to guidelines and central reader manual* | *PD_MM.119* | *Yes* | Other | No | |
| *Visit 5 vital signs assessed under different conditions than at baseline* | *PD_MM.120* | *No* | Other | No | |
| *CPET at Visit 6 conducted under different conditions than at baseline* | *PD_MM.121* | *Yes* | Other | Yes | PP2 |
| *CPET at Visit 6 not done according to guidelines and central reader manual* | *PD_MM.122* | *Yes* | Other | No | |
| *Visit 6 vital signs assessed under different conditions than at baseline* | *PD_MM.123* | *No* | Other | No | |
| *Study medication not stored correctly but subject continued to use it* | *PD_MM.124* | *Yes* | Received wrong treatment or incorrect dose | Yes | All PP. If PD date avaialble, only exclude from PP1 and PP3 if PD date prior to visit 4 efficacy assessment, and from PP2 if PD date prior to visit 6 efficacy assessment. |
| *ECHOC at EOT not performed according to guidelines* | *PD_MM.125* | *No* | Other | No | |
| *SAE not reported within 24 hours of knowledge, or not reported as per protocol* | *PD_MM.126* | *Yes* | Other | No | |
| *Other violation during treatment period and follow-up not listed above* | *PD_MM.127* | *No* | Other | No | |
| *Study drug not withheld prior to assessments at Visit 3* | *PD_MM.128* | *No* | Other | No | |
| *Study drug not withheld prior to assessments at Visit 4* | *PD_MM.129* | *No* | Other | No | |
| *Study drug not withheld prior to assessments at Visit 5* | *PD_MM.130* | *No* | Other | No | |
| *Study drug not withheld prior to assessments at Visit 6* | *PD_MM.131* | *No* | Other | No | |
| *CPET performed out of window of last QA (> 213 days since last QA)* | *PD_MM.132* | *No* | Other | No | |
| *Lab report not signed within 5 days of receipt* | *PD_MM.133* | *No* | Other | No | |
| *CPET data not sent to central reading facility within 5 working days* | *PD_MM.134* | *No* | Other | No | |

| | | | | | |
|---|---|---|---|---|---|
| *CPET data not sent to central reading facility within 5 working days* | *PD_MM.135* | *No* | Other | No | |
| *Lab report not signed within 5 days of receipt* | *PD_MM.136* | *No* | Other | No | |
| *Unable to understand and comply with the instructions, and to physically perform the CPET.* | *PD_MM.140* | *Yes* | Entered but did not satisfy criteria | Yes | PP1 and PP2 |
| *Woman of childbearing potential who did not have a negative serum pregnancy test at the Screening visit and a negative urine pregnancy test at the Randomization visit* | *PD_MM.141* | *Yes* | Entered but did not satisfy criteria | No | |
| *Woman of childbearing potential who did not agree to perform monthly pregnancy tests up to the end of the S-FU period* | *PD_MM.142* | *Yes* | Entered but did not satisfy criteria | No | |
| *Woman of childbearing potential who did not agree to use reliable contraception from at least 30 days prior to Visit 2 and up to at least 30 days after study drug discontinuation* | *PD_MM.143* | *Yes* | Entered but did not satisfy criteria | No | |
| *Known severely reduced single ventricle ejection fraction (< 30%)* | *PD_MM.144* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *Known Fontan circulation stenosis, affected area > 50% of the diameter* | *PD_MM.145* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *Known valvular defects (severe atrioventricular [AV] valve regurgitation, outflow obstruction)* | *PD_MM.146* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *Known pulmonary-venous pathway obstruction* | *PD_MM.147* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *Candidate on the active list for heart transplantation within 3 months of screening* | *PD_MM.148* | *Yes* | Entered but did not satisfy criteria | No | |
| *Clinical worsening leading to medical interventions including reoperation of Fontan circulation within 3 months of screening* | *PD_MM.149* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *Unscheduled hospitalization due to deterioration of the Fontan palliated condition within 3 months of screening* | *PD_MM.150* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *Signs and symptoms of HF requiring change in diuretic therapy within 3 months of screening* | *PD_MM.151* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *Ventricular tachyarrhythmia or supraventricular tachyarrhythmia within 3 months of screening* | *PD_MM.152* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *Peritoneal, pleural, mediastinal, or pericardial effusions within 3 months of screening* | *PD_MM.153* | *Yes* | Entered but did not satisfy criteria | No | |
| *Thrombotic or hemorrhagic complications (including thromboembolism and hemoptysis) within 3 months of screening* | *PD_MM.154* | *Yes* | Entered but did not satisfy criteria | No | |
| *Protein losing enteropathy (serum albumin < 30 g/L and total protein < 50 g/L) within 3 months of screening* | *PD_MM.155* | *Yes* | Entered but did not satisfy criteria | No | |
| *Plastic bronchitis/chyloptysis within 3 months of screening* | *PD_MM.156* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *Signs and symptoms requiring the addition of a new class of cardiovascular medication (e.g., nitrates, alpha-blockers, or ERAs) within 3 months of screening* | *PD_MM.157* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *History of syncope during exercise* | *PD_MM.158* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *Symptomatic coronary artery disease at Screening* | *PD_MM.159* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *Iron supplementation, regardless of route of administration, unless already present at stable dose for more than 3 months prior to Screening* | *PD_MM.160* | *Yes* | Entered but did not satisfy criteria | No | |
| *Exercise training program for cardiopulmonary rehabilitation in the 3-month period prior to Screening* | *PD_MM.161* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *Myocardial infarction < 6 months before Screening* | *PD_MM.162* | *Yes* | Entered but did not satisfy criteria | No | |
| *Pacemaker at Screening* | *PD_MM.163* | *Yes* | Entered but did not satisfy criteria | No | |
| *Known or suspected pulmonary veno-occlusive disease* | *PD_MM.164* | *Yes* | Entered but did not satisfy criteria | No | |
| *Known and documented severe hepatic impairment defined as Child-Pugh Score C* | *PD_MM.165* | *Yes* | Entered but did not satisfy criteria | No | |
| *Hypersensitivity to any active substance or excipient of any of the study drugs* | *PD_MM.166* | *Yes* | Entered but did not satisfy criteria | No | |
| *Treatment with a strong cytochrome P450 3A4 (CYP3A4) inducer within 1 month prior to Randomization (Visit 2)* | *PD_MM.167* | *Yes* | Entered but did not satisfy criteria | No | |
| *Treatment with a strong CYP3A4 inhibitor within 1 month prior to Randomization (Visit 2)* | *PD_MM.168* | *Yes* | Entered but did not satisfy criteria | No | |
| *Any factor or condition likely to affect protocol compliance of the subject, as judged by the investigator* | *PD_MM.169* | *Yes* | Entered but did not satisfy criteria | No | |
| *Treatment with another investigational therapy in the 3-month period prior to Screening* | *PD_MM.170* | *Yes* | Entered but did not satisfy criteria | No | |
| *Treatment with nitrates, alpha-blockers, or ERAs in the 3-month period prior to Screening* | *PD_MM.171* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *Introduction or change of dose of PH specific drugs (other than ERAs) in the 3-month period prior to Screening* | *PD_MM.172* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *Known drug or substance (e.g., alcohol) abuse, unstable psychiatric illness, or any other condition that, in the opinion of the investigator, may interfere with participation in the study* | *PD_MM.173* | *Yes* | Entered but did not satisfy criteria | No | |

