STATISTICAL ANALYSIS PLAN

FOR THE CLINICAL STUDY REPORT

SERENADE

A multi-center, double-blind, placebo-controlled Phase 2b study to evaluate the efficacy and safety of macitentan in subjects with heart failure with preserved ejection fraction and pulmonary vascular disease

Purpose of Analysis	Clinical Study Report
Investigational Drug	Macitentan / ACT-064992
Protocol Number	AC-055G202/ SERENADE
Actelion Document Number	D-17.528
Document Status/Version Number	Final Version 1.0
Date	2 April 2020
Author	PPD, Statistician
	PPD, CRO Statistician (Valos)
Reviewer	PPD, Project Statistician
Reviewer	PPD CRO Lead Expert Statistician (Valos)
Reviewer	PPD Clinical Project Physician
Reviewer	PPD Clinical Trial Scientist
Reviewer	PPD Medical Safety Officer
Reviewer	PPD Senior Clinical Scientist

Confidential

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS AND ACRONYMS

6-MWD	6-minute walk distance
6-MWT	6-minute walk test
ACE	Angiotensin converting enzyme inhibitors
AE	Adverse event
ALT	Alanine transaminase
ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
ARB	Angiotensin Receptor Blockers
AST	Aspartate transaminase
ATC	Anatomic Therapeutic Chemical
BCRP	Breast-Cancer Resistant Protein
BMI	Body Mass Index
bpm	Beats per minute
CDDM	Clinical Development Data Management
CEC	Clinical Event Committee
CI	Confidence interval
CL	Confidence limit(s)
CO	Cardiac Output
COPD	Chronic obstructive pulmonary disease
CRF	Case report form
CRO	Contract research organization
CSR	Clinical study report
CV	Cardiovascular
DAOH	Days Alive and Out of the Hospital
DAU	Daily accelerometer units
DBP	Diastolic Blood Pressure
dPAP	diastolic Pulmonary Arterial Pressure

DPG	Diastolic Pulmonary Vascular Pressure Gradient
DPS	Data presentation specification
dSAP	diastolic Systemic Arterial Pressure
DTS	Data Transfer Specifications
ECG	Electrocardiogram
ECSC	European Community for Steel and Coal
EF	Ejection Fraction
eGFR	estimated Glomerular Filtration Rate
EMA	European Medicines Agency
EOS	End-of-study
EOT	End-of-treatment
ESRD	End Stage Renal Disease
EudraCT	European Clinical Trial Database
FAS	Full Analysis Set
FC	Functional Class
FDA	(US) Food and Drug Administration
FEV	Forced Expiratory Volume
FVC	Forced Vital Capacity
GCP	Good Clinical Practice
HAESI	Hepatic adverse events of special interest
HDL	High-Density Lipoprotein
HF	Heart Failure
HFpEF	Heart failure with preserved ejection fraction
HLGT	High Level Group Term
ICE	Integrated Computer Environment
ICF	Informed Consent Form
ICH	International Council on Harmonisation
IDMC	Independent Data Monitoring Committee

IxRS	Interactive voice/web Recognition System
IRT	Interactive Response Technology
IV	Intravenous
KCCQ	Kansas City Cardiomyopathy Questionnaire
KM	Kaplan-Meier
LAV	Left Atrial Volume
LAVI	Left Atrial Volume Index
LDL	Low-Density Lipoprotein
LTA	liver test abnormality
LVPA	Light to vigorous physical activity
LVEDP	Left Ventricular End Diastolic Pressure
LVEF	Left Ventricular Ejection Fraction
MDRD	Modification of diet in renal disease
MedDRA	Medical Dictionary for Regulatory Activities
MI	Myocardial infarction
MLA	Marked laboratory abnormality
mPAP	mean Pulmonary Arterial Pressure
MRA	Mineralocorticoid receptor antagonists
mRAP	mean Right Arterial Pressure
MRI	Macitentan Run-In
MR-proANP	Mid-Regional pro-Atrial Natriuretic Peptide
mSvO2	mixed venus Oxygen Saturation
NDD	Number of days dead
NDH	Number of days in hospital
NT-pro-BNP	n-terminal pro-Brain Natriuretic Peptide
NYHA	New York Heart Association
OL	Open-label
OLE	Open-label enrollment

PAP	Pulmonary Arterial Pressure
PASP	Pulmonary Artery Systolic Pressure
PAWP	Pulmonary Artery Wedge Pressure
PD	Protocol Deviation
PGA	Patient Global Assessment
PFT	Pulmonary Function Test
PPS	Per-Protocol analysis set
PRI	Placebo Run-In
PT	Preferred Term
PTOP	Post-Treatment Observation Period
PVR	Pulmonary Vascular Resistance
RHC	Right heart catheterization
SAE	Serious adverse event
SAF	Safety set
SAP	Statistical analysis plan
SAS®	Statistical Analysis Software
SBP	Systolic Blood Pressure
SCR	Screened analysis set
SDTM	Study Data Tabulation Model
SI	International system of units
SMQ	Standardised MedDRA Query
SOC	System organ class
sPAP	systolic Pulmonary Arterial Pressure
SpO2	Oxygen saturation
sSAP	systolic Systemic Arterial Pressure
SSAS	Sub-study Analysis Set
TAPSE	Tricuspid Annular Plane Systolic Excursion
TDD	Total Daily Dose

TDI	Total Duration of Interruptions
TEAE	Treatment-Emergent Adverse Event
TFU	Total Expected follow-up time
TSH	Thyroid Stimulating Hormone
ULN	Upper limit of the normal range
USA	United States of America
VLDL	Very Low-Density Lipoprotein
V/Q	Ventilation/perfusion
WHF	Worsening of Heart Failure
WHO	World Health Organization
WHO DDE	World Health Organization Drug Dictionary Enhanced

1 INTRODUCTION

This Statistical Analysis Plan (SAP), for study AC-055G202 / SERENADE, describes in detail the methods, conduct and content of statistical analyses of efficacy, safety and quality of life endpoints planned for the final Clinical Study Report (CSR).

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 those provided by Actelion Clinical Development Data Management (CDDM) via the Integrated Computer Environment (ICE) system.

A separate programming specification and conventions document is prepared in line with this SAP to provide programming details necessary to implement the statistical analysis.

2 STUDY DESIGN AND FLOW

Protocol version 6 (6 February 2020, D-20.020) is used as a reference and the synopsis can be found in Appendix A.

2.1 Study design

This is a prospective, multi-center, double-blind, randomized, placebo-controlled, parallel-group Phase 2b study designed to evaluate whether macitentan 10 mg reduces n-terminal pro-Brain Natriuretic Peptide (NT-pro-BNP) versus placebo at Week 24 in subjects with heart failure with preserved ejection fraction (HFpEF) and pulmonary vascular disease.

It was planned that approximately 300 subjects would be randomized in a 1:1 ratio to either macitentan or placebo. Treatment allocation is stratified by NT-pro-BNP level at macitentan run-in entry (< 1000 pg/mL and ≥ 1000 pg/mL). The study is being conducted in 77 sites in 17 countries.

Following closure of Screening on 27 December 2019, it is expected that approximately 140 subjects will be randomized.

The study comprises of a screening period, a single-blind run-in period, a double-blind treatment period and a safety follow-up period as described below and depicted in Figure 1.

Run-in period:
The run-in period starts with a single-blind placebo run-in of 4 weeks (Visit 2 to the day before Visit 4), followed by a single-blind macitentan run-in of 5 weeks, from the first dose of macitentan at Visit 4 to the subject's randomization at Visit 6. The

placebo and macitentan run-in periods are single-blind as only the subject is blinded to the identity of the treatment.

Treatment period:
The double-blind treatment period consists of a core phase up to Week 24 (Visit 11), followed by an extension phase lasting for 28 weeks, and ending on the day of last study treatment intake.

Subjects who are in core phase at the time of global protocol version 6 approval will stop study treatment at Week 24 (Visit 11), will not proceed to the extension phase of the study and may be eligible to transition to SERENADE OL (AC-055G203) at this time point.

Subjects who have passed Week 24, i.e., are in the extension phase at the global protocol version 6 approval, will return for an end-of-treatment (EOT) visit within 2 months, but no later than Week 52 and will then proceed to enroll in the SERENADE OL study, if eligible, or complete end-of-study (EOS), if not eligible to enter SERENADE OL.

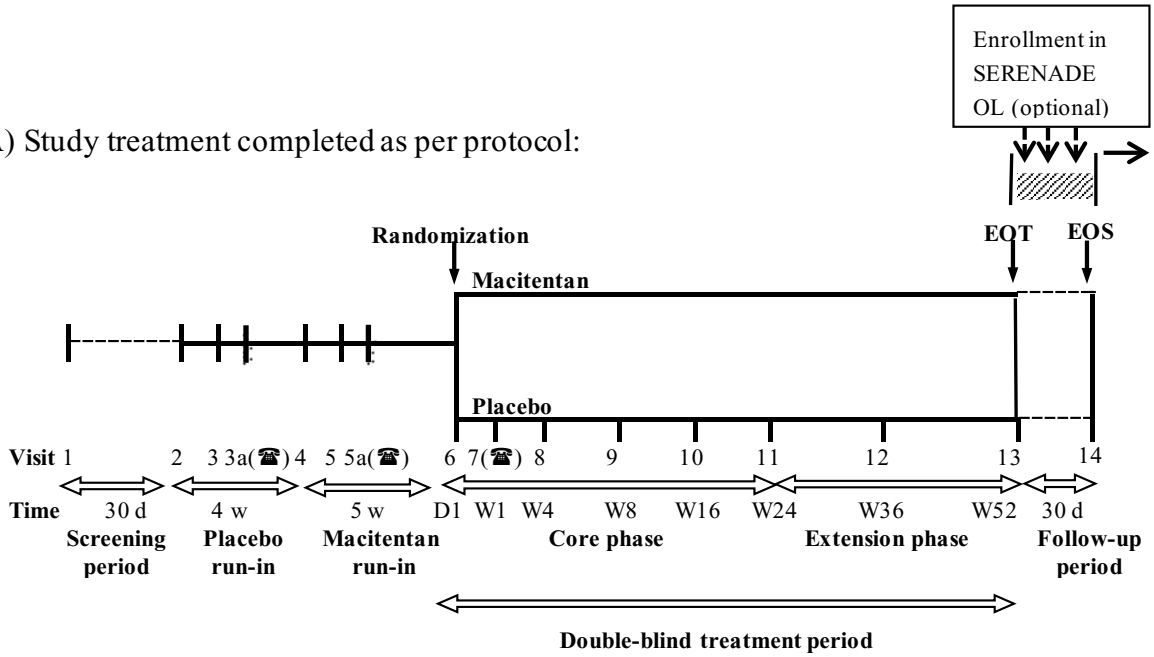
Safety follow-up period:
The safety follow-up period starts on the day after the last dose of study treatment and ends 30 days thereafter with the EOS visit, or Post-Treatment Observation Period (PTOP) PTOP1 visit for those subjects who prematurely discontinue study treatment (see the definition later in this section).

End-of-study:
EOS is reached when the safety follow-up period or, if applicable, PTOP have been completed. For an individual subject, the study is completed with the EOS visit, which is either Visit 14 (safety follow-up visit) for subjects who completed the treatment period as per protocol, or PTOP4 for subjects who prematurely discontinued study treatment (see also section "Post-treatment observation period" below) / PTOP2 (applicable for subjects who consented to global protocol version 6).

SERENADE OL extension (study AC-055G203):
Subjects who remained in SERENADE (main study) for 52 weeks after randomization may be eligible for transition to the SERENADE OL study. Subjects who consented to global protocol version 6, are eligible to transition to SERENADE OL if they remained in the main study for at least 24 weeks.

The overall study design is depicted in Figure 1.

Figure 1 Study design

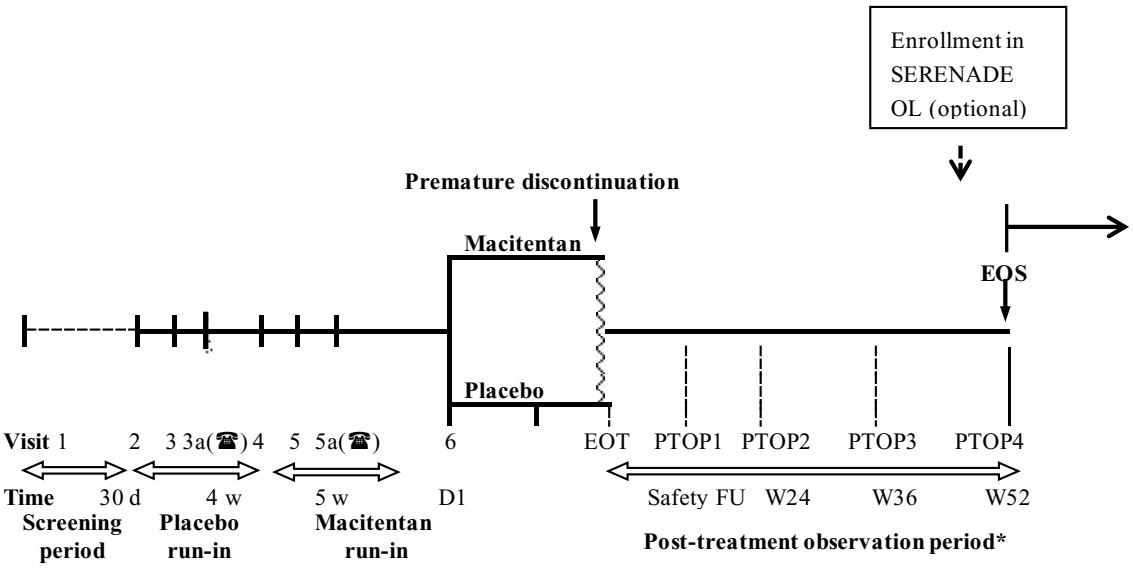


D = Day; d = days, EOT = End-of-treatment, EOS = End-of-study; W = Week; w = weeks.

Post-treatment observation period:
Subjects who prematurely discontinue double-blind study treatment are asked to enter a PTOP, which ends 52 weeks after randomization. The PTOP starts with visit PTOP1, which corresponds to the safety follow-up visit. Thereafter, visits are scheduled at Week 24 (PTOP2), Week 36 (PTOP3) and Week 52 (PTOP4), depending on time point of premature discontinuation. If the safety follow-up visit (PTOP1) falls within the time-window of any of the other PTOP visits, then the corresponding PTOP visit and the safety follow-up visit can be combined (see protocol Table 4 Visit and assessment schedule for subjects entering the PTOP). Subjects who have not completed Week 24 at the time of global protocol version 6 approval, will end the PTOP at Week 24 (PTOP2). Subjects who are past Week 24 at the time of consenting to global protocol Version 6 will return for an EOT visit within 2 months but no later than Week 52 and will then proceed to enroll in the OL study, if eligible, or complete EOS, if not eligible to enter SERENADE OL.

In Figure 2, an example of PTOP is shown for subject prematurely discontinuing the double-blind study treatment at Week 8 (Visit 9).

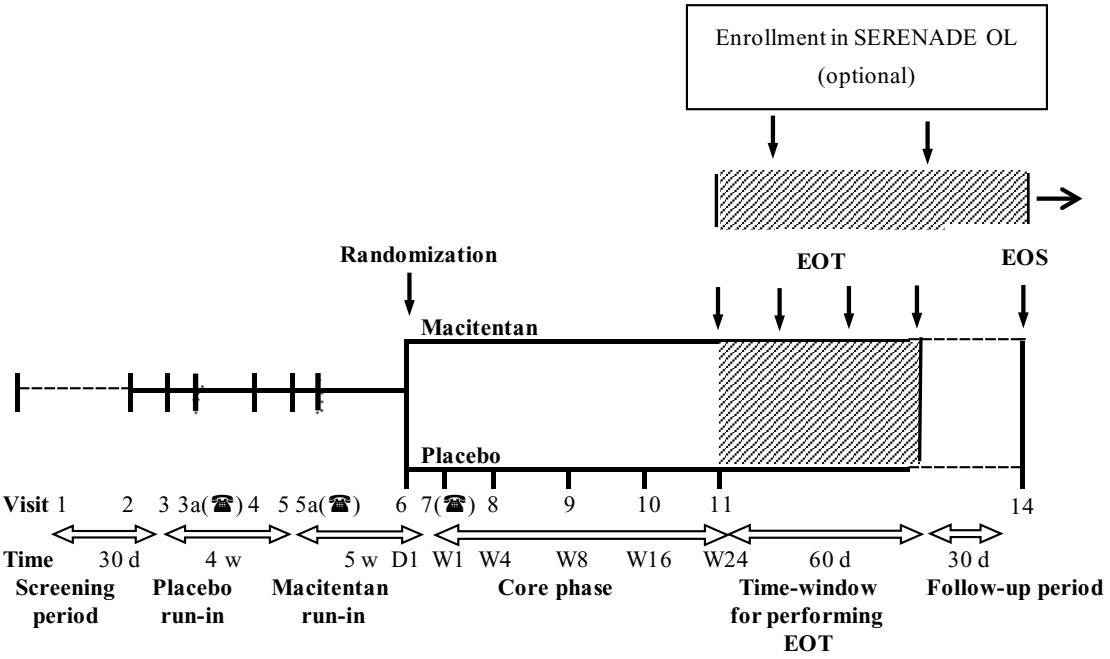
Figure 2 Study design in case of premature discontinuation of double-blind treatment



D = Day; d = days, EOT = End-of-treatment, EOS = End-of-study; W = Week; w = weeks, PTOP = Post-Treatment Observation Period.
*PTOP visits to be performed depending on the time point of premature discontinuation of study treatment. Patients who had not passed Week 24 will remain in PTOP until Week 24.