| | | | | | |
|---|---|---|---|---|---|
| *Any planned surgical intervention (e.g., organ transplant) during the study period, except minor interventions (e.g., tooth extraction)* | *PD_MM.174* | *Yes* | Entered but did not satisfy criteria | No | |
| *Any known factor or disease that may interfere with treatment compliance or full participation in the study (e.g., chemotherapy treatment for cancer) or illness with an anticipated life expectancy of less than 12 months.* | *PD_MM.175* | *Yes* | Entered but did not satisfy criteria | No | |
| *PA-Ac device not correctly assigned or initiated during Screening* | *PD_MM.176* | *Yes* | Other | No | Confirmed by DM that missing data are fully fully covered by PD_PM.303 |
| *Subject not contacted to remind them to wear PA-Ac device* | *PD_MM.177* | *No* | Other | No | |
| *PA-AC device not returned/device lost/issues with upload* | *PD_MM.178* | *Yes* | Other | No | Confirmed by DM that missing data are fully fully covered by PD_PM.303 / PD_PM.313 |
| *PA-Ac device not correctly assigned or initiated at Visit 4* | *PD_MM.179* | *Yes* | Other | No | |
| *PA-Ac device not returned/lost device/upload problems* | *PD_MM.180* | *Yes* | Other | No | Confirmed by DM that missing data are fully fully covered by PD_PM.303 / PD_PM.313 |
| *Subject not contacted to remind them to wear PA-Ac device* | *PD_MM.181* | *No* | Other | No | |
| *PA-Ac device not correctly assigned or initiated at Visit 5* | *PD_MM.182* | *Yes* | Other | No | |
| *PA-Ac device not correctly assigned or initiated for Visit 6* | *PD_MM.183* | *Yes* | Other | No | |
| *Pregnancy test was missed* | *PD_MM.184* | *Yes* | Other | No | |
| *Subject received treatment which was not approved for use* | *PD_MM.185* | *Yes* | Other | No | |
| *Study treatment not administered as per protocol (any dose of study treatment higher than the planned total daily dose in a single day will be considered an overdose)* | *PD_MM.186* | *Yes* | Received wrong treatment or incorrect dose | No | |
| *Subject treatment not discontinued although subject (legally acceptable representative) decided not to continue treatment* | *PD_MM.187* | *Yes* | Developed withdrawal/interruption criteria but not withdrawn/interrupted | No | |
| *Subject medication not discontinued although investigator considers that the subject should not continue treatment* | *PD_MM.188* | *Yes* | Developed withdrawal/interruption criteria but not withdrawn/interrupted | No | |
| *Sponsor decision to withdraw subject or terminate study and subject was not discontinued within the time indicated by the sponsor* | *PD_MM.189* | *Yes* | Developed withdrawal/interruption criteria but not withdrawn/interrupted | No | |
| *Visit was performed remotely/partly remotely* | *PD_MM.190* | *No* | Other | No | |
| *Monthly lab sample missing* | *PD_MM.191* | *Yes* | Other | No | |
| *Consecutive lab samples missing/not available without replacement/local labs being done* | *PD_MM.192* | *Yes* | Other | No | |
| *Oversampling during lab assessments* | *PD_MM.193* | *Yes* | Other | No | |
| *CPET at Visit 5 done, but not valid/questionable on review by central reading facility* | *PD_MM.314* | *Yes* | Other | No | |
| *CPET at Visit 6 done, but not valid/questionable on review by central reading facility* | *PD_MM.318* | *Yes* | Other | Yes | PP2 |
| *Screening assessment not done/missing* | *PD_PM.301* | *Yes* | Other | No | |
| *Subject does not have sufficient PA-Ac baseline data* | *PD_PM.303* | *Yes* | Other | Yes | PP3 |
| *Randomization CPET done, but not valid/questionable on review by central reading facility* | *PD_PM.304* | *Yes* | Other | Yes | PP1 and PP2 |
| *Randomization assessment not done/missing* | *PD_PM.305* | *Yes* | Other | No | |
| *Randomization labs missing/incomplete* | *PD_PM.306* | *Yes* | Other | No | |
| *Subject has no post-baseline safety assessment* | *PD_PM.307* | *Yes* | Other | No | |
| *Subject has no post-baseline efficacy assessment* | *PD_PM.308* | *Yes* | Other | No | Confirmed by DM that this is covered by PD_PP.251, PD_PP.260, PD_PP.253, PD_PM.313 |
| *Visit 3 assessment not done/missing* | *PD_PM.309* | *Yes* | Other | No | |