In Figure 3, an example of discontinuation of double-blind treatment period at Week 24 or beyond, per global protocol version 6.

Figure 3 Discontinuation of double-blind treatment period at Week 24 or beyond, per global protocol version 6



D = Day; d = days; EOT = End-of-Treatment; EOS = End-of-Study; W = Week; w = weeks.

2.2 Study visit and assessment schedule

The schedule of visits and assessments together with the protocol synopsis can be found in Appendix A (protocol Table 1, Table 2, Table 3, and Table 4).

3 OBJECTIVES

3.1 Primary objective(s)

The primary objective of the study is to evaluate whether macitentan 10 mg reduces NT-pro-BNP versus placebo at Week 24 in subjects with HFpEF and pulmonary vascular disease.

3.2 Secondary objectives

The secondary objective of the study is to evaluate the effect of macitentan 10 mg as compared to placebo on:

- Quality of life
- Daily physical activity
- Worsening of heart failure.

3.3 Other objectives

The other objectives are:

- To evaluate the effects of macitentan 10 mg as compared to placebo on:
  - Cardiovascular deaths and hospitalizations
  - New York Heart Association (NYHA) functional class (FC)
  - Clinical composite outcome measure
  - Echocardiographic measures of cardiac function and structure.

3.4 Safety objective

To evaluate the safety and tolerability of macitentan 10 mg in subjects with HFpEF and pulmonary vascular disease.

4 CHANGES OR CLARIFICATIONS TO ANALYSES PLANNED IN THE STUDY PROTOCOL

4.1 Changes to the analyses planned in the study protocol

4.1.1 Geographical regions and other subgroup analyses

Section 10.3.2.5 (subgroup analyses) in the protocol describes four geographical regions (North America, South America, Western Europe, Eastern Europe). Considering the closure of Screening on 27 December 2019, the following three geographical regions are applicable: Americas (merge of North and South America subjects), Western Europe, Eastern Europe.

The following subgroups were added with respect to the protocol section 10.3.2.5:

- Age (18–64, 65–84, ≥ 85 years)
- Renal function at Baseline: normal, mild decrease, moderate decrease, severe decrease / end stage renal disease (ESRD)
- Obesity (body mass index [BMI] ≥ 30 vs < 30).

4.1.2 Baseline

Section 10.3.1 of the protocol defines Baseline as the last non-missing value observed among all measures collected during placebo and macitentan run-ins, up to the day of randomization (Visit 6).

- For accelerometry however, evaluable measures are observed among data collected over 9 days following a visit. Therefore, the last non-missing value

observed up to day of Visit 6 would result in using the assessment following Visit 5, i.e., 9 days starting after 1 week of single-blind macitentan treatment. Instead the following will apply:

  - Accelerometry Baseline [Section 5.5.2.2] is defined as the evaluable measures observed among data collected over 9 days following Visit 3 (during placebo run-in).
  - Visit 3 is selected because these data are obtained while the subject is still on placebo. Visit 3 data are preferred to data collected at Visit 2 (placebo run-in) as any potential changes in heart failure (HF) treatment are expected to occur at the beginning of the placebo run-in, and so the clinical stability of subjects is expected to be better established at Visit 3 as compared to Visit 2.
  - In case of missing evaluable data during placebo run-in following Visit 3, the training period following Visit 2 (placebo run-in) will be used.
  - In the rare event accelerometer data are not evaluable both at/after Visit 2 (placebo run-in) and Visit 3, then accelerometer data collected during the 9 days following Visit 5 will be used.
  - Further details for defining accelerometry Baseline are described in Section 5.5.2.2.

- For safety variables [Section 5.6] the baseline is the last non-missing assessment obtained before or on the day of the start of macitentan single-blind run-in treatment. This is to capture changes in safety after first intake of macitentan and is referred to as the macitentan run-in baseline.

4.1.3 Percent of baseline in MR-proANP over time

Mid-Regional pro-Atrial Natriuretic Peptide (MR-proANP) is collected at Week 24 and Week 52 post baseline. Fewer subjects are expected to have a Week 52 assessment following the protocol amendment (global protocol version 6) to reduce the duration of the study to 24 weeks. Therefore, the random coefficient regression model scheduled in the protocol section 10.3.4.9 will be replaced by an analysis of covariance (ANCOVA) model at Week 24 adjusting for the log of the variable baseline value.

4.1.4 Display of listings

Section 10.1.8 of the protocol specifies that listings will be prepared on the Full Analysis Set (FAS). In order to include all subjects' data, including data from subjects

discontinuing during the single-blind run-in periods and screening failures, listings will be presented using the Screened (SCR) Set or PRI set where applicable [see Section 7.2].

4.2 Changes in the conduct of the study / data collection

4.2.1 The following protocol amendments were performed:

Protocol Version 2, amendment 1 - 8 February 2017

- Implementation of additional safety monitoring measures post-macitentan first dose, as requested by the FDA.

Protocol Version 3, amendment 2 - 12 April 2017

- Update of exclusion criteria and forbidden medication to address FDA request to avoid co-administration of macitentan with breast-cancer resistant protein (BCRP) substrate medications.

Protocol Version 4, amendment 3 - 10 April 2018

- To define EOS for subjects who are eligible to enter the SERENADE OL extension study (AC-055G203)
- To establish a Clinical Event Committee (CEC) to review and adjudicate worsening of heart failure events, as well as all reasons for hospitalization and death.
- To revise the inclusion and exclusion criteria and run-in failure criterion 3b.
- Changes made to statistical methods section to account for: subjects entering the SERENADE OL; the introduction of CEC; and the amendment to allow NT-pro-BNP values measure at screening to be used for stratification at the time of treatment assignment
- Change to definition of placebo run-in analysis set and macitentan analysis set.

Protocol Version 5, amendment 4 - 8 March 2019

- To introduce a new efficacy endpoint, the 6-minute walk test (6-MWT), as part of a sub-study to assess the change from baseline in exercise capacity [see Section 5.7].
- To remove the 8-hour post-macitentan first-dose safety monitoring period at the start of macitentan run-in and add 2 safety phone calls during the run-in period.
- To change the testing hierarchy of the key secondary efficacy endpoints. The Kansas City Cardiomyopathy Questionnaire (KCCQ) was moved to the first

secondary endpoint in the hierarchy, with accelerometry (which used to be the first secondary endpoint) moved then to the second position, followed by 'time to worsening HF events'.

- To introduce a new hierarchical composite exploratory efficacy endpoint which combines hard (death, hospitalizations) and soft (functional capacity, quality of life) endpoints to allow for a more complete and broader assessment of clinical benefit of macitentan in HFpEF. The composite endpoint will be analyzed using the win ratio approach that accounts for clinical priorities [Pocock 2012]. This was subsequently removed in protocol Version 6 due to the reduced sample size, and stopping assessment of functional capacity (6MWD).

Protocol Version 6, amendment 5 - 6 February 2020

- The early termination of recruitment into the study since the study failed to meet subject recruitment targets and completion of the study within reasonable timeline is not realistic. The sample size was updated to reflect the early termination of recruitment.

- To reduce the length of the double-blind treatment period to 24 weeks. The rationale for this change is that Week 24 is the pre-defined time point for assessing the primary endpoint, as well as the key secondary endpoints. The secondary endpoint of time to worsening heart failure (WHF) event however was planned to be assessed up to Week 52 to gather meaningful information for the preparation of a pivotal clinical trial development program. Due to the reduced sample size, the number of WHF events is expected to be too low for a meaningful analysis of time to WHF. The double-blind treatment period will thus be stopped at Week 24, and eligible subjects will be able to transition to SERENADE OL at this time point. Subjects who have already completed the Week 24 visit, will be scheduled to come back for an EOT visit within 60 days and will proceed to enroll in the OL study, if eligible.

- To remove the CEC which was appointed to review and adjudicate in a blinded fashion WHF events, the reasons for hospitalization and causes of death. The rationale is based on the reduction of the double-blind treatment period from 52 weeks to 24 weeks, coupled by the low occurrence of WHF events which will not allow for meaningful conclusions to be drawn. Removal of the CEC does not affect safety monitoring and therefore the decision was also endorsed by the Independent Data Monitoring Committee (IDMC).

- To re-schedule accelerometry to be performed 9 consecutive days prior to Week 24 (Visit 11/PTOP2) for subjects completing Week 24 (Visit 11) under Amendment 5 to ensure assessment performed on double-blind treatment.
- To stop sub-study assessments (6-MWT and Borg Dyspnea Index), as number of subjects participating in the sub-study is too low to allow for meaningful interpretation of results.

Note: Only the main reasons for the protocol amendments are described. No subjects were enrolled prior to amendments 1 and 2.

4.3 Clarifications concerning endpoint definitions and related variables or statistical methods

4.3.1 Previous and concomitant medications

Section 5.2.1 of the protocol specifies definitions for a previous therapy, a study-concomitant therapy and a study treatment-concomitant therapy for the purpose of capturing these in the case report form (CRF). The definitions for previous therapy and treatment concomitant therapy are adapted in this SAP Section 5.2.5 for the purpose of presenting summaries for reporting and interpreting the results in the CSR.

The SAP changes to the protocol are as follows:

- Previous therapy is relative to first dose rather than signing of informed consent.
- Treatment concomitant therapy period is extended up to 30 days after study treatment end.
- Study-concomitant therapy is not required for reporting of results in the CSR.

4.3.2 Efficacy endpoints evaluated over 52 weeks

Sections 6.1.2 and 6.1.3 of the protocol specify efficacy event endpoints as evaluated over 52 weeks. Prior to global protocol version 6 subjects were followed up to Week 52, except where subjects prematurely discontinue study. Under global protocol version 6 the follow-up period was reduced to 24 weeks, except where subjects had already passed Week 24 (Visit 11). Therefore, it is expected that the follow-up will be varied across subjects. All subjects visit assessment information available will be considered up to Week 52.

The time to event variables in sections 10.2.2.3 and 10.2.3 of the protocol are defined as events occurring until EOS/ open-label enrollment (OLE). In order not to lose any information and reduce risk of bias or informative censoring and due to the shorter

follow-up time for subjects enrolled under global protocol version 6, the approach to EOS/ OLE is applied consistently where applicable in the CSR SAP.

5 DEFINITIONS OF VARIABLES

This section provides the definitions and sources for all variables used in the analyses, including specifications for the derivations.

General recurrent definitions (e.g., study treatment start date, baseline) or unit conversion are described in detail in Section 11.

5.1 Screening and run-in failures and disposition

5.1.1 Screening failure

Screening failures are all subjects with answer 'NO' to the question 'Did the subject enter the single-blind run-in period?' in the "Eligibility for Placebo Run-in" CRF.

The primary reason for screening failure is provided in the same CRF ('Subject withdrew consent', 'Not eligible as per inclusion/exclusion criteria', 'Lost to follow-up' or 'Other'). In the rare case that the reason for screening failure is not provided in the CRF for a subject considered a screening failure, the reason is categorized as 'Unknown'.

It is permitted to re-screen subjects twice, if the reason for non-eligibility is transient (e.g., abnormal laboratory test, insufficient wash-out period for a forbidden medication). All screening assessments except pulmonary function tests (PFTs) must be repeated at the time of re-screening. If the screening echocardiography was done within 2 months of the re-screening visit, it doesn't need to be repeated.

Subjects who failed screening more than once will be counted only once (and only the reason for the last screening failure will be included in the summary table). If a subject is screened more than once and subsequently enters the single-blind placebo run-in period, the subject is not considered a screening failure.

5.1.1.1 Unmet eligibility criteria for Placebo Run-in

Inclusion criteria not met, and exclusion criteria met are collected in the CRF for subjects who are not eligible for placebo run-in as per the inclusion/Exclusion criteria (CRF; Eligibility for Placebo Run-in). For re-screened subjects it is taken from the latest re-screening. Note a subject may have multiple criteria which are reported separately to the primary reason for screening failure described above in Section 5.1.1.

5.1.2 Single-blind placebo run-in disposition and failures

A subject is considered to have withdrawn during placebo run-in / not entered into macitentan run-in if he/she entered the single-blind placebo run-in period but eventually did not enter the single-blind macitentan run-in period as identified from the answer to the CRF question "Did the subject enter the macitentan run-in period?" (CRF; Eligibility - Macitentan Run-in).

If the answer to this question is "No", the reasons provided in the CRF: 'Subject withdrew consent', 'Not eligible as per inclusion/exclusion criteria', 'Placebo run-in failure', 'Lost to follow-up' or 'Other' (CRF; Eligibility - Macitentan Run-in) are used as the primary reason for not entering macitentan run-in period.

Run-in failure criteria are collected on the CRF for subjects who met any run-in failure criteria during the placebo run-in (CRF; Eligibility - Macitentan Run-in).

5.1.2.1 Unmet eligibility criteria for Macitentan Run-in

Exclusion criteria are collected on the CRF for subjects who met any of the exclusion criteria during the placebo run-in (CRF; Eligibility - Macitentan Run-in). Note a subject may have multiple criteria which are reported separately to the primary reason for not entering macitentan run-in described above in Section 5.1.2.

In addition, subjects on a stable dose of heart failure (or cardiovascular) medication for 7 days prior to the Macitentan Run-in Visit 4 are captured on the same CRF (as "yes" /" no").

5.1.3 Single-blind macitentan run-in disposition and failures

A subject is considered to have withdrawn during macitentan run-in / not entered into double-blind if he/she entered the single-blind macitentan run-in period but eventually was not randomized as identified from the answer to the question "Was the subject randomized?". If the answer to this question is "No", the reasons provided in the CRF: 'Subject withdrew consent', 'Not eligible as per inclusion/exclusion criteria', 'Macitentan run-in failure', 'Lost to follow-up' or 'Other' (CRF; Randomization) are used as the primary reason for not entering double-blind period.

Run-in failure criteria are collected on the CRF for subjects who met any run-in failure criteria during the macitentan run-in (CRF; Eligibility - Double-blind Treatment Period).

5.1.3.1 Unmet eligibility criteria for Double blind treatment period

Exclusion criteria are collected on the CRF for subjects who met any of the exclusion criteria during the macitentan run-in (CRF; Eligibility - Double-blind Treatment

Period). Note a subject may have multiple criteria which are reported separately to the primary reason for not entering double-blind described above in Section 5.1.3.

In addition, a subject on stable cardiovascular medication for 7 days prior to Visit 6 (randomization) is captured on the same CRF (as "yes" /" no").

5.2 Subject characteristics

If a subject is screened more than once, only the data assessed during the latest successful re-screening attempt (subject did not fail screening / entered placebo run-in) are displayed and included in any analysis. Exceptions to this rule are applied where data from earlier visits may be used if not available at the latest successful re-screening attempt as follows:

- Historical data such as right heart catheterization (RHC) and PFTs, the data coming from an earlier screening are included in the corresponding analysis if not available at the successful (re-)screening attempt.
- For echocardiography, the data coming from an earlier screening visit are considered.
- NT-pro-BNP (from central laboratory only).
- Characteristics that will not change including sex, race, ethnicity and height may be used from previous screening visits.

5.2.1 Demographics

Demographic data at screening comprise age (years; continuous and categorical), sex, ethnicity, race, height (cm), weight (kg) and BMI (kg/m2; continuous and categorical). Demographics of the subjects are taken from the Screening visit (Visit 1) or Re-Screening visit (age, weight, height, BMI) where subject is re-screened.

The following categories are defined for demographic variables:

- **Age (years)**: 18–64, 65–84, ≥ 85

Note the above categories for age are the same as those required for disclosure to EudraCT and ClinicalTrials.gov.

- **BMI (kg/m^2):** < 18.5; ≥ 18.5 – < 25; ≥ 25 – < 30; ≥ 30 – < 40; ≥ 40
- **BMI (kg/m^2):** < 30; ≥ 30, for the obesity evaluation
- **Race** (Black or African American / American Indian or Alaska Native / Native Hawaiian or Other Pacific Islander / Asian / White / Other / Not applicable)

- **Geographical region:** considering the closure of Screening on 27 December 2019, the following geographical regions are applicable:
  - Americas: North America (USA), South America (Argentina and Brazil)
  - Western Europe: Austria, Denmark, France, Germany, Spain, Sweden, United Kingdom. In addition, Israel is also considered part of this region
  - Eastern Europe: Bulgaria, Czech Republic, Hungary, Poland, Romania, Russia.