| | | | | | |
|---|---|---|---|---|---|
| *Visit 3 labs missing/incomplete* | *PD_PM.310* | *Yes* | Other | No | |
| *Visit 4 assessment not done/missing* | *PD_PM.311* | *Yes* | Other | No | |
| *Visit 4 labs missing/incomplete* | *PD_PM.312* | *Yes* | Other | No | |
| *Subject does not have sufficient PA-Ac data at Visit 4* | *PD_PM.313* | *Yes* | Other | Yes | PP3 |
| *Visit 5 assessment not done/missing* | *PD_PM.315* | *Yes* | Other | No | |
| *Visit 5 labs missing/incomplete* | *PD_PM.316* | *Yes* | Other | No | |
| *Subject does not have sufficient PA-Ac data for Visit 5* | *PD_PM.317* | *Yes* | Other | No | |
| *Visit 6 assessment not done/missing* | *PD_PM.319* | *Yes* | Other | No | |
| *Visit 6 labs missing/incomplete* | *PD_PM.320* | *Yes* | Other | No | |
| *Subject does not have sufficient PA-Ac data for Visit 6* | *PD_PM.321* | *Yes* | Other | No | |
| *Visit 7 assessment not done/missing* | *PD_PM.322* | *Yes* | Other | No | |
| *Visit 7 labs missing/incomplete* | *PD_PM.323* | *Yes* | Other | No | |
| *Initiation of forbidden therapy up to the week 16 visit* | *PD_PM.324* | *Yes* | Received a disallowed concomitant treatment | Yes | all PP |
| *Aminotransferase elevation ≥ 3 x ULN and study drug not interrupted or discontinued* | *PD_PM.325* | *Yes* | Developed withdrawal/interruption criteria but not withdrawn/interrupted | No | |
| *Aminotransferase elevation ≥ 3 x ULN and retests not done per protocol* | *PD_PM.326* | *Yes* | Other | No | |
| *Hemoglobin < 100 g/L with a decrease from baseline of ≥ 50 g/L and retest not done per protocol* | *PD_PM.327* | *Yes* | Other | No | |
| *Hemoglobin < 80 g/L with no underlying blood loss and study treatment not interrupted* | *PD_PM.328* | *Yes* | Developed withdrawal/interruption criteria but not withdrawn/interrupted | No | |
| *Re-introduction of treatment before*<br>*•Aminotransferases returned to pre-treatment or normal range*<br>*•Hemoglobin returned to pre-treatment or normal range* | *PD_PM.329* | *Yes* | Developed withdrawal/interruption criteria but not withdrawn/interrupted | No | |
| *Blood pressure dropped < 90 mm Hg (< 85 mm Hg for subjects < 18 years old and < 150 cm of height and study drug not interrupted* | *PD_PM.330* | *Yes* | Developed withdrawal/interruption criteria but not withdrawn/interrupted | No | |
| *Subject got pregnant during the treatment period or within 30 days after treatment discontinuation* | *PD_PM.331* | *Yes* | Other | No | |
| *Woman of childbearing potential not using reliable contraception* | *PD_PM.332* | *Yes* | Other | No | |
| *Administration of incorrect study treatment (i.e. Subject did not get medication bottle assigned by IRT)* | *PD_PM.333* | *Yes* | Received wrong treatment or incorrect dose | Yes (if wrong kit contained non-randomized drug) | all PP (if wrong drug). If PD date available (either date of first intake from an incorrect bottle of - if this data is not available, use date the wrong bottle was dispensed ), only exclude if occurred prior to EP assessment date |
| *Study treatment compliance between 2 visits < 80% or > 120%* | *PD_PM.334* | *No* | Received wrong treatment or incorrect dose | No | |
| *Study treatment compliance for the whole treatment period < 80% or > 120%* | *PD_PM.335* | *Yes* | Received wrong treatment or incorrect dose | Yes | PP2 |
| *Study treatment compliance up to the Week 16 visit < 80% or > 120%* | *PD_PM.336* | *Yes* | Received wrong treatment or incorrect dose | Yes | PP1 and PP3 |
| *Subject did not enter the PTOP after premature EOT* | *PD_PM.337* | *No* | Other | No | |
| *Subject was not called every 90+/-7 days during PTOP* | *PD_PM.338* | *No* | Other | No | |
| *Subject was not called at EOP* | *PD_PM.339* | *No* | Other | No | |

| | | | | | |
|---|---|---|---|---|---|
| *Subject was not called at VSFU* | *PD_PM.340* | *No* | Other | No | |
| *Initiation of forbidden therapy after the week 16 visit* | *PD_PM.341* | *Yes* | Received a disallowed concomitant treatment | Yes | PP2 |
| *Subject got pregnant prior to Week 16 (visit 4)* | *PD_PM.342* | *Yes* | Other | No | |
| *Screening lab not done/missing* | *PD_PM.350* | *Yes* | Other | No | |
| *Male or female < 12 years old* | *PD_PM.351* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *Not a Fontan-palliated subject with either LT-TCPC, or EC-TCPC surgery > 1 year before Screening* | *PD_PM.352* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *Subject not in NYHA FC II or III (assessed by the investigator using the Specific Activity Scale* | *PD_PM.353* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *Peripheral oxygen saturation ($SpO_2$) of < 88% at rest at Screening* | *PD_PM.354* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *Respiratory limitation with a breathing reserve of < 10%* | *PD_PM.355* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *Iron deficiency defined as ferritin < 10 µg/L at Screening* | *PD_PM.356* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *Body mass index (BMI) for age reference (z scored) equivalent to BMI > 35 kg/m$^2$ (for adults) at Screening* | *PD_PM.357* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *Systolic BP < 90 mmHg (< 85 mmHg for subjects < 18 years old and < 150 cm of height) at rest or during CPET at Screening* | *PD_PM.358* | *Yes* | Entered but did not satisfy criteria | No | |
| *Hemoglobin < 75% of the lower limit of normal assessed by central laboratory at Screening* | *PD_PM.359* | *Yes* | Entered but did not satisfy criteria | No | |
| *Serum AST and/or ALT > 3 x upper limit of normal range assessed by central laboratory at Screening* | *PD_PM.360* | *Yes* | Entered but did not satisfy criteria | No | |
| *Severe renal impairment (estimated creatinine clearance < 30 mL/min/1.73m$^2$) assessed by central laboratory at Screening* | *PD_PM.361* | *Yes* | Entered but did not satisfy criteria | No | |
| *Peak $VO_2$ < 15 mL/kg/min at Baseline* | *PD_PM.362* | *Yes* | Entered but did not satisfy criteria | Yes | all PP |
| *Systolic BP < 90 mmHg (< 85 mmHg for subjects < 18 years old and < 150 cm of height) at rest or during CPET at Baseline* | *PD_PM.363* | *Yes* | Entered but did not satisfy criteria | No | |
| *PK sampling taken after the morning dose of study treatment* | *PD_PM.364* | *Yes* | Other | No | |
| *Subject had treatment interruption(s) in the 2 weeks prior to PK sample being collected* | *PD_PM.365* | *Yes* | Received wrong treatment or incorrect dose | No | |
| *Subject got pregnant and study drug was not interrupted* | *PD_PM.366* | *Yes* | Developed withdrawal/interruption criteria but not withdrawn/interrupted | No | |
| *Treatment interruption of > 4 consecutive weeks up to Week 16* | *PD_PM.367* | *Yes* | Received wrong treatment or incorrect dose | Yes | all PP |
| *Treatment interruption of > 4 consecutive weeks after Week 16* | *PD_PM.368* | *No* | Received wrong treatment or incorrect dose | Yes | PP2 |
| *Subject not Randomized but received study drug* | *PD_PP.245* | *Yes* | Other | Yes | all PP |
| *No CPET training test conducted* | *PD_PP.246* | *No* | Other | No | |
| *Randomization occurred outside of Screening period* | *PD_PP.249* | *No* | Other | No | |
| *Subject has no PK assessment* | *PD_PP.250* | *Yes* | Other | No | |
| *CPET at Visit 4 not done* | *PD_PP.251* | *Yes* | Other | Yes | PP1 |
| *CPET at Visit 5 not done* | *PD_PP.252* | *Yes* | Other | No | |
| *CPET at Visit 6 not done* | *PD_PP.253* | *Yes* | Other | Yes | PP2 |
| *Treatment interruption of > 2 consecutive weeks and treatment not permanently discontinued* | *PD_PP.254* | *No* | Developed withdrawal/interruption criteria but not withdrawn/interrupted | No | Covered in PD_PM.367/PD_PM.368 |
| *Visit 3 performed out of window (Day 57 ± 7 days)* | *PD_PP.255* | *No* | Other | No | |
| *Visit 4 performed out of window (Day 113± 7 days)* | *PD_PP.256* | *Yes* | Other | No | |
| *Visit 5 performed out of window (Day 225 ± 7 days)* | *PD_PP.257* | *No* | Other | No | |
| *Visit 6 performed out of window (Day 365 ± 7 days)* | *PD_PP.258* | *No* | Other | No | |
| *Visit 7 performed out of window (EOT + 30-35 days)* | *PD_PP.259* | *No* | Other | No | |