Age, sex, race, and ethnicity are collected in the "Demographics" CRF. Country is determined from the site number and provide in the SDTM.DM dataset.

The values of height, weight and BMI are recorded in the "Height, Body Weight & BMI" CRF. Conversion from imperial units to metric (international system of units [SI]) units is performed automatically in the CRF applying the conversion rules detailed in Section 11.9.

BMI is calculated automatically in the CRF using the formula specified in Section 11.9 and will not be (re-) calculated for the analysis.

5.2.2 Baseline disease characteristics

The baseline disease characteristics are listed below:

Specific Medical History ("Specific Medical History" CRF):

- **Time since most recent hospitalization for heart failure** (years): expressed in number of years and determined in relation to the date of screening. It is calculated by firstly taking the difference between the date of screening and the date of most recent HF hospitalization, and secondly by dividing the result by 365.25.
- **Number of heart failure hospitalizations within the last 12 months**
- **Anemia** (Yes / No)
- **Hypertension** (Yes / No)
- **Diabetes mellitus type II** (Yes / No)
- **Chronic obstructive pulmonary disease (COPD;** Yes / No)
- **Atrial fibrillation/flutter** (Yes / No): subject is classified as having atrial fibrillation/flutter (Yes) if atrial fibrillation = Yes or atrial flutter = Yes, regardless if ongoing at screening or not. If both atrial fibrillation and atrial flutter are missing then atrial fibrillation/flutter is classified as missing, otherwise = No.

- **Ischemic heart disease (Yes / No)**

New York Heart Association Functional Class - NYHA FC ("NYHA Functional Class" CRF): only Class II and III are allowed per inclusion criteria.

Echocardiography (provided by central echocardiographic laboratory):

- **Left Ventricular Ejection Fraction (LVEF) [%]:** classified as ≥ 40 - <50 and ≥ 50%. These categories represent the mid-range Ejection Fraction (EF) and preserved EF. In the unexpected case we have LVEF < 40%, this sub-category will also be displayed.
- **Left Atrial Volume Index (LAVI)** [mL/m^2]
- **Lateral E/e'**
- **Pulmonary Artery Systolic Pressure (PASP)** [mmHg].

Laboratory:

- **NT-pro-BNP** (pg/mL): value at screening from central laboratory only
- **NT-pro-BNP** used as stratification factor as per the Interactive voice/web Recognition System (IxRS) value provided as continuous and categorical (< 1000 pg/mL and ≥ 1000 pg/mL)
- **Creatinine** *(*µmol/L, *SI units)*
- **Estimated Glomerular Filtration Rate(eGFR)** (mL/min/1.73m^2) using the Modification of Diet in Renal Disease formula [see Section 11.9].

Presence of valvular disease ("Medical History" CRF):

Any disease or diagnosis corresponding to one of the following preferred terms (PTs): 'Mitral Valve Repair', 'Mitral Valve Replacement', 'Mitral Valve Valvotomy', 'Mitral Valve Commissurotomy', 'Aortic Valve Repair', 'Aortic Valve Replacement', 'Aortic Valve Valvotomy'. The definitions are based on a combination of MedDRA PTs) and Standardized MedDRA Queries (SMQs) which can be updated according to MedDRA version / SMQs updates and documented in the data presentation specification (DPS).

Signs and symptoms of special interest ("Signs/Symptoms of special interest" CRF):

- **Peripheral edema** (Yes / No / Not assessed)
- **Jugular venous distension** (Yes / No / Not assessed).

5.2.3 Other baseline characteristics

Other baseline characteristics include the following:

- **Pulmonary function tests** ("Pulmonary Function Tests" CRF), including results of test performed during screening and historical PFT performed within 12 months prior to Screening:
  - FEV_1 (L)
  - FEV_1 predicted (L)
  - FEV_1 percent of predicted value (%)*
  - Forced Vital Capacity (FVC) (L)
  - FEV_1/FVC (ratio)*, equations are those from the European Community for Steel and Coal (ECSC) [Quanjer 1993].

  *Calculated values provided in the CRF are not (re-) calculated.

- **Right heart catheterization** ("Right Heart Catheterization" CRF) - where historical RHC values are available these are collected at Screening on the CRF form:
  - Heart rate (bpm)
  - Cardiac Output (CO) [L/min]
  - Systolic Pulmonary Arterial Pressure (sPAP) [mmHg]
  - Diastolic Pulmonary Arterial Pressure (dPAP) [mmHg]
  - Mean Pulmonary Arterial Pressure (mPAP)[mmHg]
  - Mean Right Atrial Pressure (mRAP)[mmHg]
  - Pulmonary Artery Wedge Pressure (PAWP)[mmHg]
  - Left Ventricular End Diastolic Pressure (LVEDP)[mmHg]
  - Systolic Systemic Arterial Pressure (sSAP) [mmHg]
  - Diastolic Systemic Arterial Pressure (dSAP) [mmHg]
  - Mixed Venus Oxygen Saturation (mSvO2) at rest [%]

The following RHC values are calculated automatically within the CRF according to the algorithm in Section 11.9:

- Pulmonary Vascular Resistance (PVR) [wood units]*
- Diastolic Pulmonary Vascular Pressure Gradient (DPG) [mmHg]*.

*Calculated values provided in the CRF are not (re-) calculated.

- **Renal function**: the following categories are defined for renal function based on the eGFR (mL/min/1.73m^2) from the abbreviated modification of diet in renal disease (MDRD) study [FDA 2010]:
  - Normal: ≥ 90

- Mild Decrease: 60 – < 90
- Moderate Decrease: 30 – < 60
- Severe Decrease: 15 – < 30
- ESRD: < 15.

5.2.4 Medical history

Medical history is collected during the Screening visit (Visit 1, or re-screening visit where subject is re-screened) on the "Medical History" CRF. The original terms reported by the investigators are coded using the latest implemented version of MedDRA.

Note: specific medical history [Section 5.2.4.1] are not collected on the main medical history CRF. Therefore, both sources of information are used in the general summary of medical history using the coded terms.

5.2.4.1 Specific medical history

Pre-defined specific medical history, including current medical conditions of special interest, for this population are collected during the Screening visit (Visit 1) on the "Specific Medical History" CRF. These predefined terms are also coded according to MedDRA.

5.2.5 Previous and concomitant therapies

All therapies as collected in the "Previous / Concomitant Medication" CRF including the therapies collected for subjects entering the PTOP.

The original terms used by the investigators to describe therapies are assigned PTs and anatomic therapeutic chemical (ATC) classification code using the latest version of the WHO Drug dictionary enhanced (WHO DDE).

Start and end dates that are incomplete or missing are handled according to the rules in Section 12.

5.2.5.1 Previous therapies

Previous therapy is defined as any therapy that was started prior to start date of study treatment and with an end date prior to the start date of study treatment (i.e., first dose of single-blind placebo run-in treatment, see definition in Section 11.2.4) or end date equal to start of study treatment with the flag 'Ongoing at start of treatment?' ticked "No".

If the end date is missing and the start date is prior to the date of study treatment start, then the therapy is considered as previous if 'Ongoing at start of treatment?' is ticked "No".

Rules for handling missing and partial dates are detailed in Section 12.

5.2.5.2 Baseline therapies

Baseline therapy is defined as any therapy that was started prior to start date of study treatment with an end date after start date of study treatment or the end date is missing and 'Ongoing at start of treatment?' ≠'No'. It is used in describing therapies of special interest [see Section 5.2.5.6.].

5.2.5.3 Study-concomitant therapies

Not applicable, not required for presentation of therapies.

5.2.5.4 Study-treatment concomitant therapies

Study-treatment concomitant therapy considers study treatment overall periods including single-blind placebo and macitentan run-in periods and double-blind treatment period.

Study-treatment concomitant therapies are all the therapies that are ongoing at start of study treatment or initiated on or after study treatment start date and up to study treatment end (EOT) date [Section 11.2.5] + 30 days (inclusive).

5.2.5.5 Post Study-treatment therapies

Post study-treatment concomitant therapies are all the therapies that are initiated after the EOT date + 30 days.

5.2.5.6 Therapies of special interest

Therapies of special interest are based on recommendations for treatment of patients with HFpEF and include the following:

- Loop diuretics, thiazides, other diuretics [see definition in Appendix B]
- Beta blockers (Level 2 text: Beta blocking agents / ATC code: C07A)
- Calcium channel blockers (Level 2 text: calcium channel blockers / ATC code: C08C and C08D)
- ACE inhibitors or ARB (one category) (Level 2 text: AGENTS ACTING ON THE RENIN-ANGIOTENSIN SYSTEM / ATC codes: C09A, B, C and D)
- MRAs (Level 2 text: POTASSIUM-SPARING AGENTS / ATC code: C03D)
- Cardiac glycosides (ATC code: C01A).

The definitions above can be updated according to WHO DDE updates.

5.2.5.7 Diuretics

Diuretics are determined from all therapies as collected on "Previous / Concomitant Medication" CRF with ATC code starting with C03 (DIURETICS) and where ingredients include the diuretics listed in Appendix B [Subset of ingredients selected from Standardised Drug Groupings (Diuretics)]. These selected diuretic therapies are then coded according to the ingredient so that for example "Furosemide sodium" and "Furosemide" are both assigned type = "Furosemide".

Note for combination diuretics each component will be entered in the CRF separately together with the respective dose of each component so that the diuretic type and corresponding total daily dose (TDD) can be determined and to avoid double counting.

Diuretics are further partitioned for placebo run-in, macitentan run-in and double-blind treatment periods to determine frequency of changes within a period. They are defined as:

Placebo run-in concomitant diuretics: diuretic therapy initiated during the placebo run-in period [see definition in Section 11.5.3].

Macitentan run-in concomitant diuretics: diuretic therapy initiated during the macitentan run-in period [see definition in Section 11.5.4].

Double-blind treatment concomitant diuretics: diuretic therapy initiated during the double-blind treatment period [see definition in Section 11.5.2].

Baseline diuretics for placebo run-in: includes diuretic(s) taken on day prior to first dose of placebo run-in treatment [see Section 11.2.6].

Baseline diuretics for macitentan run-in: includes diuretic(s) taken on day prior to first dose of macitentan run-in treatment [see Section 11.2.8].

Baseline diuretics for double-blind treatment: includes diuretic(s) taken on day prior to first dose of double-blind treatment [see Section 11.2.10].

Changes (increase / decrease) in diuretics in a period compared to the respective baseline are derived for the placebo run-in, macitentan run-in and double-blind concomitant diuretics as follows:

Increases in diuretic categorized as:

- Increase total daily dose (TDD): TDD (at the time of new diuretic prescription) > baseline TDD. Calculated for oral diuretics only. Each time this happens is considered as an increase;
- New diuretic where the diuretic type(s) ≠ baseline diuretic type(s) (each time this happens is considered as an increase);
- New IV diuretic where route = INTRAVENOUS (each time this happens is considered as an increase).

Decreases in diuretic:

- TDD: TDD at the time of new diuretic prescription < baseline TDD. Each time this happens is considered as a decrease.

Note any interruption of a diuretic and its re-introduction at the same dose is not considered as a change.

TDD mg/day: is calculated for all oral diuretics according to frequency and the dose (mg). Where dose units are recorded as g or ug, the dose is converted to mg by multiplying or dividing by a factor of 1000, respectively. Other reported units cannot be converted to mg and therefore TDD cannot be derived and is set to zero. TDD (mg/day) is derived for each oral diuretic using coded values as follows:

Frequency (coded)[a]	Total daily dose (mg / day)
Once a day	Dose
Twice per day	2 × dose
<X> times per day	<X> × dose
Every other day	Dose / 2
Twice per week	2 × dose / 7
<X> times per week	<X> × dose / 7
Every week	dose / 7
Every <X> weeks	Dose /(7 × <X>)
Every month	Dose / 30.5
Every year	Dose / 365
Missing, other codes	0

a: values coded to CONTINUOUS, ONCE, AS NEEDED, UNKNOWN and OTHER cannot be used to determine TDD and therefore TTD will be set to zero.

Diuretics given via any other route (or where missing route) are not used in deriving TDD. TDD is set to zero if the above frequency coded terms are not assigned or if either frequency or dose are missing or not coded. Where more than one diuretic of the same type is taken in the same period / overlapping period then these are additive.

Diuretics of different types may not be summed for the TDD, i.e., furosemide 80 mg TDD should not be combined with Spironolactone 100 mg TDD.

The start / end date of the diuretic flagged as a change identifies the duration of the change in the diuretic therapy. The start date is used to determine number of changes in the defined periods below. The end date is used in determining proximity timing of change relative to other study assessment including NT-pro-BNP at wWeek 24.

5.2.6 Other subject characteristics

- **Women of childbearing potential (Yes / No):** collected on the "Demographics / Re-Screening Demographics" CRFs together with reason for not being of childbearing potential. Any change in status is collected on the "Women of Childbearing Potential - Status Change" CRF.

- **Contraceptive methods:** For females of child bearing potential, the contraceptive method used is recorded on dedicated "Contraception methods" CRF. Options include hormonal contraceptives as well as other birth control methods. The original terms used to describe hormonal contraceptives are coded using PTs from the most recent version of the WHO Drug dictionary.

All contraceptive methods are listed together with the data for women of childbearing potential.

5.3 Study treatment exposure and compliance

5.3.1 Exposure

The exposure is evaluated in terms of study treatment duration including study treatment interruptions and then in terms of actual weeks exposed to study treatment, excluding any interruptions. The double-blind treatment is the main period of interest to be evaluated. Other treatment periods are defined to evaluate the total exposure to macitentan for all subjects during the study and to assess compliance during macitentan run-in period. The data are recorded in the "Study Drug Log" CRF. See Section 11.2 for the definitions of treatment start / end date.

Duration of double-blind treatment (weeks): time (in days) elapsed between the double-blind treatment start date and the double-blind study treatment end date + 1 day divided by 7, regardless of any treatment interruptions, i.e., [(Double-blind treatment end - Double-blind treatment start +1)/7].

Exposure of double-blind treatment (weeks): treatment duration adjusted for interruption defined as:

Duration of double-blind treatment - Total Duration of Interruptions (TDI),

where TDI is defined as the sum of duration of interruption over all interruption periods in double-blind treatment. For a single period, the duration of interruption (days) is time elapsed between the Interruption Start Date and the Interruption End Date, i.e., (Interruption End Date - Interruption Start Date), where Interruption Start Date is the Study Drug Log End Date with corresponding reason for treatment end being temporarily interrupted. Interruption End Date is the next chronological Study Drug Log Start Date after Interruption Start Date. In the event of partial or missing treatment start or end date, exposure is missing. If Interruption End Date cannot be derived as there is no next chronological Study Drug Log entry the exposure is missing.

Exposure of double-blind treatment up to Week 24 (weeks): is derived in the same way as above except double-blind treatment end date is replaced by Week 24 (Visit 11) date for subjects who continue treatment beyond Week 24. This exposure is adjusted only for interruptions occurring prior to Week 24 (Visit 11) date. If an interruption has start date < Visit 11 date and end date > Visit 11 date, then only the difference between Visit 11 date and start date is used for that interruption.

Cumulative subject year exposure (years): is calculated by summing the duration of double-blind treatment for all subjects divided by 52.

Total duration of macitentan treatment (weeks): is calculated, for all subjects who receive at least 1 dose of macitentan, as the time (in days) elapsed between the single-blind macitentan treatment start date and the last dose date of macitentan (double-blind or single-blind) + 1 day divided by 7, regardless of any treatment interruptions.

Duration of singe-blind placebo run-in treatment (weeks): is the time (in days) elapsed between the single-blind placebo treatment start date and the single-blind study treatment end date + 1 day divided by 7, regardless of any treatment interruptions.

Duration of singe-blind macitentan run-in treatment (weeks): is the time (in days) elapsed between the single-blind macitentan treatment start date and the single-blind macitentan treatment end date + 1 day divided by 7, regardless of any treatment interruptions.

5.3.2 Compliance with study treatment

Double-blind study treatment compliance (%) is based on the study drug dispensing and accountability data recorded in the CRF "Study Drug Dispensing & Accountability". It is determined using drug accountability data for the double-blind treatment, calculated as the total tablets dispensed at any time from Visits 6 - Visit 12 minus total tablets returned for the corresponding bottle dispensed:

Double-blind study-treatment compliance =

[(total number of tablets dispensed at any time from Visit 6 onward– total number of tablets returned from Visit 8 onward) / Total number of tablets that should have been taken during double-blind period] × 100.

The number of capsules that should have been taken during the double-blind treatment period is double-blind treatment duration (in days) as defined in Section 5.3.1, since macitentan 10 mg is taken once daily.

Double-blind study treatment compliance (%) up to Week 24 is similarly calculated.

Double-blind study-treatment compliance up to Week 24 =

[(total number of tablets dispensed from Visit 6 to Visit 10 – total number of tablets returned from Visit 8 to Visit 11) / Total number of tablets that should have been taken during double-blind period within Week 24] × 100.

The number of capsules that should have been taken during the double-blind treatment period up to Week 24 is double-blind treatment duration up to Week 24 (in days) as defined in Section 5.3.1.

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

Compliance to macitentan single-blind treatment during run-in period: is calculated in the CRF to assess this run-in criterion and will not be (re-) calculated for the analysis.

5.3.3 Study treatment discontinuation

Study treatment discontinuation for each period are identified in CRF as occurring during the single-blind placebo run-in, single-blind macitentan run-in, or during the double-blind treatment period. Premature discontinuation are collected in the "Study Drug Log" CRF are identified as those with a treatment end date and associated

reason 'What was the reason for treatment end?' answered 'Premature Discontinuation'. For example, a subject is considered to have prematurely discontinued double-blind study treatment if the 'reason for treatment end' in the Study drug Log CRF is 'PREMATURE DISCONTINUATION' and the 'Study Period' is 'DOUBLE-BLIND TREATMENT'.

Reason for study treatment discontinuation are collected in the "Premature Discontinuation of Study Treatment' CRF, full list of reasons is as follows:

- Death
- Lost to follow-up
- Pre-specified study treatment discontinuation criteria
- Exclusion criteria met during run-in
- Run-in failure
- Subject decision (adverse event [AE], Lack of efficacy, No reason provided, Other)
- Physician decision (AE, Lack of efficacy, Other)
- Sponsor decision (Study termination, Other).

Note: with global protocol version 6, subjects will stop study treatment at Week 24 (Visit 11) and may be eligible to transition to SERENADE OL at this time point. Subjects who have passed Week 24 at the time of consenting to global protocol version 6 will be scheduled to come back for an EOS visit within 2 months but no later than Week 52 and will then proceed to enroll in the OL study, if eligible.

The reason for treatment discontinuation to be recorded in the eCRF for these subjects is 'sponsor decision', unless an additional reason led to discontinuation of study treatment (e.g., AE).

Note: a subject should prematurely discontinue treatment only once during the study, i.e., should not enter subsequent treatment periods following a premature discontinuation.

Subjects are evaluated by the investigator for completion of treatment within each treatment period via the CRF where the question 'What was the reason for treatment end?' is answered 'Completed as per protocol'.

5.3.4 Study treatment adjustments or interruptions

Study treatment dose adjustments are not permitted.

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' (CRF; "Study Drug Log"). Study treatment interruptions are identified in CRF for the treatment period: single-blind placebo run-in, single-blind macitentan run-in or double-blind treatment.

5.4 Study discontinuation

As per Section 8 of the protocol, a subject who completes 52 weeks of follow-up including the safety follow-up period is considered to have completed the study as per protocol, regardless of whether he/she has completed the double-blind treatment period as per protocol (subjects who prematurely discontinue double-blind study treatment are asked to enter the post-treatment observation period identified with available PTOP visit date).

Therefore:

- subjects entering the double-blind treatment period who did not prematurely discontinue the treatment [Section 5.3.3] with available EOS / Visit 14 [Section 11.5.8] are considered completers;
- subjects prematurely discontinuing the double-blind study treatment who entered PTOP with available PTOP4 Visit are considered completers;
- subjects who enter the SERENADE OL are considered completers, except where a subject has a protocol deviation assigned (PD226) indicating that the subject did not complete the 52 weeks of follow-up before entering the SERENADE OL (this deviation is not applicable to subjects who signed the Informed Consent Form (ICF) v6 prior to entering OL extension) or, if the subject signed the ICF v6, he/she has a protocol deviation assigned (PD336) indicating that entered the SERENADE OL prior to Week 24 (i.e., Visit 11 or PTOP2);
- subjects who do not enter the double-blind treatment period / are not randomized should continue with safety follow-up and complete EOS / Visit 14. These subjects are not considered completers in the summary of disposition.

For subjects who prematurely discontinue the study, the reasons for study discontinuation are reported on a dedicated "Study Discontinuation" CRF and are as follows:

- Death

- Lost to follow-up
- Subject decision/Withdrawal of consent (AE, Lack of efficacy, No reason provided, Other)
- Physician decision (AE, Lack of efficacy, Other)
- Sponsor decision (Study termination, Other).

The date of study discontinuation corresponds to the reason; date of death (as entered in the "Death" CRF); date of last successful contact for subjects lost to follow-up; date of subject decision, physician decision or when the subject was informed of sponsor decision as reported in the CRF "Study Discontinuation".

In the rare event a subject does not complete the study according to the definition above and there is no entry in the "Study Discontinuation" CRF, the subject is classified as a prematurely discontinuation with reason given as "unknown".

5.4.1 Study disposition by visit

With the change in duration of double-blind treatment period per global protocol version 6 (see protocol section 3.1.1.8) subject follow-up will be more varied as:

1. Subjects will stop study treatment at Week 24 (Visit 11).
2. Subjects who have passed Week 24 at the time of consenting to global protocol version 6 will be scheduled to come back for an EOT visit within 2 months but no later than Week 52.
3. Subjects previously completed study treatment at Week 52 (Visit 13), protocol version 5 or earlier.
4. Subjects may prematurely discontinue study at any time.

Therefore, subject meeting criteria 1 above will have visit assessments up to Week 24. Subjects meeting criteria 2 will have assessments performed up to at least Week 24 and subsequent visit assessments will depend on the timing of the subject enrollment and approval of the protocol amendment at their local level.

The disposition of subjects in the double-blind period will be determined by visit based on time of discontinuation and randomization.

5.5 Efficacy variables

Efficacy endpoints of the study are summarized below:

Table 1 Efficacy endpoints

Efficacy endpoints	***Planned in the protocol***
Primary efficacy endpoint	
• Percent of baseline NT-pro-BNP assessed at Week 24.	*Yes*
Secondary efficacy endpoints	
• Change from baseline to Week 24 in the clinical summary score (as assessed by the Kansas City Cardiomyopathy Questionnaire [KCCQ]).	*Yes*
• Change from baseline to Week 24 in accelerometer-assessed proportion of time spent in light to vigorous physical activity based on a threshold of > 100 activity counts per minute.	*Yes*
• Time to worsening heart failure (WHF) event over 52 weeks (Up to EOS / OLE)	*Yes*
Other efficacy endpoints	
• Number of days alive and out of the hospital (DAOH) assessed over 52 weeks (up to EOS / OLE)	*Yes*
• Time to first occurrence of the composite endpoint, which is defined as either HF death or HF hospitalization over 52 weeks (up to EOS / OLE)	*Yes*
• Time to first occurrence of a composite of CV death or CV hospitalization over 52 weeks (up to EOS / OLE)	*Yes*
• Number of hospital admissions (for HF) over 52 weeks (up to EOS / OLE)	*Yes*
• NYHA FC (improved/worsened/stable) at each post-baseline assessment	*Yes*
• Clinical composite outcome measure ('worsened', 'unchanged', 'improved') based on NYHA class, patient global assessment and occurrence of death or HF hospitalization at each post-baseline assessment	*Yes*
• Change from baseline in the KCCQ overall summary score as well as clinical summary score and physical limitations score over time	*Yes*
• Percent of baseline NT-pro-BNP over time	*Yes*
• Change from baseline in accelerometer-assessed physical activity variables over time	*Yes*
• Proportion of time spent in physical activity at each time assessment	*Yes*
• Change from baseline in echocardiography left and right heart function at Weeks 24 and 52	*Yes*

Efficacy endpoints	***Planned in the protocol***
• Percent of baseline MR-proANP over time	*Yes*
• Patient global assessment (PGA)	*No*

CV = Cardiovascular; DAOH = Number of days alive and out of the hospital; EOS = End-of-study; FC = Functional Class; HF = Heart Failure; KCCQ = Kansas City Cardiomyopathy Questionnaire; MR-proANP = Mid-Regional pro-Atrial Natriuretic Peptide; NT-pro-BNP =n-terminal pro-Brain Natriuretic Peptide; NYHA = New York Heart Association; OLE = Open-label enrollment; PGA = Patient global assessment; WHF = worsening heart failure.

Prior to global protocol version 6 subjects were followed up to Week 52, except where subjects prematurely discontinue study. Under global protocol version 6 the follow-up period was reduced to 24 weeks, except where subjects had already passed Week 24 (Visit 11). Therefore, it is expected that the follow-up will be varied across subjects. All subjects efficacy assessment information available will be considered up to Week 52.

Except where otherwise specified, baseline [Section 11.3] for efficacy variables is defined as the last non-missing value observed among all measures collected during placebo and macitentan run-ins, up to and including the start of double-blind treatment date [Section 11.2.10]. Therefore, the baseline is expected to be the Visit 6 (Randomization Day 1) assessment unless missing or as specified otherwise for accelerometer data [Section 5.5.2.2] or as determined by the scheduled visit assessments (Screening for echocardiography, see Appendix A, protocol Table 1).

Efficacy data for NT-pro-BNP are provided via a central laboratory (Covance). The results of the NT-pro-BNP are considered to be potentially un-blinding. As per protocol Section 11.3, the clinical trial team are blinded to the NT-pro-BNP results from all visits post randomization Visit 6, until un-blinding of the study post-database lock.

All assessments recorded after start of double-blind treatment [Section 11.2.10] will be assigned to the most appropriate study visit according to the best fitting time-window for that assessment as detailed in Section 11.1. Assessments recorded before or on double-blind treatment start date (i.e., Visit 1/Screening, Visit 4/W4-R, Visit 6/W9-R/D1) will not be assigned with the same criteria, but evaluated according to the scheduled visit assigned.

5.5.1 Primary efficacy variable(s)

The primary efficacy variable is the percent of baseline NT-pro-BNP at Week 24. It is calculated as the ratio of the Week 24 NT-pro-BNP value over baseline value, expressed in percentage, i.e., as 100 times the Week 24 value divided by the baseline

value. Reduction in NT-pro-BNP (ratio < 100) is associated with a clinical improvement.

For subjects without an available NT-pro-BNP value at Week 24, the last available value observed before Week 24 (including baseline) will be carried forward and considered for the main analysis [see Section 10.6.2].

5.5.2 Secondary efficacy variables

5.5.2.1 Quality of life: KCCQ Clinical summary score at Week 24

The first key secondary efficacy variable is the clinical summary score as assessed by the Kansas City Cardiomyopathy Questionnaire (KCCQ), expressed as change from baseline to Week 24:

- KCCQ clinical summary score at Week 24 - Baseline KCCQ score.

Baseline score is the Visit 6 value (according to the schedule of assessment this is the first time KCCQ is assessed). Hence, if value at Visit 6 is missing then baseline is missing and change from baseline cannot be determined for this subject.

For subjects who died before Week 24, the missing score at Week 24 will be set to 0: otherwise, the last available value observed before Week 24 (including baseline) is carried forward and considered for the analysis.

The algorithm for the derivation of the KCCQ clinical summary scores is detailed in Appendix C.

5.5.2.2 Accelerometer-assessed light to vigorous physical activity at Week 24

The second key secondary efficacy variable is the proportion of time spent in light to vigorous physical activity (LVPA) based on a threshold of > 100 activity counts per minute, expressed as change from Baseline to Week 24.

Physical activity is assessed via accelerometer device. The accelerometer is given to the subject at Visit 2 (D1-R start of placebo run-in) and the subject will wear the device daily during waking hours until Visit 3 (W1-R). This period is considered a training period. Subsequent assessments via accelerometry are performed for 9 consecutive days following the visit [Week 1-R placebo; Week 5-R macitentan; Week 4; Week 16; Week 24] as well as for 9 consecutive days during waking hours prior to Visit 11 (Week 24) for subjects enrolled under global protocol version 6.

Baseline is defined as the non-missing, evaluable measures observed among data collected over 9 days following Visit 3 (during placebo run-in), because these data are obtained while the subject is still on placebo. Visit 3 data are preferred to data collected at Visit 2 as any potential changes in HF treatment are expected to occur at

the beginning of the placebo run-in, and so the clinical stability of subjects is expected to be better established at Visit 3 as compared to Visit 2. However, in case of missing non-evaluable data during placebo run-in, the training period following Visit 2 (during placebo run-in) may be used. In the rare event accelerometer data is not evaluable at Visit 2 and Visit 3 then accelerometer data collected during Visit 5 is used. The data obtained during the 9 days following Visit 5 is the least favored choice since the subject is already on single-blind macitentan during this period. However, using this data is considered acceptable for the Baseline definition, as the recently initiated single-blind macitentan treatment would have then only no or limited impact on the accelerometer parameters.

Daily time spent in LVPA is determined as the sum of time (minutes) in light, moderate or vigorous (including very vigorous) activity per day.

- Proportion of time in LVPA = sum of daily time in LVPA / sum weartime

Derivations are determined using only complete days. To be considered **evaluable**, physical activity should have been measured for at least 4 complete days (out of the 9 consecutive days) at a specific time point of assessment.

The 9-day time period is selected by taking the first (in chronological order) interval of 9 consecutive days with at least 4 complete daily time periods following the visit (or the last 9 consecutive days preceding Week 24 visit for assessment according to protocol version 6) and not including the midpoint of the day-interval between the visit and the following one.

A complete day is defined as a record of at least 7 hours of data, i.e., if "WearMinutes" >=7*60 (after excluding the periods when the device was apparently not worn). These limitations allow for obtaining reliable results [Robertson 2011].

Accelerometry data

Efficacy data for accelerometry are collected electronically and are transferred to Actelion by the vendor (Actigraph).

The data are processed and provided by ActiGraph. Data are provided as "Minute epochs" (60 seconds) and "Day epochs" (24 hours) including activity counts in 3 dimensions (X, Y and Z axis); vector magnitude (square root of the sum of the squares of X, Y and Z axis); and "wear time" (minutes) calculated using the Troiano 2008 approach [Troiano 2008]. The time (minutes) spent in sedentary, light, moderate and vigorous (including very vigorous) activity is provided for the "Day epoch" according to the algorithm of Freedson 1998 [Freedson 1998], excluding time where the device was not worn ("WearFilteredCutpoints"). The cutpoints are based on the activity counts on the vertical Y axis. In addition, ActiGraph provide the daily LVPA and daily activity accounts including non-wear time. Physical activity is evaluated

according to wear time unless specified otherwise. Actigraph also provide the SERENADE specific parameters for the Day epoch: (1) number of episodes of activity > 100 activity counts per minute of at least 1 minute duration and; (2) number of episodes of activity > 100 activity counts per minute of at least 5 minute duration. Derivations for physical activity (mean daily activity and proportions of time) are determined in the analyses using evaluable days at each visit.

[Note according to current literature and standard of care: for wear-time and valid day criteria, this would entail the subject wearing the device for a minimum of 4 days, and to have a minimum of 10 hours of wear time per day. The protocol listed valid day as 7 hours per day and will be applied for the analyses unless specified otherwise].

5.5.2.3 Time to first occurrence of worsening of heart failure event

The third key secondary efficacy variable is the time to first occurrence of WHF event.

Occurrence of a WHF event is assessed by the investigator using the 2015 ACC/AHA definition (see protocol section 7.2.2.3 for details).

A WHF event includes HF death, hospitalization for WHF or an urgent visit for WHF. WHF events determined by the investigator are reported in the CRF.

Time to first occurrence of WHF event is expressed in days and calculated as the onset date of the first WHF event minus date of randomization plus 1 or, for censored subjects, as EOS date minus date of randomization plus 1. The onset date of the first WHF event is the earliest between the following three dates:

- the date of death reported on the 'Death' form in the CRF where classification of primary cause of death is 'HF';
- the earliest date of WHF hospitalization entered on a 'WHF Event' CRF for a WHF event type of 'WHF Hospitalization';
- the earliest date of urgent WHF visit entered on a 'WHF Event' CRF for a WHF event type of 'Urgent WHF Visit'.

All worsening of heart failure events occurring until EOS / OLE are considered, irrespective of subjects' compliance to assigned therapies.

Subjects without any worsening of heart failure event up to EOS / OLE are right-censored at their EOS / OLE date [Section 11.2.12].

Prior to global protocol version 6 subjects were followed up to Week 52 (except where subjects prematurely discontinue study). Under global protocol version 6, subjects are followed up to at least Week 24. For the time to first occurrence of WHF

event, all events up to the EOS / OLE for each subject are considered, regardless of their duration of follow-up.

5.5.3 Other efficacy variables

5.5.3.1 Number of days alive and out of the hospital

The number of days alive and out of the hospital (DAOH) is calculated for each randomized subject based on the total expected follow-up time, (TFU) the number of days in hospital and the time elapsed between death date and expected follow-up time (number of days dead).