| | | | | | |
|---|---|---|---|---|---|
| *Peak $VO_2$ measurements at Week 16 (Visit 4) not available* | *PD_PP.260* | *Yes* | Other | Yes | PP1 |
| *Treatment interruption of > 4 consecutive weeks and treatment not permanently discontinued* | *PD_PP.261* | *No* | Developed withdrawal/interruption criteria but not withdrawn/interrupted | No | Covered in PD_PM.367/PD_PM.368 |
| *Monthly lab performed out of window* | *PD_PP.262* | *No* | Other | No | |

**Actelion Pharmaceuticals Ltd.**

# Statistical Analysis Plan for COVID-19 Impact Assessment

## Prospective, multi-center, double-blind, randomized, placebo-controlled, parallel-group study assessing the efficacy and safety of macitentan in Fontan-palliated adult and adolescent subjects

**Protocol AC-055H301; Phase 3**

**JNJ-67896062 / ACT-064992 (macitentan)**

| | |
|---|---|
| **Status:** | Approved |
| **Date:** | 23 August 2021 |
| **Prepared by:** | Actelion Pharmaceuticals Ltd., a Janssen Pharmaceutical Company of Johnson & Johnson |
| **Document No.:** | EDMS-307892, 2.0 |

**Compliance:** The study described in this report was performed according to the principles of Good Clinical Practice (GCP).

# TABLE OF CONTENTS



# LIST OF IN-TEXT TABLES

## VERSION HISTORY

**Table 1: SAP Version History Summary**

| SAP Version | Approval Date | Change | Rationale |
|---|---|---|---|
| 1.0 | 8 April 2021 | Not Applicable | Initial release |
| 2.0 | 23 August 2021 | Sections 5.1.2 | Clarification of visit mapping in case of missing or invalid assessments. |
| | | Section 5.3.2.1 | 1. Addition of "Pre/During COVID-19" as Week 16 assessment may fall in Pre-COVID19 period while Week 52 in "During COVID-19" period.<br>2. After consultation with the Study Results Committee constituted of external clinicians with expertise in subjects with Functional Single Ventricle Heart Disease on 13-Aug-2021, an additional analysis will be performed excluding subjects who had reported COVID-19 infection. |
| | | Section 5.3.2.2 | Clarification for the derivation of the 9-day time window |

# 1. INTRODUCTION

This Statistical Analysis Plan (SAP) describes the analyses to assess the impact of the COVID-19 pandemic in the double-blind study AC-055H301 (RUBATO), following the guidance within the COVID-19 Standard Reporting Guideline Principles. This document is a separate Appendix of the Clinical Study Report (CSR) SAP of this study.

This SAP refers to the documents listed in Table 2.

**Table 2: Study Documents**

| Document | Date, Version |
|---|---|
| Study Protocol AC-055H301 (RUBATO) | Final version 10 |
| Study Protocol AC-055H301 COVID-19 Appendix | Approved version, 17 June 2020 (D-20.158) |
| SAP for CSR AC-055H301 (RUBATO) | Final approved version (EDMS-RIM-264216) |
| Data Presentation Specifications Part 1 AC-055H301 (RUBATO) | Final version |

Source data for the analyses will be Analysis Data Model (ADaM) data sets derived for the AC-055H301 (RUBATO) CSR SAP.

Except where otherwise specified, all definitions and conventions used in the CSR SAP of this study will be applied.

## 1.1. Objectives

To assess the impact of COVID-19 pandemic on:

1. disposition information
2. compliance
3. extent of exposure
4. protocol deviations
5. prior and concomitant medications
6. efficacy (ie, primary and secondary efficacy endpoints)
7. selected safety.

## 1.2. Study Design

The study design is described in section 3.1 of the study protocol.

# 2. STATISTICAL HYPOTHESES

Not applicable for the scope of this document.

The statistical hypotheses for the trial objectives are described in the CSR SAP (section 1.3 for general consideration and sections 5.2.3.1, 5.3.1.3, and 5.3.2.3 for primary and secondary efficacy endpoints, respectively).

# 3. SAMPLE SIZE DETERMINATION

Not applicable for the scope of this document.

The sample size of this study AC-055H301/RUBATO is based on the justification provided in section 1.4 of the CSR SAP.

# 4. POPULATIONS (ANALYSIS SETS) FOR ANALYSIS

The same analysis sets provided in section 2.2 of the CSR SAP will be used.

# 5. STATISTICAL ANALYSES

## 5.1. General Definitions

### 5.1.1. COVID-19 periods of the pandemic

World Health Organization Director-General's opening remarks at the media briefing on COVID-19 on 11 March 2020 declared that the assessment of COVID-19 can be characterized as a pandemic. Therefore, the COVID-19 periods related to the pandemic could be identified as follows:

**Table 3: COVID-19 periods of the pandemic**

| Period | Definition |
|---|---|
| Pre COVID-19 | Defined as the period before 11 March 2020 unless there is prior evidence of COVID-19 by the data collection. In that case, 11 March 2020 is replaced by the date of the first evidence of COVID-19 (see definition below). |
| During COVID-19* | Defined as the period between 11 March 2020 (or earlier in case of prior evidence of COVID-19) and the end date of the pandemic (limits included). |
| Post COVID-19* | Defined as the period starting immediately after the end date of pandemic onwards. |

* Given that the pandemic still ongoing at the time of this document, no end date of the pandemic is available. Therefore, an open end is set for the 'During COVID-19' period and the 'Post COVID-19' period cannot be considered for the analysis.