The **TFU** is determined as the number of days elapsed between the randomization date and the EOS (EOS - randomization date + 1) for the alive subjects or TFU=395 days (52 weeks × 7 days + 30 days safety follow-up + 1 day) for subjects who die during the study / OLE date. Under global protocol version 6 TFU=199 days (24 weeks × 7 days + 30 days safety follow-up + 1 day). For further details see Section 11.2.12.

The **number of days in hospital** (NDH) is obtained by adding the durations (discharge date - admission date + 1) of each individual hospital stay for which the admission date is on or after the randomization date. Subjects never hospitalized have 0 days in hospital.

The **number of days dead** (NDD) are calculated for subjects who die during the study as the number of days elapsed between their death date and what would have been their expected study duration if they did not die. It is calculated as:

$$NDD = 395 - (death\ date - randomization\ date + 1)$$

The NDH and NDD are then subtracted from the TFU to derive DAOH for each subject:

$$DAOH = TFU - NDH - NDD$$

The percentage of DAOH (%DAOH) is also calculated for each subject by dividing DAOH by the TFU.

5.5.3.2 Time to first occurrence of HF death or HF hospitalization

The variable of interest is the time to first occurrence of HF death or HF hospitalization.

Occurrence of HF death and hospitalization events are assessed by the investigator and recorded in the CRF ("Death" form and "Hospitalization" form respectively).

HF deaths are all deaths with a primary reason of 'HF' entered into the 'Death' CRF.

HF hospitalization is defined per appendix 6 of the protocol as:

- Subject is admitted to the hospital with a primary diagnosis of HF
- Length of stay is at least 24 h (or extends over a calendar date).

HF hospitalization are defined as the hospitalizations recorded on the 'Hospitalization' CRF where length of stay ≥ 1 day (end date - start date ≥ 1) **AND** primary diagnosis of HF determined as the following criterion selected in the CRF:

'HF' is specified where the main reason for hospitalization is cardiovascular = 'Yes'.

All HF hospitalizations and HF deaths with an admission / death date on or after randomization until EOS / OLE are considered, irrespective of subjects' compliance to assigned therapies.

Subjects without any HF hospitalization or HF death up to EOS / OLE are right-censored at their EOS / OLE date [Section 11.2.12].

Time to first occurrence of HF death or HF hospitalization is expressed in days and calculated as the onset date of the first HF death or HF hospitalization minus the date of randomization plus 1 or, for censored subjects, as the EOS / OLE date minus the date of randomization plus 1.

Prior to global protocol version 6 subjects were followed up to Week 52 (except where subjects prematurely discontinue study). Under global protocol version 6, subjects are followed up to at least Week 24. For the time to first occurrence of HF death or HF hospitalization, all events up to the EOS / OLE for each subject are considered, regardless of their duration of follow-up.

5.5.3.3 Time to first occurrence of CV death or CV hospitalization

The variable of interest is the time to first occurrence of cardiovascular (CV) death or CV hospitalization.

Occurrence of CV death and hospitalization events are assessed by the investigator and recorded in the CRF ("Death" form and "Hospitalization" form respectively).

CV deaths are all deaths with a primary reason of 'CV: Acute MI', 'CV: Sudden cardiac death', 'CV: HF', 'CV: Stroke', 'CV: Procedure', 'CV: Hemorrhage', or 'CV: Other' entered into the 'Death' CRF.

CV hospitalization is defined per appendix 6 of the protocol as:

- Subject is admitted to the hospital with a primary diagnosis of HF, MI, stroke, resuscitated sudden death, CV procedure, CV hemorrhage or cardiovascular hospitalization not included in the above categories but with specific, known cause.
- Length of stay is at least 24 h (or extends over a calendar date).

CV Hospitalization are the hospitalizations recorded on the 'Hospitalization' CRF where length of stay ≥ 1 day (end date - start date ≥ 1) **AND** the primary diagnosis CV determined as the question 'Was the main reason for hospitalization cardiovascular?' ticked 'Yes'.

All CV hospitalizations and CV deaths with an admission / death date on or after randomization until EOS date are considered, irrespective of subjects' compliance to assigned therapies.

Subjects without any CV hospitalization or CV death up to EOS / OLE are right-censored at their date of EOS / OLE date [Section 11.2.12].

Time to first occurrence of CV death or CV hospitalization is expressed in days and calculated as the onset date of the first CV death or CV hospitalization minus the date of randomization plus 1 or, for censored subjects, as the EOS / OLE date minus the date of randomization plus 1.

Prior to global protocol version 6 subjects were followed up to Week 52 (except where subjects prematurely discontinue study). Under global protocol version 6, subjects are followed up to at least Week 24. For the time to first occurrence of CV death or CV hospitalization, all events up to the EOS / OLE for each subject are considered, regardless of their duration of follow-up.

5.5.3.4 Number of HF hospital admissions over 52 weeks

The total number of recurrent HF hospitalizations, i.e., the cumulative number of distinct HF hospitalization episodes (different admission dates) as defined in Section 5.5.3.2, is computed for each subject, based on investigator assessment.

All hospitalizations occurring until EOS / OLE date are considered, irrespective of subjects' compliance to assigned therapies.

Subjects who do not experience any HF hospitalization before EOS / OLE date will be considered as having 0 HF hospitalization.

Prior to global protocol version 6 subjects were followed up to Week 52 (except where subjects prematurely discontinue study). Under global protocol version 6,

subjects are followed up to at least Week 24. For the analysis of the number of HF hospital admissions over 52 weeks all events up to the EOS / OLE for each subject are considered, regardless of their duration of follow-up.

5.5.3.5 NYHA FC (improved/worsened/stable) at each post-baseline assessment

Changes from baseline in NYHA FC values (I=1, II=2, III=3, IV=4) are categorized as improved, worsened or stable at every post-baseline assessment according to the following rules:

An improvement corresponds to a decrease in NYHA class by at least one level whereas a worsening corresponds to an increase in NYHA class by at least one level. Subjects remaining in the same NYHA class as the one reported at baseline are categorized as stable.

The proportion of subjects in each category is calculated at each post-baseline assessment based on the number of subjects with non-missing data (i.e., those having a reported value of I through IV).

Subjects who died (any cause) are classified as worsened (NYHA FC = V).

5.5.3.6 Clinical composite outcome ('worsened', 'unchanged', 'improved')

The clinical composite outcome measure provides an overall evaluation of whether a subject's condition is considered to have improved, worsened, or unchanged after a pre-specified period of time [Packer 2001].

The clinical composite outcome measure is derived based on NYHA FC, patient global assessment (PGA) and occurrence of death or HF hospitalization [see definition in Section 5.5.3.2] at each post-baseline assessment.

PGA is a seven-point patient self-evaluation scale rating compared to Visit 6, collected in the CRF pages 'Patient Global Assessment (PGA)'. Favorable changes are 'MARKEDLY IMPROVED', 'MODERATELY IMPROVED', 'SLIGHTLY IMPROVED' responses. 'REMAINED UNCHANGED' implies an unchanged condition. Other responses consist of an unfavorable change.

Improved, worsened, or unchanged conditions are defined as below:

- **Improved:** experienced a favorable change in NYHA class [see Section 5.5.3.5 for definition] or in the PGA from the baseline and did not experience death or HF hospitalization during the planned duration of treatment.

- **Worsened:** experienced death or HF hospitalization during the planned duration of treatment or reported worsening of their NYHA class or an unfavorable change in PGA compared to baseline.
- **Unchanged:** neither improved nor worsened compared to baseline, as defined above.

The clinical composite outcome will be calculated and described for the different following periods of time:

- Over 8 weeks (Day 2 up to Day 56 (+ 4 day window for Week 8 assessment)
- Over 16 weeks (Day 2 up to Day 112 (+ 8 day window for Week 16 assessment)
- Over 24 weeks (Day 2 up to Day 168 (+ 8 day window for Week 24 assessment)

For a given period, only the deaths and hospitalizations occurring during the studied period will be taken into account in the clinical composite outcome calculation. Similarly, the last non-missing NYHA and PGA assessments performed during the studied period will be considered. To allow for the assessment schedule window, + 4 days and + 8 days will be added to each period according to the visit schedule [see Appendix A, protocol Table 2].

5.5.3.7 Change from baseline in the KCCQ overall summary score and KCCQ scores over time

Changes from baseline are derived at each visit for the following KCCQ summary scores:

- Total score
- Clinical summary score
- Physical limitations score.

For subjects who died during the study, the scores will be set to 0 after the date of death.

The algorithm for the derivation of the summary scores is detailed in Appendix C.

5.5.3.8 Percent of baseline in NT-pro-BNP over time

The percent of baseline NT-pro-BNP is calculated at each post-baseline visit using the same calculation described for the primary efficacy variable [see Section 5.5.1].

No imputation for missing data will be performed.

5.5.3.9 Change from baseline in accelerometer-assessed physical activity variables over time

The accelerometer-assessed variables of interest are:

- Proportion of time spent in light to vigorous (including very vigorous) physical activity (LVPA) based on a threshold of > 100 activity counts per minute [see Section 5.5.2.2].
- Mean daily number of episodes of activity > 100 activity counts per minute of at least 1 minute duration*. The mean daily number of episodes is derived for each visit as the sum of episodes over complete days divided by number of complete days.
- Mean count per minute of daily activity, derived at each visit taking the total activity counts in the Y axis (Day epoch - "WearFilteredTotalAxisYCounts") for a complete day and divided by the total wear minutes in that day ("WearMinutes" variable in Data Transfer Specifications [DTS]) across all complete days (at least 4 days out of 9 days). In summary, to compute the mean counts per minute of daily activity, it is necessary to identify all complete days; if less than 4 complete days out of 9 are identified then the endpoint is missing, otherwise, if at least 4 complete days out of 9 are identified before a visit, then calculate the endpoint on the complete days only.
- Mean daily number of episodes of activity > 100 counts per minute of at least 5 minutes duration* is derived for each visit as the sum of episodes over complete days divided by number of complete days.
- Mean daily accelerometer units (DAU) is derived at each visit based on the total vector magnitude (Day epoch - variable "WearfilteredTotalVectorMagnitude" in DTS) and dividing by total wear minutes. Mean DAU is then derived for each visit for the complete days.

[*Actigraph provides the number of episodes for each day (Day epoch) that accelerometry is performed for 9 days following each visit].

Longitudinally collected over time and expressed as changes from baseline.

5.5.3.10 Proportion of time spent in physical activity at each time assessment

The proportion of time spent in each category of physical activity [Section 5.5.2.2] is calculated for each category and assessment time using evaluable days.

Proportion time = total time spent in the category / total wear time:

- Sedentary physical activity (< 100 activity counts per minute)
- Light physical activity (100 to ≤ 1951 activity counts per minute)
- Moderate physical activity (1952–5724 activity counts per minute)

– Vigorous (including very vigorous) physical activity (≥ 5725 activity counts per minute).

5.5.3.11 Change from baseline in echocardiography left and right heart function at Week 24 and Week 52.

Standard 2D/Doppler echocardiography is performed at screening (Visit 1), Week 24 (Visit 11/PTOP2) and Week 52 (Visit 13/EOT). The change is defined as post-baseline value minus baseline value.

Echocardiographies are read centrally and echocardiography parameters are provided by the echocardiography laboratory.

The variables of interest for the analyses include the following:

- Tricuspid annular plane systolic excursion (TAPSE) [mm]
- Left atrial volume (LAV) [mL]
- Septal e' velocity [cm/sec]
- Pulmonary artery systolic pressure (PASP) [mmHg]
- Lateral E/e' ratio
- Septal E/e' ratio

Note: other echocardiographic parameters provided by the echocardiography laboratory will be listed.

5.5.3.12 Percent of baseline in MR-proANP over time

The percent of baseline MR-proANP is calculated at each post-baseline visit using the same methodology described for the primary efficacy variable [see Section 5.5.1].

5.5.3.13 Patient Global Assessment

PGA is performed at Randomization (Visit 6), Week 8 (Visit 9), Week 16 (Visit 10), Week 24 (Visit 11/PTOP2), Week 36 (Visit 12/PTOP3) and Week 52 (Visit 13/EOT).

The assessment is based on a 7-point patient self-evaluation as detailed in Section 5.5.3.6.

5.6 Safety variables

For the safety and tolerability endpoints, the **main analyses** are based on comparison of the double-blind treatment groups, placebo versus macitentan for subjects included in the safety set [Section 7.1.6]. Longitudinal safety data collected for subjects

included in the safety set will be described over time from screening to the end of study.

As the safety set includes a pre-selected population of subjects tolerating the run-in periods, a few analyses are presented separately including subjects who do not enter the double-blind period, i.e., on PRI and MRI analysis sets [Sections 7.1.2 and 7.1.3].

For analyses related to change during the double-blind treatment period, the macitentan run-in baseline [Section 11.3.1] is used, i.e., the last non-missing assessment obtained before or on the day of the start of macitentan single-blind run-in treatment.

Subjects who discontinue treatment early or complete the treatment period as per protocol enter a 30 days safety follow-up period on the day after the last dose of study treatment. The **main analysis** of safety variables considers the time from first dose of double-blind treatment to end of double-blind treatment period plus 30 days [see definition in Section 11.5]. Selected analyses are defined separately for the single-blind placebo [Section 11.5.3] and macitentan run-in [Section 11.5.4] periods as well as for the combined macitentan treatment period [Section 11.5.5]. The periods are defined in Section 11.5 and are inclusive of start and end dates.

Safety endpoints are defined in section 6.2 of the protocol as:

Safety endpoints	***Planned in the protocol***
• All-cause death up to 30 days after study treatment discontinuation	*Yes*
• Number of all-cause hospital admissions up to 30 days after study treatment discontinuation	*Yes*
• Treatment-emergent AEs/ SAEs up to 30 days after study treatment discontinuation	*Yes*
• AEs leading to premature discontinuation of study treatment	*Yes*
• Change in vital signs (systolic and diastolic arterial blood pressure and pulse rate) and body weight up to 30 days after study treatment discontinuation	*Yes*
• Treatment-emergent marked laboratory abnormalities up to 30 days after study treatment discontinuation	*Yes*
• Change in laboratory parameters from baseline to all assessed time points during the study	*No*
• Change from baseline in eGFR up to 30 days after study treatment discontinuation	*Yes*
• Decrease from baseline in SBP of ≥ 5% and SBP < 100 mmHg up to 30 days after study treatment discontinuation.	*Yes*
• Treatment emergent ECG findings	*No*

5.6.1 Adverse events

An AE is defined as any event reported by the investigator in the "Adverse Event" CRF. All AEs that occur after signing of the ICF and up to 30 days after study treatment discontinuation are reported in the CRF "Adverse Event". For subjects entering the PTOP, serious adverse events (SAEs) are reported until the end of the PTOP period. The original terms used by the investigators to describe AEs are assigned PT for classification and tabulation using the latest implemented version of MedDRA.

5.6.1.1 Treatment-emergent adverse events

Treatment-emergent AEs (TEAEs) are defined as any AE that started during the study treatment period (i.e., from first dose of single-blind placebo treatment, see Section 11.5.1), or any worsening of AE (see below for definition) that started prior

to the study treatment period with an "Intensity change date" reported during the study treatment period.

Double-blind TEAEs are defined for the double-blind treatment period as any AE that started, or intensity worsened during the double-blind period. This will be used as the main period for presenting summaries for comparison of treatment groups.

Treatment-emergent AEs are also defined for the placebo run-in and macitentan run-in periods as well as combined macitentan treatment period (single-blind plus double-blind macitentan treatment period [see Section 11.5.5]). These latter periods are used for a few selected presentations of AE and SAE summaries.

A change in intensity of an AE is documented in the adverse event CRF together with the intensity change date. A worsening (✓) in intensity is determined according to the following table for intensity change in AE:

Previous intensity	**New intensity**			
	Mild	Moderate	Severe	Missing
Mild	×	✓	✓	✓
Moderate	×	×	✓	✓
Severe	×	×	×	×
Missing	×	✓	✓	✓

Double-blind TEAEs:

- any AE with an onset date during double-blind treatment period, start and end dates of period included [see definition in Section 11.5.2].
- any worsening of AE that started prior to the start of double-blind treatment with an "Intensity change date" reported during the double-blind treatment period (start and end dates included) [see definition in Section 11.5.2].

Placebo run-in TEAEs:

- any AE with an onset date between the start and end dates (included) of the placebo run-in period [see definition in Section 11.5.3].
- any worsening of AE that started prior to the placebo run-in start date with an "Intensity change date" reported between the start and end dates (included) of the placebo run-in period [see definition in Section 11.5.3].

Macitentan run-in TEAEs:

- any AE with an onset date between the start and end dates (included) of the macitentan run-in period [see definition in Section 11.5.4].
- any worsening of AE that started prior to the macitentan run-in start date with an "Intensity change date" reported between the start and end dates (included) of the macitentan run-in period [see definition in Section 11.5.4].