Prior evidence of COVID-19 (ie, before 11 March 2020) can be identified as follows:

- **Date of premature study treatment discontinuation / study withdrawal**: reasons for premature study treatment discontinuation / study withdrawal are identified in the corresponding eCRF modules prefixed with the text "COVID-19 related", as appropriate (ie, either for AE, or medical or non-medical reason, whether subject or physician decision).
- **Onset date of adverse events (AEs)**: COVID-19 related AEs are identified if the corresponding preferred terms (PTs) are included in the list as provided in Appendix 2 List of COVID-19 related AEs.
- **Date of death**: if the death is COVID-19 related, the primary cause of death should contain the text "COVID-19" in the "Death" eCRF module. (Note: "COVID-19" should then also be reported in the corresponding (S)AE).
- **Protocol deviations (PDs)**: if a PD is linked to COVID-19, the PD term should be prefixed with the text "COVID-19 related".

Each assessment/visit will be associated to the COVID-19 periods of the pandemic as defined above considering their date respect to the cut-off date of 11 March 2020 (or earlier, as applicable). For example, if the peak $VO_2$ assessment at Week 16 was performed before the cut-off date of 11 March 2020, and there is no prior evidence of COVID-19, that assessment is associated to the "Pre COVID-19" period, otherwise it is associated to the "During COVID-19" period.

Note that it is possible that 'COVID-19 related" data occurred in the "Pre COVID-19" period, due to the fact that the study is running since late 2017 and the pandemic start date selected for the analyses (11 March 2020) does not necessarily coincide with the COVID-19 virus circulation start date in all countries.

### 5.1.2. Visit mapping

To minimize the amount of missing data and the possibility to delay the assessments until the on-site visits can be resumed according to the COVID-19 protocol appendix, the efficacy analysis will be based on scheduled visits, considering all scheduled assessments after the treatment start date. In case of missing or invalid assessment at scheduled visit, extended time windows for Week 16 and Week 52 will be considered as follows:

- for subjects who prematurely discontinue study treatment, the End-of-Treatment visit (if available) will be re-assigned to a scheduled visit if conducted during the time window specified in the table below.
- unscheduled visits (if available) or Visit 5/Week 32 (if available) will be also re-assigned to a scheduled visit if conducted during the time window specified in the table below.

The following rules are applied to assign End-of-Treatment or unscheduled visits (including Visit 5/Week 32) to analysis visits. If a subject has 2 or more actual visits with an available value in the same visit window the visit closest to the target day will be used as the protocol visit for that visit window. For peak $VO_2$, if a subject has 2 or more actual visits in the same visit window, the visit with the valid value closest to the target day will be used as the protocol visit for that visit window. The other additional visit(s) will not be used in the summaries or analyses by visit, but they can be used for determination of other endpoints (eg, time to event endpoints, marked laboratory abnormalities). If 2 actual visits are equidistant from the target day within a visit window, the later visit is used. Except otherwise specified, if more than one assessment falls on the same day then the worst assessment (ie, lowest for peak $VO_2$) is used.

**Table 4: Mapping of premature End-of-Treatment visit or unscheduled visits**

| Parameter | Scheduled Visit Number | Time Interval (label on output) | Time Interval (Day)* | Target Time Point (Day) |
|---|---|---|---|---|
| Premature End-of-Treatment or unscheduled visits re-mapping | 2 | Baseline | <=1 | 1 |
| | 4 | Week 16 | 2 to 224 | 113 |
| | 6 | Week 52 | 225 to open end | 365 |

*Relative to Study Day 1.

## 5.2. Subject information

The last subject was randomized in RUBATO in 30 June 2020 with the screening activities announced to close on 23 June 2020. In total there are 8 subjects randomized after or on 11 March 2020.

### 5.2.1. Subject Disposition

No subject was reported to fail the screening because of COVID-19, therefore the impact of COVID-19 in the screening activities will not be evaluated in the different pandemic periods.

#### 5.2.1.1. Study and Study Treatment Disposition

Reasons for premature termination of the study due to COVID-19 are captured in the "Study discontinuation" eCRF module.

Reasons for premature study treatment discontinuation due to COVID-19 are captured in the "Premature Discontinuation of Study Treatment" eCRF module.

The corresponding summaries for study and study treatment disposition as planned in the CSR SAP section 4.3.2 will show details for subjects who reported reasons for discontinuation due to COVID-19. Therefore, no specific summary will be provided.

Listings of subjects will be provided for the following categories:

- Subjects who terminated study prematurely due to COVID-19
- Subjects who discontinued study treatment prematurely due to COVID-19.

### 5.2.2. Medical History

The impact of COVID-19 in the medical history will not be evaluated as there are only 8 subjects randomized after or on 11 March 2020.

### 5.2.3. Compliance

#### 5.2.3.1. Treatment Compliance

The analysis of treatment compliance defined in the main CSR SAP (overall up to EOT, and up to Week 16) is further performed in the subgroup of subjects impacted by COVID-19 (ie, for those subjects who experienced at least a COVID-19 related AE). A subject listing of study treatment compliance will be provided.

#### 5.2.3.2. Study Assessment Compliance

The number and percentage of subjects in the following categories will be summarized by treatment group and overall on the FAS, to display the following information:

- Number and % of subjects who missed at least one post-baseline scheduled assessment identified by the reporting of at least one of the following protocol deviations PD_PM.309 (Visit 3 assessment not done/missing), PD_PM.311 (Visit 4 assessment not done/missing),

PD_PM.315 (Visit 5 assessment not done/missing), PD_PM.319 (Visit 6 assessment not done/missing), PD_PM.322 (Visit 7 assessment not done/missing)

  - Number and % of subjects who missed at least one post-baseline scheduled assessment due to COVID-19, identified by the reporting of at least one of the above protocol deviations with the prefix "COVID-19 related"

- Number and % subjects who missed scheduled Visit 4/Week 16 assessment, identified by the reporting of the protocol deviation PD_PM.311 (Visit 4 assessment not done/missing)

  - Number and % of subjects who missed scheduled Visit 4/Week 16 assessment due to COVID-19, identified by the reporting of the protocol deviation PD_PM.311 with the prefix "COVID-19 related"

- Number and % of subjects who missed CPET at scheduled Visit 4/Week 16, identified by the reporting of the protocol deviation PD_PP.251 (CPET at Visit 4 not done)

  - Number and % of subjects who missed CPET at scheduled Visit 4/Week 16 due to COVID-19, identified by the reporting of the protocol deviation PD_PP.251 with the prefix "COVID-19 related"

- Number and % of subjects with insufficient PA-Ac data at scheduled Visit 4/Week 16, identified by the reporting of the protocol deviation PD_PM.313 (Subject does not have sufficient PA-Ac data at Visit 4)

  - Number and % of subjects with insufficient PA-Ac data at scheduled Visit 4/Week 16 due to COVID-19, identified by the reporting of the protocol deviation PD_PM.313 with the prefix "COVID-19 related"

- Number and % subjects who missed scheduled Visit 6/Week 52, identified by the reporting of the protocol deviation PD_PM.319 (Visit 6 assessment not done/missing)

  - Number and % of subjects who missed scheduled Visit 6/Week 52 due to COVID-19, identified by the reporting of the protocol deviation PD_PM.319 with the prefix "COVID-19 related"

- Number and % of subjects who missed CPET at scheduled Visit 6/Week 52, identified by the reporting of the protocol deviation PD_PP.253 (CPET at Visit 6 not done)

  - Number and % of subjects who missed CPET at scheduled Visit 6/Week 52 due to COVID-19, identified by the reporting of the protocol deviation PD_PP.253 with the prefix "COVID-19 related"

- Number of remote visits per subject, identified by the reporting of the protocol deviation PD_MM.190 (Visit was performed remotely/partly remotely)

- Number of remote visits per subject due to COVID-19, identified by the reporting of the protocol deviation PD_MM.190 with the prefix "COVID-19 related"

A subject listing of study assessment compliance will be provided based on the PDs reporting.

### 5.2.4. Extent of Exposure

A subject is considered to have a dose not administered if he/she has a study treatment interruption (ie, if a reason associated with a study treatment end is 'Temporarily interrupted due to an AE' or 'Temporarily interrupted not due to an AE' in the "Study Drug Log" eCRF module).

Due to data collection (eg, no direct link between study treatment interruption and AEs), a dose not administered because of COVID-19 cannot be derived from the "Study Drug Log", therefore no summary will be provided.

A listing of subjects with any temporary study treatment interruptions (ie, as reported in the "Study Drug Log" eCRF module) will be reported as well as a separate listing of subjects who reported a COVID-19 related AEs with action taken with study treatment="Drug interrupted".

### 5.2.5. Protocol Deviations

Protocol deviations COVID-19 related are identified with the text "COVID-19 related" as a prefix of the PD description.

Subjects with important (ie, major) "COVID-19 related" PDs will be summarized by displaying counts and percentages of subjects with at least an important "COVID-19 related" PD within each category (as defined in the CSR SAP) by treatment group and overall. A separate similar summary of important "COVID-19 related" PDs will be provided by geographical region and site.

All reported "COVID-19 related" PDs will be described in a subject listing. Important PDs will be flagged accordingly.

### 5.2.6. Prior and Concomitant Medications

Previous and study treatment concomitant therapy are defined in the CSR SAP (sections 4.7.1 and 4.7.2, respectively).

Individual subject's listing of all concomitant medications for all subjects who experienced at least a COVID-19 related AE will be provided. Previous and study treatment concomitant therapies are flagged accordingly.

## 5.3. Efficacy Analyses

All efficacy analyses will be based on the full analysis set (FAS), not controlling for the adaptive design and not based on the population wise splitting approach as applied for the main analyses of primary and secondary efficacy endpoints.

For both the primary and secondary endpoints, the corresponding estimand section describes data handling rules. Baseline definitions follows section 2.5.1 of the CSR SAP. To minimize the

amount of missing data and the possibility to delay the assessments until the on-site visits can be resumed according to the COVID-19 protocol appendix, the assessments at Week 16 and Week 52 will be based on scheduled Visit 4 and Visit 6, respectively. In case of missing scheduled visit/assessment, premature End-of Treatment and unscheduled visits (including Visit 5/Week 32 if available) will be re-mapped to a scheduled visit as described in Section 5.1.2 (unless otherwise specified).

An overview of visits/assessments not performed because of COVID-19 will be provided by treatment group and overall for primary and secondary efficacy endpoints.

### 5.3.1. Primary Efficacy Endpoint

Absolute values at baseline and at Week 16 as well as absolute change in peak $VO_2$ from baseline to Week 16 will be summarized using descriptive statistics, where appropriate.

1. Additional Estimand (based on scheduled visits).

| | |
|---|---|
| **A.** | **Population:** FAS |
| **B.** | **Variable:** peak $VO_2$ as change from baseline to scheduled Week 16 visit (see Section 5.1.2 of this document). |
| **C.** | **Intercurrent events and strategies**: same as primary estimand (see section 5.2.2 of the CSR SAP). |
| **D.** | **Population-level summary measure:** same as the ANCOVA model of the primary estimand (see section 5.2.2 of the CSR SAP) carried out on the total sample size (without the population wise splitting approach for the control of the adaptive design). |

2. To account for the pandemic and to assess the robustness of the results in the presence of the changes made during the pandemic, a Mixed Model Repeated Measures (MMRM) will use all data available to estimate the effect at Week 16.

**A. Population:** All randomized Fontan-palliated adult and adolescent subjects defined through protocol inclusion/exclusion criteria (ie, FAS).

**B. Variable:** peak $VO_2$ as change from baseline to Week 16.

**C. Intercurrent events:**

- Death occurring before Week 52: composite strategy (worst-case).

In case of death prior to Week 52, all scheduled assessments following the date of death will be imputed by a worst-case imputation (ie, lowest observed change in peak $VO_2$ from baseline value across both treatment groups and over all assessments).

- Non-missing peak $VO_2$ value at Week 16 or Week 52 assessed as 'invalid' by the central reading facility: composite strategy (minimum of $25^{th}$ percentile and no change)

In case an assessment at Week 16 or Week 52 has been evaluated as "invalid" by the independent central reading facility, it will be ignored. This will result in imputing the change from baseline to Week 16 or Week 52 with the $25^{th}$ percentile of the observed changes in peak $VO_2$ from baseline value (across both treatment groups and over all assessments). If this leads to an improvement (change >0), a change of 0 (no change) will be imputed.

- Premature study treatment discontinuation before Week 52 not "COVID-19" related.

According to the protocol, if treatment is discontinued prior to Week 16, and the subject enters the Post-Treatment Observation Period (PTOP), it is requested that the subject returns for a CPET assessment at Week 16 but not at Week 52.

If the premature study treatment discontinuation occurs prior to Week 16 and it is not "COVID-19" related:

- If a valid assessment is available at Week 16 (see Section 5.1.2), this will result in imputing only the change from baseline to Week 52 with the 25$^{th}$ percentile of the observed changes in peak $VO_2$ from baseline value (across both treatment groups and over all assessments). If this leads to an improvement (change>0), a change of 0 (no change) will be imputed.
- If no valid assessment is available at Week 16 (see Section 5.1.2), all scheduled assessments following the date of event will be imputed by applying the 25$^{th}$ percentile of the observed changes in peak $VO_2$ from baseline value (across both treatment groups and over all assessments). If this leads to an improvement, a change of 0 (no change) will be imputed.