Combined macitentan TEAEs:

- any AE with an onset date between the start and end dates (included) of the combined macitentan treatment period [Section 11.5.5].
- any worsening AE that started prior to the combined macitentan treatment period start date with an "Intensity change date" reported between the start and end dates (included) of the combined macitentan treatment period [Section 11.5.5].

5.6.1.2 Frequency of treatment-emergent adverse events

TEAEs reported more than once (as qualified by the same PT) for a subject within a study period are counted only once in the frequency table for the given study period.

5.6.1.3 Intensity of treatment-emergent adverse events

For treatment-emergent AEs reported more than once (as qualified by the same PT) for a subject within a study period with different intensities, the worst intensity is considered for the given study period. In case of missing intensity, "*severe*" is imputed.

5.6.1.4 Relationship of treatment-emergent adverse events

Relationship to study treatment is defined as 'related' or 'not related'. An AE is considered related if the causality is checked as 'related' by the investigator. For treatment-emergent AEs reported more than once (as qualified by the same PT) for a subject within a study period, the worst relationship (i.e., 'related') is considered. AEs with missing relationship are considered in any analysis as 'related'.

5.6.2 Deaths

The date/time of death and associated primary cause are recorded in the CRF "Death" form.

The original terms used by the investigator to describe death (i.e., primary cause of death) are assigned PTs for classification and tabulation using latest implemented version of MedDRA.

Additional classification of primary cause of death is performed by the investigator based on pre-specified causes (CV vs non-CV) available in the same CRF.

Treatment-emergent deaths are assigned to periods (double-blind treatment, placebo and macitentan run-in, combined macitentan) in the same way as for TEAEs according to the date of death.

5.6.3 Serious adverse events

An AE is considered serious if the tick box 'Yes' for 'Serious?' is checked on the AE CRF. If the information on seriousness is missing, the AE is assumed to be a SAE for the purpose of the summaries.

Treatment-emergent SAEs are assigned to study treatment periods as for AEs [Section 5.6.1.1]. SAEs collected during PTOP are reported on the same CRF as other AEs / SAEs and handled the same way as SAEs collected during the double-blind treatment period.

5.6.4 Adverse events leading to discontinuation of study treatment

An AE is considered as leading to discontinuation of study treatment if the tick box 'Drug withdrawn' of 'Action taken with study treatment' is checked on the "Adverse Event" CRF.

5.6.5 Other significant adverse events

5.6.5.1 AEs with fatal outcome

An AE with fatal outcome is any AE with Outcome = "Fatal" in CRF.

TEAEs with fatal outcome are presented for subjects in the safety set [see Section 10.9.2.4].

5.6.5.2 AEs of special interest

AEs of special interest identified are given below. The definitions are based on a combination of MedDRA PTs and SMQs which can be updated according to MedDRA version / SMQs updates.

<u>Oedema and fluid retention:</u>

- Any AE with PTs listed in the SMQ "Haemodynamic oedema, effusions and fluid overload (SMQ)" OR
- Any AE with PTs containing "Pulmonary congestion"

- excluding events with PTs containing "site"

Anaemia:

- Any AE with a PT within the SMQ "Haematopoietic erythropenia" OR
- Any AE with a PT within the SMQ "Haematopoietic cytopenias affecting more than one type of blood cell " (excluding two unspecific PTs: "blood disorder", "blood count abnormal") OR
- an event with any PT containing the text "anaemia".

Hepatic AEs of special interest (HAESI):

- Any AE with PT, system organ class (SOC) or HLGT containing the text specified in Appendix D.

The latest version of the MedDRA dictionary will be applied and hence any PTs that change or are no longer applicable will be documented in the DPS.

5.6.6 Vital signs and body weight

Vital signs (systolic and diastolic blood pressures [SBP and DBP], pulse rate) and body weight are measured pre-dose at all visits and reported in the "Vital Signs" CRF. Triplicate SBP and DBP and radial pulse measurements are measured in a supine or sitting position. The average of the triplicate values [provided in SDTM] are used for the analysis.

At Visit 4/W4-R, SBP, DBP and pulse rate are also monitored 4 hours post-macitentan first dose, and 8-hours post-macitentan first dose. These assessments are listed and not included in any summaries. At this visit, the SpO2 (%) is also monitored pre-dose, 4-hours post macitentan first-dose, 8-hours post-macitentan first dose and at additional time points during visit 4, if needed by the investigator. These assessments are only listed. Note these first-dose safety monitoring assessments (4 and 8-hour post-macitentan first-dose) are no longer applicable for subjects enrolled under protocol version 5 or later.

All assessments recorded during double-blind treatment period [Section 11.5.2] excluding Visit 6/W9-R/D1 will be assigned to the most appropriate study visit according to the best fitting time-window for that assessment as detailed in Section 11. If more than one value falls on the same date and time (and parameter) then the one with the last sequential number in SDTM will be used. Assessments recorded before or on double-blind treatment start date (i.e., Visit 1/Screening, Visit 2/D1-R, Visit 3/W1-R, Visit 4/W4-R, Visit 6/W9-R/D1) will not be assigned with the same criteria, but reported according to scheduled visit.

Absolute change from macitentan run-in baseline [Section 11.3.1] to each visit in SBP, DBP, pulse rate, and body weight defined as:

(Value at visit (Visit 6 - Visit 13)) - (value at macitentan run-in baseline)

is evaluated.

5.6.6.1 Decrease in SBP

The decrease from baseline during double blind-period in SBP ≥ 5% and SBP < 100 mmHg (both conditions must occur at the same time) at any time point of assessment starting from one day after start of double-blind study treatment up to EOT + 30 days will be derived.

Similarly, decrease from baseline during macitentan run-in period is determined from one day after single-blind macitentan run-in start date [see definition in Section 11.2.8] and up to start date of double-blind treatment (inclusive). If double blind treatment was not received, then the period is up to 30 days after the end of the single-blind macitentan run-in date [see definition in Section 11.2.9].

5.6.7 Electrocardiogram

A standard 12-lead ECG is performed at Visit 1/Screening, Visit 4/W4-R, Visit 6/W9-R/D 1, Visit 8/W4, Visit 10/W16, Visit 11/W24 and Visit 13/W52-EOT.

All assessments recorded after start of double-blind treatment [Section 11.2.10] will be assigned to the most appropriate study visit according to the best fitting time-window for that assessment as detailed in Section 11.1. Assessments recorded before or on double-blind treatment start date (i.e., Visit 1/Screening, Visit 4/W4-R, Visit 6/W9-R/D1) will not be assigned with the same criteria, but evaluated according to the scheduled visit assigned.

The ECGs are interpreted locally, and the presence of 'atrial fibrillation' or 'atrial flutter' or 'sinus rhythm' are recorded in the CRF ("12-Lead ECG" form). Unscheduled assessments may be performed at any time during the study if appropriate based on the investigator judgment.

The following variable is also defined:

Treatment-emergent qualitative ECG abnormalities (findings): defined as presence (yes/ no) of any finding (atrial fibrillation or atrial flutter or sinus rhythm) during the double-blind treatment period [see Section 11.5.2]. Defined as any finding and separately for each type.

5.6.8 Laboratory

Hematology and chemistry tests are performed at Visit 1/Screening, Visit 4/W4-R, Visit 6/W9-R/D1, Visit 8/W4, Visit 9/W8, Visit 10/W16, Visit 11/W24, Visit 12/W36, Visit 13/W52-EOT, Visit 14/EOS, Visit PTOP1. In between Visit 9/W8 and Visit 12/W36, monthly aspartate transaminase / alanine transaminase (AST/ALT) monitoring is recommended. Unscheduled assessments may be performed at any time during the study if appropriate based on the investigator judgment.

Data are evaluated in SI units unless specified otherwise as provided by the central laboratory. In case of local laboratory, values are provided in conventional units and converted to SI units. The tests converted to SI are available in SDTM for the analysis.

Note: Laboratory results recorded as '< xxx' (or similar with ≤, >, or ≥) will be considered as 'xxx' for analysis, e.g., a result for ALT of '< 1' is considered as 1 for the analysis. The values are listed including the < or > sign.

For the analysis, local laboratory values are included in the calculation of marked laboratory abnormalities (MLAs) [Section 5.6.8.2]. All assessments recorded after double-blind treatment start date [Section 11.2.10] up to EOT + 30 days will be assigned to the most appropriate study visit according to the best fitting time-window for that assessment as detailed in Section 11.1. 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 date and time (and laboratory) then the one with the last sequential number in SDTM will be used.

Laboratory parameters include:

Hematology:

- Hemoglobin (SI Unit: g/L; Conventional unit: g/dL)
- Hematocrit (SI Unit: L/L; Conventional unit: %)
- Erythrocyte count (reticulocyte count) (SI Unit: 10^{12}/L; Conventional unit: 10^{6}/µL)
- Leukocytes count (SI Unit: 10^{9}/L; Conventional unit: 10^{3}/µL)
- Neutrophils count (SI Unit: 10^{9}/L; Conventional unit: 10^{3}/µL)
- Lymphocytes count (SI Unit: 10^{9}/L; Conventional unit: 10^{3}/µL)
- Monocytes count (SI Unit: 10^{9}/L; Conventional unit: 10^{3}/µL)
- Eosinophils count (SI Unit: 10^{9}/L; Conventional unit: 10^{3}/µL)
- Basophils count (SI Unit: 10^{9}/L; Conventional unit: 10^{3}/µL)
- Platelet count (SI Unit: 10^{9}/L; Conventional unit: 10^{3}/µL)

Chemistry:

- ALT (U/L)
- AST (U/L)
- Alkaline phosphatase (U/L)
- Total and direct bilirubin (SI unit: µmol/L; Conventional unit: mg/dL)
- Creatinine (SI unit: µmol/L; Conventional unit: mg/dL)
- Blood urea nitrogen (SI unit: mmol/L; Conventional unit: mg/dL)
- Uric acid (SI unit: µmol/L; Conventional unit: mg/dL)
- Sodium, potassium, chloride, calcium, magnesium (mmol/L)
- Total protein, albumin (SI unit: g/L; conventional unit: g/L)
- Albumin / Globulins ratio
- eGFR (mL/min/1.73m^2) using the Modification of Diet in Renal Disease formula

The absolute observed value and change from baseline to each visit for **selected laboratory parameters** (i.e., erythrocytes, hemoglobin, hematocrit, AST, ALT, total bilirubin and alkaline phosphatase) are defined as:

Values at visit (Visit 8 to Visit 13) - (value at macitentan run-in baseline).

5.6.8.1 Fasted state laboratory values

Laboratory values at screening / re-screening are collected only in fasted state. The value at the latest successful (re-) screening are used:

- Glucose (SI unit: mmol/L; Conventional unit: mg/dL)
- Thyroid stimulating hormone (TSH; SI unit: mIU/L; Conventional unit: µIU/mL)
- T3, T4
- Lipid profile (SI unit: µmol/L; Conventional unit: mg/dL):
  - low-density lipoprotein (LDL) cholesterol
  - high-density lipoprotein (HDL) cholesterol
  - very low-density lipoprotein (VLDL) cholesterol
  - total cholesterol
  - triglycerides.

5.6.8.2 Marked laboratory abnormalities

The following MLA are derived according to the protocol.

Double-blind MLAs are all marked laboratory abnormalities occurring from one day after the double-blind study treatment start [see definition in Section 11.2.10] and up to 30 days after the end of double-blind treatment date [see definition in Section 11.2.11].

Double-blind treatment-emergent MLAs are double-blind MLAs that were not present at macitentan run-in baseline.

Macitentan run-in MLAs are all marked laboratory abnormalities occurring from one day after the single-blind macitentan run-in start date [see definition in Section 11.2.8] and up to start date of double-blind treatment (inclusive). If double-blind treatment was not received, then period is up to 30 days after the end of single-blind macitentan run-in date [see definition in Section 11.2.9].

Macitentan run-in treatment emergent MLAs are macitentan run-in MLAs that were not present at macitentan run-in baseline [see Section 11.3.1].

Placebo run-in MLAs are all marked laboratory abnormalities occurring from one day after the single-blind placebo run-in start date [see definition in Section 11.2.6] and up to start date of single-blind macitentan treatment (inclusive). If single-blind macitentan treatment was not received, then period is up to 30 days after the end of single-blind placebo run-in date [see definition in Section 11.2.7].

Placebo run-in treatment emergent MLAs are placebo run-in MLAs that were not present at Screening.

Table 2 Definition of marked laboratory abnormalities

Parameter	LL marked	LLL marked	HH marked	HHH marked
Hematology				
Hemoglobin *(baseline value within normal range or below LLN)*	< 100 g/L	< 80 g/L	> 20 g/L above ULN	> 40 g/L above ULN
Hemoglobin *(baseline value > ULN)*	< 100 g/L	< 80 g/L	> 20 g/L above baseline	> 40 g/L above baseline
Hematocrit	< 0.28 L/L for females < 0.32 L/L for males	< 0.20 L/L	> 0.55 L/L for females > 0.60 L/L for males	> 0.65 L/L

Parameter	LL marked	LLL marked	HH marked	HHH marked
Leukocytes	< 3.0 × 10^9/L	< 2.0 × 10^9/L	> 20.0 × 10^9/L	> 100.0 × 10^9/L
Neutrophils	< 1.5 × 10^9/L	< 1.0 × 10^9/L	NA	NA
Lymphocytes	< 0.8 × 10^9/L	< 0.5 × 10^9/L	> 4.0 × 10^9/L	> 20.0 × 10^9/L
Eosinophils	NA	NA	> 5x10^9/L	NA
Platelets	< 75 × 10^9/L	< 50 × 10^9/L	> 600 × 10^9/L	> 999 × 10^9/L
Chemistry				
ALT*	NA	NA	> 3 × ULN	> 5 × ULN
AST*	NA	NA	> 3 × ULN	> 5 × ULN
Alkaline phosphatase	NA	NA	> 2.5 × ULN	> 5 × ULN
Total bilirubin	NA	NA	> 2 × ULN	> 5 × ULN
Creatinine	NA	NA	> 1.5 × ULN	> 3 × ULN
(baseline value within normal range or below LLN)				
Creatinine *(baseline value > ULN)*	NA	NA	> 1.5 × above baseline	> 3 × above baseline
BUN	NA	NA	> 2.5 × ULN	> 5 × ULN
Uric acid	NA	NA	> 590 µmol/L	> 720 µmol/L
Sodium	NA	< 130 mmol/L	> 150 mmol/L	> 155 mmol/L
Potassium	< 3.2 mmol/L	< 3.0 mmol/L	> 5.5 mmol/L	> 6.0 mmol/L

Parameter	LL marked	LLL marked	HH marked	HHH marked
Magnesium	< 0.5 mmol/L	< 0.4 mmol/L	NA	> 1.23 mmol/L
Calcium	< 2.0 mmol/L	< 1.75 mmol/L	> 2.9 mmol/L	> 3.1 mmol/L
Albumin	< 30 g/L	< 20 g/L	NA	NA

* For ALT and AST, additional threshold (HHHH) are reported namely > 8 × ULN

ALT = alanine aminotransferase; AST = aspartate aminotransferase; BUN = blood urea nitrogen; LLN = lower limit of the normal range; NA = not applicable; ULN = upper limit of the normal range

5.6.8.3 Additional liver laboratory abnormalities

The following liver test abnormalities (LTAs) are defined for MRI and double-blind periods:

- (ALT > 3 × ULN) **OR** (AST > 3 × ULN)
- (ALT > 5 × ULN) **OR** (AST > 5 × ULN)
- (ALT > 8 × ULN) **OR** (AST > 8 × ULN)
- Total Bilirubin > 2 × ULN
- {(ALT > 3 × ULN) **OR** (AST > 3 × ULN)} **AND** (Total Bilirubin > 2 × ULN at any time).

Double-blind LTAs and macitentan run-in LTAs are derived in the same way as for the double-blind MLAs and macitentan run-in MLAs respectively [Section 5.6.8.2].

The categories "ALT or AST > 3 × ULN", "ALT or AST > 5 × ULN", "ALT or AST > 8 × ULN" are not mutually exclusive and subjects may be counted in more than one category. The highest ALT or AST value at any time point in the period is considered in the evaluation of these categories, as defined above.

5.6.9 Other safety variables

5.6.9.1 Number of all-cause hospital admissions up to 30 days after study treatment discontinuation

All-cause hospitalizations are reported in the dedicated CRF "Hospitalization".

The total number of events for each subject (e.g., all-cause hospitalizations) is calculated for the run-in periods [see definition in Sections 11.5.3 and 11.5.4] and for the double-blind treatment period [see definition in Section 11.5.2].

5.6.9.2 Change from baseline in Glomerular Filtration Rate up to 30 days after study treatment discontinuation

The eGFR (mL/min/1.73m^2) using the Modification of Diet in Renal Disease formula is assessed at all visits with laboratory test [see Section 5.6.8].

Unscheduled visits may be performed at any time during the study if appropriate based on the investigator judgment.

As described in laboratory Section 5.6.8, the assessments will be assigned to the most appropriate study visit.

Data are evaluated as provided by the central laboratory.

The change in eGFR from macitentan run-in baseline to each visit is defined as:

Value at visit (Visit 6/ Visit 8/ Visit 13) - (value at macitentan run-in baseline).

5.6.9.3 Physical examination

Physical examination is performed at all scheduled visits, collected in the CRF "Physical Examination" with general assessment performed. If an abnormality is found (i.e., Result = "Abnormal"), further details describing the signs and symptoms related to the abnormality are specified. The abnormality is clinically significant if the question "Clinically significant" is answered "Yes".

Physical examination includes the examination of the general appearance, head, eyes, ears, nose, throat, neck (including thyroid), cardiovascular, respiratory, gastrointestinal, lymphatic, genitourinary, rectal, extremities, neurologic, psychiatric, skin, musculoskeletal.

5.6.9.4 Heart failure sign and symptoms of special interest

The following signs and symptoms of special interest are assessed at every visit as part of the physical examination (CRF "Signs/Symptoms of Special Interest"):

- Peripheral edema (defined as increased tissue fluid indicated by perceptible pitting indentation on lower leg, foot, or sacrum after palpation)
- Pulmonary rales/crackles/crepitations
- Abdominal distention or ascites (in the absence of primary hepatic disease)
- S3 gallop
- Orthopnea
- Nocturia
- Paroxysmal nocturnal dyspnea
- Nocturnal cough
- Jugular venous distention
- Hepatojugular reflux.

The above are collected in CRF as 'Present at visit' = 'Yes /No /Not assessed'.

5.7 Sub-study

Sub-study endpoints defined in this section are applicable only to subjects who were enrolled under protocol version 5 and eligible for the sub-study. Data of the sub-study assessments (6-MWT and Borg Dyspnea Index) will be described presenting listings and descriptive summaries (only in the case of at least 20 subjects with available baseline / post baseline assessments).

A training 6-MWT was conducted at Visit 2 and a qualifying 6-MWT conducted at Visit 3. To obtain meaningful results for the 6-MWT, additional eligibility criteria were defined as listed below. Subjects who met these eligibility criteria were included in the 6-MWT sub-study and 6-MWT was performed at subsequent visits (Visit 4/ W4-R, Visit 6/W9-R/D1, Visit 8/W4, Visit 9/W8, Visit 10/W16, Visit 11/W24 and Visit 13/W52-EOT). Subjects who were included in the 6-MWT sub-study were allowed to withdraw (e.g., if a subject becomes wheelchair bound) from participation in the sub-study while remaining in the main study.

Eligibility criteria for participation in the 6-MWT sub-study were:

- 6-minute walk distance (6-MWD) $\geq$100 m AND $\leq$ 450 m at Visit 3
- Ambulatory (not wheelchair / scooter dependent)

The efficacy variable is defined as follows:

- The observed and absolute change (meters) from baseline over time in 6-MWD, as measured by the 6-MWT

Where baseline is defined as the last non-missing value observed among all measures collected during placebo run-in and macitentan run-in, up to and including the start of double-blind treatment date [Section 11.2.10]. Therefore, the baseline is expected to be the Visit 6 (Randomization Day 1) assessment. All assessments, including unscheduled if any, will be assigned to the most appropriate visit time point according to the best fitting time window for the assessment [see Section 11.1, Table 7].

The Borg Dyspnea Index is assessed after each 6-MWT rating dyspnea on a scale from 0 to 10.

5.8 Quality of life variables

The KCCQ is a validated health related quality of life measure for heart failure. The KCCQ is a 23-item, self-administered instrument that quantifies physical function, symptoms (frequency, severity and recent change), social function, self-efficacy and knowledge, and quality of life [see appendix 3 of the protocol for questionnaire]. It is assessed at Visit 6/W9-R/D1, Visit 9/W8, Visit 10/W16, Visit 11/W24, Visit 12/W36, Visit 13/W52 and PTOP2/PTOP3/PTOP4.

All assessments recorded during double-blind treatment period [Section 11.5.2] excluding Visit 6/W9-R/D1 will be assigned to the most appropriate study visit according to the best fitting time-window for that assessment as detailed in Section 11.1.

For subjects who died during the study after randomization, the scores will be set to 0 for the visits after the time-window that includes death date. For the visit whose time-window include death date, the scores will be set to 0 only in case there is no KCCQ assessment included in its time-window.

Questions refer to subjects' heart failure and how it may affect their life with reference to a period of "over the past 2 weeks".

In the KCCQ, different scores can be derived, in which higher scores reflect better health status. KCCQ scoring instructions are provided in Appendix C.

The quality of life variables of interest derived from KCCQ that will be summarized are:

- Physical limitations score
- Clinical summary score
- Overall summary score.

6 DEFINITION OF PROTOCOL DEVIATIONS

The description of each protocol deviation (PD) is agreed in the sponsor protocol deviation code list. According to this document, each PD is classified as important or not and categorized into periods as follows:

- PDs at screening and during run-in (for subjects who entered run-in only)
- PDs during double-blind treatment period and follow-up (including PTOP)
- PDs applicable to all study periods.

The data are available in SDTM.

Protocol deviation PD_MM.329 'Administration of incorrect study treatment (i.e., subject did not get the medication bottle assigned by IXRS)' will be classified to 'Incorrect bottle received: treatment received different from assigned treatment' (PD_MM.3291) or 'Incorrect bottle received: assigned treatment received' (PD_MM.3292).

The PDs which lead to exclusion from the Per-Protocol analysis set (PPS) are in Section 7.1.5.

7 ANALYSIS SETS

7.1 Definitions of analysis sets

7.1.1 Screened Analysis Set

The SCR analysis set includes all subjects who are screened and have a subject number.

7.1.2 Placebo run-in Set

The placebo run-in (PRI) Set includes all screened subjects who enter the placebo run-in and receive at least one dose of study treatment in single-blind placebo run-in.

7.1.3 Macitentan run-in Set

The macitentan run-in (MRI) Set includes all subjects who have completed the placebo run-in, enter the macitentan run-in and receive at least one dose of study treatment in single-blind macitentan run-in.

7.1.4 Full analysis Set

The FAS includes all subjects randomized to double-blind study treatment [see definition in Section 11.8]. 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 taken into account.

7.1.5 Per-Protocol analysis Set

The PPS comprises all subjects from the FAS who received double-blind study treatment and who complied with the protocol sufficiently to allow reliable assessment of the treatment effect on the primary efficacy endpoint. Subjects are evaluated according to the study treatment they have been assigned to.

Criteria for exclusion from PPS include:

Insufficient compliance to study drug in the first 24 Weeks determined from PD_PP.223 (Study treatment compliance during the core phase of the double-blind treatment period duration < 80% or >120%)

Insufficient exposure to study drug in the first 24 Weeks determined from 1 or more of the following:

- PD_PM.225 (Interruption of double-blind study treatment for more than 4 consecutive weeks without permanent study treatment discontinuation up to and including Week 24 (core phase)).
- Study drug discontinued (or interrupted) more than 7 consecutive days prior to Week 24 assessment of NT-pro-BNP.

Non-protocol compliant assessment of NT-pro-BNP at baseline or Week 24: determined from one of more of the following:

- PD_PM.215 (NT-pro-BNP sample not collected or not evaluable at Visit 11).
- PD_PM.152 [(At least 1 visit with missing or non-evaluable NT-pro-BNP sample during Screening (Visit 1) and/or Run-in (Visit 4, Visit 6)]. Where PD is applied to all 3 Visits.

Presence of any protocol deviation questioning the diagnosis of the studied disease and other PDS affecting the assessment of the primary efficacy are presented in Table 3.

Table 3 Exclusions from Per-Protocol analysis Set

PD Code	Description / Condition	Additional
PD_PP.101	No informed consent signed and dated by subject	
PD_PP.109	NYHA class I or IV or missing at Visit 1	
PD_PP.111	Subjects without structural heart disease (left atrial enlargement or Left ventricular hypertrophy) assessed by echocardiography	
PD_PP.112	No NT-pro-BNP / BNP value fulfilling the protocol criteria within 3 months prior to Screening or at Screening	
PD_PP.113	Subject does not meet the criteria for pulmonary vascular disease or right ventricular dysfunction	
PD_PM.116	Any prior measurement of LVEF < 40%	
PD_PP.117	Significant unrepaired structural valvular heart disease	
PD_PP.118	Hypertrophic, restrictive, or infiltrative cardiomyopathy	
PD_PP.119	Pericardial disease	
PD_PP.120	Acute coronary syndrome or unstable coronary artery disease or has undergone coronary artery bypass graft (CABG) or percutaneous coronary intervention (PCI) within 3 months of screening.	
PD_PP.121	Known indication for PCI or CABG	
PD_PP.124	History of or anticipated heart transplant or anticipated/implanted ventricular assist device	
PD_PP.125	Transient ischemic attack (TIA) or stroke within 3 months of Screening	
PD_PP.126	Systolic blood pressure (SBP) ≥180 mmHg or diastolic blood pressure (DBP) ≥110 mmHg	If present in all 3 periods: Screening, placebo run-in, macitentan run-in
PD_PM.127	Significant parenchymal lung disease	
PD_PP.129	Clinically significant congenital heart disease that could be the cause of the patient's symptoms and signs of heart failure.	
PD_PP.142	Any known factor or disease that might interfere with treatment compliance, study conduct, or interpretation of the results, such as drug or alcohol dependence or psychiatric disease.	
PD_PP.143	Known concomitant life-threatening disease with a life expectancy < 12 months.	
PD_PM.160	Subject not on stable Heart Failure/cardiovascular therapy during a minimum of 7 consecutive days immediately prior to randomization (Visit 6)	
PD_PM.201	Subject met at least one run-in failure criterion but was randomized	
PD_PP.221	Not on at least one oral diuretic (any type) at time of randomization (visit 6)	

PD Code	Description / Condition	Additional
PD_PM.310	Subject received a PAH specific therapy (i.e., ERA, PDE-5 inhibitors, prostanoids, IP receptor agonists, guanylate cyclase stimulators) between Screening and EOS	Except therapies starting after Week 24 NT-pro-BNP
PD_PM.314	Subject received investigational product other than study treatment between Screening and EOS	Except therapies starting after Week 24 NT-pro-BNP
PD_MM.329	Administration of incorrect study treatment (i.e., subject did not get the medication bottle assigned by IxRS). Plus, additional condition that treatment received is different from assigned treatment.	Except study treatment dispensed after Week 24 NT-pro-BNP.

BNP = Brain Natriuretic Peptide; CABG = coronary artery bypass graft; DBP = diastolic blood pressure; EOS = End-of-study; ERA = endothelin receptor antagonist; FC = Functional Class; IxRS = Interactive voice/web Recognition System; LVEF = Left Ventricular Ejection Fraction; NT-pro-BNP = n-terminal pro-Brain Natriuretic Peptide; NYHA = New York Heart Association; OLE = Open-label enrollment; PAH = pulmonary arterial hypertension; PCI = percutaneous coronary intervention; PDE-5 = phosphodiesterase-5; SBP Systolic blood pressure; TIA = Transient ischemic attack.

7.1.6 Safety Set

The Safety Set (SAF) includes all subjects from FAS who received at least one dose of double-blind study treatment. Subjects are evaluated according to the study treatment received. The treatment received may be different from the treatment assigned at randomization (randomized treatment) only in the case of a dispensing error that was sustained throughout the subject`s entire double-blind treatment period. Short-term dispensing errors will not qualify for a change from the randomized treatment group.

7.1.7 Sub-study Analysis Set

Applicable only to subjects who were enrolled under protocol version 5.

The Sub-study Analysis Set (SSAS) includes all subjects randomized to double-blind study treatment and eligible to enter the Sub-study [see Section 5.7]. In order to the intention-to-treat principle, subjects are evaluated according to the study treatment they have been assigned to (which may be different from the study treatment they have received).

7.2 Usage of the analysis sets

Table 4 Overview of the different analysis sets and their usage

Analyses/Data Displays	Screened analysis set (SCR)	Placebo Run-in set (PRI)	Macitentan Run-in set (MRI)	Full analysis set (FAS)	Safety set (SAF)	Per protocol analysis set[a] (PPS)	Sub-study Analysis set (SSAS)
Disposition	✓	✓	✓	✓			
Inclusion /exclusion criteria	✓						
Demographic characteristics				✓		✓	
Baseline characteristics				✓		✓	
Medical history				✓			
Previous /concomitant medications				✓			
Treatment exposure		✓	✓	✓	✓	✓	
Primary endpoint				✓		✓	
Subgroup Analyses				✓			
Secondary endpoints				✓		✓*	
Exploratory endpoints				✓			
Safety endpoints		✓	✓		✓		
Sub-study							✓
Subject listings	✓**	✓		✓			

a Main analysis of the primary and secondary efficacy endpoints is based on the full analysis set. Demographics, baseline characteristics and treatment exposure are presented for the PPS only when the PSS differs from the FAS by more than 15 subjects in both randomized treatment groups (≥ 10%).

* selected main analyses for physical activity and quality of life.

** select listings including screening failure information

The FAS is used for the analyses of all the efficacy variables as well as for the description of the study population at baseline.

The PPS is used to perform specific sensitivity analysis on the primary efficacy variable.

The SAF is used for the analyses of the safety variables.

The SCR set is used for the description of subject disposition and some select listings. Unless specified otherwise, individual listings are prepared on the PRI or FAS.

The PRI set is used to present the reasons leading to discontinuation of placebo run-in as well as a few select safety analyses.

The MRI set is used to document the reasons leading to discontinuation of macitentan run-in as well as a few select safety analyses.

The SSAS is used for presenting listings and descriptive summaries of the sub-study data (6-MWT and Borg Dyspnea Index) assessed for subjects included in the sub-study (applicable only to subjects enrolled under global protocol version 5).

8 DEFINITION OF SUBGROUPS

The category in **bold type** is the reference category used in the statistical analysis of these subgroups.

The following subgroups are defined according to the protocol.

- Geographical region (**Americas**, Western Europe, Eastern Europe as defined in Section 5.2.1]
- Sex (Male versus **Female**)
- NYHA FC (II versus **III**) at screening as defined in Section 5.2.2
- Atrial Fibrillation (yes versus **no**) as collected on Specific MH CRF
- NT-pro-BNP **(< 1000** pg/mL versus ≥ 1000 pg/mL) as per the IxRS stratification
- Age (**18–64**, 65–84**,** ≥ 85 years)
- Renal function at Baseline: **Normal**, mild decrease, moderate decrease, severe decrease / ESRD
- Obesity (BMI ≥ 30 vs **< 30**).

9 GENERAL STATISTICAL METHODOLOGY

SAS (Statistical Analysis System®) version 9.4 is used for all the statistical analysis.

This section describes in general terms the statistical analysis methods which are applied in the subsequent sections. Specific details will be addressed in Section 10.

In addition, all efficacy data will be summarized for observed and change from baseline using descriptive statistics by visit time point where applicable.

9.1 Statistical methods for ANCOVA analysis

The statistical method used to analyze the primary efficacy variable takes into account the expected underlying log normal distribution of NT-pro-BNP [Solomon 2012]. As a consequence, the treatment effect will be tested on a log-scale using an ANCOVA and expressed, on the original scale, as the ratio between treatment groups of the geometric means of individual percent of baseline at Week 24 (macitentan over placebo).

The model is adjusting for the log of the ratio: baseline NT-pro-BNP over NT-pro-BNP value at subject's entry into macitentan run-in (as per Interactive Response Technology [IRT]), with the aim to better control the potential unbalance between treatment groups of the NT-pro-BNP levels observed before entering the macitentan run-in. Such unbalance may not be fully captured by the baseline NT-pro-BNP value, which is expected to be the value observed the day of randomization for the majority of subjects, and thus already influenced by macitentan. Despite the randomization stratification based on the NT-pro-BNP value collected at subject's entry into macitentan run-in, a continuous covariate quantifying individual NT-pro-BNP evolution during the macitentan run-in is deemed more sensitive and therefore more appropriate than a binary covariate based uniquely on the stratification factor.

Due to the underlying log-normal distribution of the primary efficacy variable, an analysis of covariance will be applied on log-transformed data.

y_{ij} denotes the log-transformed percent of baseline in NT-pro-BNP at Week 24 for subject j in treatment group i

x_{ij} denotes the log-transformed ratio: baseline NT-pro-BNP over NT-pro-BNP value at subject's entry into macitentan run-in, for subject j in treatment group i

TRT_i denotes for each subject the treatment group i he/she belongs to. Equal 1 for macitentan and 0 for placebo

α denotes the common intercept

e_{ij} denotes the random error term, assumed to follow a normal distribution with mean 0 and standard error σ

The model could be written:

$$y_{ij} = \alpha + \beta_1 TRT_i + \beta_2 x_{ij} + e_{ij}$$

2-sided hypotheses are focusing on the model coefficient β_1.

The null hypothesis is that macitentan and placebo effects on NT-pro-BNP are the same. The alternative hypothesis is that the effect of macitentan on NT-pro-BNP differs from placebo effect.