If the premature study treatment discontinuation occurs after Week 16 and before Week 52 and it is not "COVID-19" related (and no valid assessment is available at Week 52 based on the definition in Section 5.1.2), this will result in imputing the change from baseline to Week 52 with the 25$^{th}$ percentile of the observed changes in peak $VO_2$ from baseline value (across both treatment groups and over all assessments). If this leads to an improvement (change>0), a change of 0 (no change) will be imputed.

**D. Population-level summary measure:** difference in mean change in peak $VO_2$ from baseline to Week 16 between macitentan 10 mg and placebo, as estimated from the MMRM.

Missing data at Week 16 and Week 52 arising from any other reason (eg, premature study treatment discontinuation due to COVID-19, missed visits because of COVID-19) are assumed to be missing at random, and as such, will require no further imputation, since the MMRM approach will implicitly impute these values under a missing at random assumption.

The model will include randomized treatment group, time (via a categorical variable for visit), treatment-by-time interaction, baseline-by-time interaction, and peak $VO_2$ at baseline as fixed effects. An unstructured variance-covariance matrix will be specified.

At Week 16, the resulting LS means and corresponding 95% and 99% CLs obtained in each treatment group, and the LS means differences (95% and 99% CLs) for macitentan 10 mg vs placebo will be provided.

3. Subgroup analysis based on the COVID-19 periods of the pandemic.

In order to assess the consistency of the treatment effect across the COVID-19 periods of the pandemic, the subgroup "COVID-19 periods" will be added, categorizing the assessments as follows:

- "Pre COVID-19 period" if Week 16 peak $VO_2$ assessment falls in the "Pre COVID-19" period
- "During COVID-19 period" if Week 16 peak $VO_2$ assessment falls in the "During COVID-19" period

The COVID-19 periods of the pandemic are defined in Table 3 of Section 5.1.1.

The subgroup analysis will be carried out the same way as the other subgroup analyses as described in section 5.2.3.5 of the CSR SAP, based on scheduled visits.

### 5.3.2. Secondary Efficacy Endpoint(s)

#### 5.3.2.1. Change from baseline over 52 weeks in peak $VO_2$

Absolute values at baseline and over Week 52 as well as absolute change in peak $VO_2$ from baseline over Week 52 will be summarized using descriptive statistics, where appropriate.

1. Additional Estimand (based on scheduled visits).

| | |
|---|---|
| **A.** | **Population:** FAS. |
| **B.** | **Variable:** peak $VO_2$ as average change from baseline to scheduled Week 16 and scheduled Week 52 (see Section 5.1.2 of this document) |
| **C.** | **Intercurrent events and strategies**: same as main estimand (see section 5.3.1.2 of the CSR SAP) |
| **D.** | **Population-level summary measure:** Same as the MMRM model of the main estimand (see section 5.3.1.2 of the CSR SAP), carried out on the total sample size (without the population wise splitting approach for the control of the adaptive design). |

2. To account for the pandemic and to assess the robustness of the results in the presence of the changes made during the pandemic, a Mixed Model Repeated Measures (MMRM) will use all data available to estimate the effect at Week 52 as done for the primary efficacy endpoint.

   - Same population, variable peak $VO_2$ as change from baseline to Week 52 (scheduled visit), same intercurrent events, and same handling of missing data as defined for the primary efficacy endpoint (Section 5.3.1 of this document).
   - Same MMRM model as defined for primary efficacy endpoint (Section 5.3.1) carried out on the total sample size (without the population wise splitting approach for the control of the adaptive design).

   At Week 52, the resulting LS means and corresponding 95% and 99% CLs obtained in each treatment group, and the LS means differences (95% and 99% CLs) for macitentan 10 mg vs placebo will be provided.

3. Subgroup analysis based on the COVID-19 periods of the pandemic.

Subgroup analyses will be carried out similarly as described for the primary efficacy endpoint in Section 5.3.1 above, by applying the MMRM. The assessments will be categorized as follows:

- "Pre COVID-19 period" if both Week 16 and Week 52 peak $VO_2$ assessment falls in the "Pre COVID-19" period
- "Pre/During COVID-19" if peak $VO_2$ assessment at Week 16 falls in the "Pre COVID-19" period and the one at Week 52 falls in the "During COVID-19" period
- "During COVID-19 period" if both Week 16 and Week 52 peak $VO_2$ assessment falls in the "During COVID-19" period

The COVID-19 periods of the pandemic are defined in Table 3 of Section 5.1.1.

After consultation with the Study Results Committee constituted of external clinicians with expertise in subjects with Functional Single Ventricle Heart Disease on 13-Aug-2021, an additional analysis will be performed excluding subjects who had reported a COVID-19 infection during the course of the study and consequently could have compromising assessments impacted by the prior infection. These subjects are identified based on the reporting of the AEs containing the text "COVID-19" (as qualified by the preferred term) and with the corresponding AE onset date before Week 16 or Week 52 assessments. This analysis will be performed only for this secondary efficacy endpoint over 52 weeks as COVID-19 infections were reported towards the end of the subject's treatment period, much later the evaluation of Week 16. The analysis will be performed based on the estimand at analysis point #1 (with a different population based on FAS excluding subjects described above).

#### 5.3.2.2. Change from baseline to Week 16 in mean count per minute of daily PA-Ac

Absolute values at baseline and at Week 16 as well as absolute change in mean count per minute of daily PA-Ac from baseline to Week 16 will also be summarized using descriptive statistics, where appropriate.

For Week 16, the 9-day time period is first selected as described in CSR SAP section 5.3.2. In case the 9-day time period is not available, an extended time window as described in Section 5.1.2 is considered.

1. Additional Estimand (based on extended time window for Week 16).

| | |
|---|---|
| **A.** | **Population:** FAS. |
| **B.** | **Variable:** Mean count per minute of daily PA-Ac as change from baseline to extended Week 16 time window (see Section 5.1.2 of this document) |
| **C.** | **Intercurrent events and strategies**: same as main estimand (see section 5.3.2.2 of the CSR SAP) |
| **D.** | **Population-level summary measure:** Same as the ANCOVA model of the main estimand (see section 5.3.2.2 of the CSR SAP) carried out on the total sample size (without the population wise splitting approach for the control of the adaptive design). |

2. To account for the pandemic and to assess the robustness of the results in the presence of the changes made during the pandemic, a Mixed Model Repeated Measures (MMRM) will use all data available to estimate the effect at Week 16.

**A. Population:** All randomized Fontan-palliated adult and adolescent subjects defined through protocol inclusion/exclusion criteria (ie, FAS).

**B. Variable:** mean count per minute of daily PA-Ac as change from baseline to Week 16.

**C. Intercurrent events:**

- Death occurring before Week 52: composite strategy (worst-case).

In case of death prior to Week 52, all scheduled assessments following the date of death will be imputed by a worst-case imputation (ie, lowest observed change from baseline value across both treatment groups and over all assessments).

- Premature study treatment discontinuation before Week 52 not "COVID-19" related.

If the premature study treatment discontinuation occurs prior to Week 16 and it is not "COVID-19" related:

  - If a valid assessment is available at Week 16 (see Section 5.1.2), this will result in imputing only the change from baseline to Week 52 with a change of 0 (no change).
  - If no valid assessment is available at Week 16 (see Section 5.1.2), all scheduled assessments following the date of event will be imputed by applying a change of 0 (no change).

If the premature study treatment discontinuation occurs after Week 16 and before Week 52 and it is not "COVID-19" related (and no valid assessment is available at Week 52 based on the definition in Section 5.1.2), this will result in imputing the change from baseline to Week 52 with a change of 0 (no change).

**D. Population-level summary measure:** difference in mean change mean count per minute from baseline to Week 16 between macitentan 10 mg and placebo, as estimated from the MMRM.

Missing data at Week 16 and Week 52 arising from any other reason (eg, premature study treatment discontinuation due to COVID-19, missed visits because of COVID-19) are assumed to be missing at random, and as such, will require no further imputation, since the MMRM approach will implicitly impute these values under a missing at random assumption.

The model will include randomized treatment group, time (via a categorical variable for visit), treatment-by-time interaction, baseline-by-time interaction, and mean count per minute at baseline as fixed effects. An unstructured variance-covariance matrix will be specified.

At Week 16, the resulting LS means and corresponding 95% and 99% CLs obtained in each treatment group, and the LS means differences (95% and 99% CLs) for macitentan 10 mg vs placebo will be provided.

3. Subgroup analysis based on the COVID-19 periods of the pandemic.

Subgroup analyses will be carried out similarly as described for the primary efficacy endpoint in Section 5.3.1 above.

## 5.4. Safety Analyses

All safety analyses will be based on the safety analysis set (SS) based on actual treatment received, unless otherwise specified. Listings will be provided on the FAS.

All definitions and conventions used in the CSR SAP section 6 of this study (AC-055H301/RUBATO) will be applied, unless otherwise specified.

### 5.4.1. Adverse Events

COVID-19 related AEs are identified if the corresponding PTs are included in the list as provided in Appendix 2 List of COVID-19 related AEs.

Overall summary table of COVID-19 related AEs will be provided, containing number and percentages of subjects having experienced at least 1 occurrence for the following categories of AEs:

- Treatment-emergent COVID-19 related AEs
- Severe treatment-emergent COVID-19 related AEs
- COVID-19 related AEs with fatal outcome
- Treatment-emergent serious COVID-19 related AEs (SAEs)
- COVID-19 related AEs leading to premature discontinuation of study treatment

A separate summary table will be provided by Special Interest Category (ie, COVID-19 related AEs) and PT for the following categories of AEs:

- Treatment-emergent COVID-19 related AEs
- COVID-19 related AEs with fatal outcome
- Treatment-emergent serious COVID-19 related AEs (SAEs)
- COVID-19 related AEs leading to premature discontinuation of study treatment.

A summary table will be provided, containing number and percentages of subjects who died on the SS. Frequencies for the following parameters will be included in the summary table:

- Number of subjects who died
- Cause of death (eg, *cause term 1, cause term 2, COVID-19 cause term etc*)

The summary will be based on the "Death" eCRF module. The summary will be provided if at least 3 deaths have occurred overall. The summary table is presented by descending order of PT according to the incidence of deaths in the macitentan 10 mg treatment group. An individual subject listing will be provided as well.

In addition to the summary tables, a listing will be provided for subjects who had experienced COVID-19 related AEs.

### 5.4.2. Clinical laboratory test

The results from COVID-19 testing will be captured as COVID-19 related AEs and reported as described in Section 5.4.1 of this document.

Options for the conduct of safety laboratory assessments are described in protocol section 7.2.4.1 and include laboratory samples taken at study site, at a satellite laboratory, collected via home healthcare visits service, or taken and analyzed locally. These alternatives continue to be valid also during the COVID-19 periods of the pandemic (ie, “Pre COVID-19” and “During COVID-19” as defined in Table 3 of Section 5.1.1).

A summary display will show the percentage of laboratory data (considering all parameters) that were conducted at central and local laboratory during the COVID-19 periods by treatment group and overall. This summary output will be presented for all assessments conducted during the course of the study and for key scheduled assessments (ie, Week 16 and Week 52).

## 6. SUPPORTING DOCUMENTATION

### 6.1. Appendix 1 List of Abbreviations

| | |
|---|---|
| AE | Adverse Event |
| ANCOVA | Analysis of Covariance |
| COVID | Corona Virus Disease |
| CSR | Clinical Study Report |
| DPS | Data Presentation Specifications |
| eCRF | electronic Case Report Form |
| FAS | Full Analysis Set |
| FDA | Food and Drug Administration |
| MedDRA | Medical Dictionary for Regulatory Activities |
| MMRM | Mixed Measures Repeated Model |
| PD | Protocol Deviation |
| PT | Preferred Term |
| SAE | Serious Adverse Event |
| SAP | Statistical Analysis Plan |
| SS | Safety Analysis Set |
| WHO | World Health Organization |

### 6.2. Appendix 2 List of COVID-19 related AEs

AEs are considered as COVID-19 related if their coded PTs are: Asymptomatic COVID-19, Congenital COVID-19, Coronavirus infection, Coronavirus test positive, COVID-19, COVID-19 immunisation, COVID-19 pneumonia, COVID-19 prophylaxis, COVID-19 treatment, Exposure to SARS-CoV-2, Multisystem inflammatory syndrome in children, Occupational exposure to SARS-CoV-2, Post-acute COVID-19 syndrome, SARS-CoV-2 antibody test positive, SARS-CoV-2 carrier, SARS-CoV-2 RNA decreased, SARS-CoV-2 RNA fluctuation, SARS-CoV-2 RNA increased, SARS-CoV-2 sepsis, SARS-CoV-2 test false negative, SARS-CoV-2 test positive, SARS-CoV-2 viraemia, Suspected COVID-19, Vaccine derived SARS-CoV-2 infection,

Antiviral prophylaxis, Antiviral treatment, Coronavirus test, Coronavirus test negative, COVID-19 screening, Exposure to communicable disease, Pneumonia viral, SARS-CoV-2 antibody test, SARS-CoV-2 antibody test negative, SARS-CoV-2 RNA, SARS-CoV-2 RNA undetectable, SARS-CoV-2 test, SARS-CoV-2 test false positive, SARS-CoV-2 test negative.