H_0: $\beta_1 = 0$

vs

H_1: $\beta_1 \neq 0$.

The null hypothesis will be tested by a 2-sided Wald test with a 2-sided significance level of 0.10.

For primary endpoint, the statistical model will be implemented by the SAS® code detailed in the DPS Part 1 document.

ANCOVA is also applied (under assumptions of normal distribution) for secondary and others efficacy analysis specified below:

- Change from baseline to Week 24 in the clinical summary score (as assessed by the KCCQ). Adjusting for the variable baseline value and for the log of the ratio (baseline NT-pro-BNP over NT-pro-BNP value at subject's entry into macitentan run-in (as per IRT).
- Change from baseline to Week 24 in accelerometer-assessed proportion of time spent in light to vigorous (including very vigorous) physical activity (LVPA) based on a threshold of > 100 activity counts per minute. Adjusting for the variable baseline value and for the log of the ratio (baseline NT-pro-BNP over NT-pro-BNP value at subject's entry into macitentan run-in (as per IRT)).
- Number of days alive and out of the hospital (DAOH and %DAOH) assessed over 52 weeks. Adjusting need to be for the log of the ratio only (baseline NT-pro-BNP over NT-pro-BNP value at subject's entry into macitentan run-in [as per IRT]).

The validity of classical parametric ANCOVA depends on several assumptions, including normality of error terms, equality of error variances for different treatments, equality of slopes for the different treatment regression lines and linearity of regression. For the primary and secondary endpoints analyzed with an ANCOVA model the assumptions will be assessed (creating diagnostic plots) and where they will not be valid, a non-parametric ANCOVA similar to that described in section 7.7 of Stokes [Stokes 2000] will be performed.

Details on the procedure described in Stokes [Stokes 2000] are given in the DPS Part 1 document.

9.2 Statistical methods for time-to-event data

The analysis of time to first occurrence of WHF will be carried out using a Cox proportional hazards model adjusting for the log of the ratio: baseline NT-pro-BNP over NT-pro-BNP value per IRT.

Data will be listed and summarized in tables and figures, including: number of events, number of censored observations, number of subjects at risk, and Kaplan-Meier (KM) estimates of the survival function for time-to-event variables, hazard ratios and corresponding 90% confidence intervals (CIs). The null hypothesis, tested by means of the log-rank test comparing macitentan 10 mg versus placebo, is rejected upon achieving a statistically significant difference at a 2-sided significance level such that the log-rank p-value is lower than $\alpha = 0.10$.

The graphical representation follows the recommendations from Pocock [Pocock 2002]. Two-sided CIs at specific time points will be constructed, with confidence limits (CLs) calculated using Greenwood's formula for the estimate of the standard error. Median time to event (as well as 25th and 75th percentiles) for each group will be provided with the corresponding 2-sided CIs calculated using the method of Brookmeyer [Brookmeyer 1982].

KM estimates of the survival function are presented at the following relevant time points: Weeks 16, 24, 36, and 52.

The statistical models will be implemented by the SAS® code described in the DPS Part 1 document.

This statistical method with the same adjustment will be also applied on other efficacy variables:

- Time to first occurrence of HF death or HF hospitalization over 52 weeks (up to EOS / OLE);
- Time to first occurrence of CV death or CV hospitalization over 52 weeks (up to EOS / OLE).

An investigation into the assumption of proportional hazards for treatment will be performed informally using a plot of the complementary log-log of the survival against the log of time (for each group). If the hazards are proportional, the lines should be approximately parallel.

9.3 Statistical methods for count variables

The number of recurrent HF hospital admissions will be analyzed using a Negative binomial regression model adjusting for the log of the ratio: baseline NT-pro-BNP over NT-pro-BNP value per IRT. The time on treatment will also be considered and included as an offset variable.

Adjusted Estimates within each treatment group and between treatment groups will be displayed along with 2-sided 90% CIs, after inverse transformation using the exponential function.

The statistical models will be implemented by the SAS® code detailed in the DPS Part 1 document.

9.4 Statistical methods for categorical variables

The proportion of subjects having NYHA FC and clinical composite outcome improved, having worsened or being stable and the proportion of subjects in each of the 3 PGA categories of favorable change, unfavorable change and unchanged [Section 5.5.3.6] will be calculated at each post-baseline assessment by treatment group. The relative risk (macitentan over placebo) of subjects having worsened will be displayed with 2-sided 90% CIs.

The statistical models will be implemented by the SAS® code detailed in the DPS Paret 1 document.

NYHA FC: to calculate the relative risk, “stable” and “improved” categories will be put together as a single category “not worsened”.

Clinical composite outcome: to calculate the relative risk, “unchanged” and “improved” categories will be put together as a single category “not worsened”.

PGA: to calculate the relative risk, “unchanged” and “favorable change” categories will be put together as a single category “not worsened”.

9.5 ANCOVA model for repeated measurements

The efficacy endpoints where this model will be implemented are: Quality of life scores over time, Percent of baseline in NT-pro-BNP over time, change from baseline in accelerometer-assessed physical activity over time, change from baseline in Glomerular Filtration Rate.

A random coefficient regression model (with random slopes and intercepts) will be applied. It includes treatment, time (expressed in weeks from the start of double-blind treatment) and time-by-treatment as factors and the baseline value as a covariate. For NT-pro-BNP analysis only, the model includes treatment, time and time-by-treatment

as factors and the log of the ratio: baseline NT-pro-BNP over NT-pro-BNP value per IRT as a covariate.

The choice of the random coefficient regression model instead of the traditional repeated measures analysis of variance (fixed coefficient model assuming that everybody changes at the same rate) was justified by the advantages of using a method allowing for individual differences in the initial status and rate of change (slope) and allowing the analysis of unbalanced data (subjects are not required to have the same number of observations or to be observed at the same time points).

For NT-pro-BNP analysis, the treatment effect expressed as geometric means ratio and its associated 90% 2-sided CIs will be estimated based on the same model by inversely transforming using the exponential function the least squares means and 90% CIs obtained in log scale.

No imputation of missing values is planned.

The statistical models will be implemented by the SAS® code detailed in the DPS Part 1 document.

10 STATISTICAL ANALYSES

10.1 Overall testing strategy

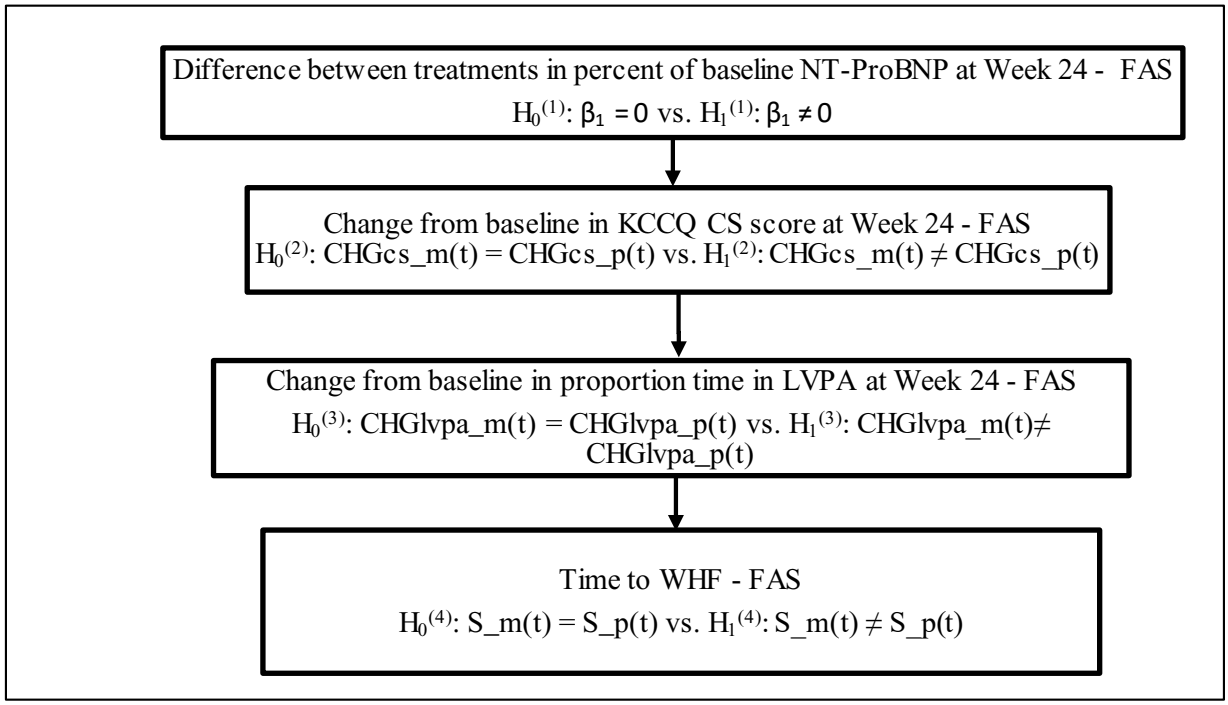
The effect of macitentan as compared to placebo will be tested using a mixed-models ANCOVA adjusting for the NT-pro-BNP evolution observed during the macitentan run-in.

The resulting p-value will be compared to an overall 2-sided type I error of 10%.

The primary and three secondary efficacy variables will be tested at 2-sided $\alpha = 0.10$, using a hierarchical testing approach to address multiplicity concern in the FAS. The secondary efficacy variables will be tested in a sequential manner following the order listed in Section 5.5.2.

The hierarchical strategy will be applied to control the experiment-wise α level [Dmitrienko 2010] with overall 2-sided $\alpha = 0.10$. Where a test null hypothesis cannot be rejected in favor of macitentan 10 mg, then the test p value and subsequent test p values in the hierarchical order will be presented for information only.

Figure 4 Hierarchical testing strategy



KCCQ CS = Kansas City Cardiomyopathy Questionnaire Clinical Summary score; CHGcs(t) = Change from baseline in KCCQ Clinical Summary score; CHGlvpa(t) = Change from baseline in proportion of time spent in light to vigorous physical activity based on a threshold of > 100 activity counts per minute;WHF = worsening heart failure; S(t) = WHF survival function.

10.2 General rules for data presentations

This section describes the general rules applied for all data displays.

Unless otherwise specified in this document:

- All listings will be sorted by randomized treatment group (or not-randomized), then by geographical region (order according to Section 5.2.1), country (alphabetical order), site, subject number and when appropriate by visit / date of assessment. The order site/ subjects within the not-randomized subset will be according to the latest run-in period achieved: MRI, PRI and screening failures last. All data collected will be displayed, including unscheduled visits (if any). Listings will be presented on the SCR set unless specified otherwise.
- In summary tables and graphical representations, subjects will be grouped by randomized treatment group. All scheduled visits for a variable will be displayed including assessments during run-in periods.

- In summary tables presented for the FAS or SAF, "macitentan" treatment group is displayed as first column from the left. "Placebo" is the second/last column from the left. A third column (i.e., "Total") with the pool across the treatment groups (macitentan+placebo) is added for subject disposition, demographic and baseline characteristics.
- The absolute change from baseline to Visit/Week X is defined as the difference between the post-baseline Visit X value and the baseline value.
- Selected summary tables presented for run-in treatment periods: placebo run-in is displayed first followed by macitentan run-in to reflect the order of the run-periods.
- Data will be summarized in tables including: number of non-missing observations, frequency with percentage per category (percentages based on the number of non-missing observations) for categorical safety variables,
- Data will be summarized in tables including: number of missing observations and frequency with percentage per category (percentages based on the total number of observations) for categorical variables other than safety variables.
- Number of events, number of censored observations, number of subjects at risk, and KM estimates of the survival function, hazard ratios and corresponding 90% CIs will be displayed for time-to-event variables.
- Geometric means and coefficients of variation will also be displayed for NT-pro-BNP analysis.

10.3 Display of subject disposition, protocol deviations and analysis sets

Subject disposition, PDs and analysis sets will be displayed in accordance with the ICH E3 guidelines [ICH 1995] and the sponsor CSR template.

10.3.1 Screening failures and run-in disposition and failures

Counts and percentages of screening failures [see definition in Section 5.1.1] and associated reason will be presented in a summary table. For subjects that fail screening more than once, only the reason of the last failure will be reported in the summary table. All screening failure reasons will be included in the listings.

Counts and percentages of single-blind placebo and macitentan run-in withdrawals [see definitions in Sections 5.1.2 and 5.1.3] will be presented overall and by reason. A subject listing of withdrawals and associated reasons according to period is provided by geographical region, country and site.

The number and percentage of subjects with 1 or more criterion for placebo run-in failure [Section 5.1.2] or macitentan run-in failure [Section 5.1.3] will be presented overall and per criterion. For subjects who failed placebo run-in and then entered and failed macitentan run-in, the summary will include only their first failure during the placebo run-in. Listings will include all run-in failure criteria met, sorted by geographical region, country, site and subject.

The number and percentage of subjects with 1 or more unmet eligibility criterion for placebo run-in [see definition in Section 5.1.1.1], macitentan run-in [see definition in Section 5.1.2.1] and double-blind treatment period [see definition in Section 5.1.3.1] will be presented by period overall and per criterion.

All subjects with unmet eligibility criteria will be listed by geographical region, country and site. Subjects randomized with unmet criteria will be flagged in the listing.

10.3.2 Subject disposition

The number of subjects randomized, treated (received at least one dose of double-blind study treatment), completed treatment, entered in PTOP, completed the study will be all summarized by treatment group and overall for the FAS.

Subjects prematurely withdrawn from the study (post randomization) with reasons for premature study withdrawal will be summarized by treatment group and overall for the FAS.

The following information is summarized, presenting the number and percent of subjects overall for the SCR set:

- Screened
- Entered the single-blind placebo run-in
- Treated with single-blind placebo
- Entered the single-blind macitentan run-in
- Treated with single-blind macitentan
- Eligible for the double blind
- Randomized
- Received double-blind treatment
- Completed treatment
- Completed study
- Eligible for the sub-study
- Participated in the sub-study

All subject disposition variables will be listed.

In addition, the disposition of subjects in the double-blind period will be summarized by visit on the FAS.

Subjects' randomization is summarized by geographical region, country and site.

Randomization information (number, date, stratification factor, planned and actual treatment group) will be provided in a subject listing for the FAS. The listing will also include the NT-pro-BNP observed value used in the randomization stratification factor and also at entry into MRI period (with flags for different values).

If any randomization code is broken before database closure, the event that triggered the request as well as the action taken following the unblinding are collected in the CRF "Subject Unblinding" and will be provided in a subject listing.

10.3.3 Protocol deviations

Protocol deviations [as defined in Section 6] will be summarized by category, displaying counts and percentages of subjects with at least one PD by treatment group and overall for the FAS. A similar table is presented for important protocol deviations.

This summary table is sorted, within each category, by overall frequency, in descending order; if a tie occurs, the tied characteristics will be sorted alphabetically.

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

10.3.4 Analysis sets

The number and percentage of subjects in each analysis set will be summarized in a table, by treatment group and overall on the SCR set. Subject participation in each analysis set will be listed.

Reasons for excluding subjects from the PPS will be listed and summarized in a table on the FAS.

10.4 Analyses of subject characteristics

This section provides the statistical analyses performed to describe subject characteristics defined in the corresponding section above.

Summaries for demographic and baseline characteristics will be performed on the FAS.

All data will be reported in subject listings.

10.4.1 Demographics

Continuous and categorical demographics [see definition in Section 5.2.1] will be summarized, by treatment group and overall, using descriptive statistics. Demographic characteristics are also summarized by subgroups of geographical region as defined in Section 8.

Demographics will also be summarized for subjects in each of the PRI and MRI analysis sets split by those who enter / don't qualify for the next period.

10.4.2 Baseline disease characteristics

Continuous and categorical baseline disease characteristics [see definition in Section 5.2.2] will be summarized, by treatment group and overall, using descriptive statistics on the FAS. NT-pro-BNP is partitioned by subjects with / without atrial fibrillation according to ECG at screening.

Baseline disease characteristics will also be summarized by subgroups of geographical regions as defined in Section 8.

Baseline disease characteristics are also summarized for subjects in each of the PRI and MRI analysis sets split by those who enter / don't qualify for the next period.

In addition, NT-pro-BNP values as per IxRS are displayed together with values assessed at subject's entry into macitentan run-in and displayed in a shift table if any mis-strata occurred.

All the baseline data collected in the "Specific Medical History" CRF and "Signs/Symptoms of special interest" CRF will be provided in a subject listing.

10.4.3 Other baseline characteristics

Other baseline characteristics (RHC and PFTs in Section 5.2.3) will be listed only.

Renal function [Section 5.2.3] based on the eGFR (mL/min/1.73m^2) is summarized as both continuous and categorical.

10.4.4 Medical history

All medical history [see definition in Sections 5.2.4 and 5.2.4.1] will be summarized together by treatment group and overall for the FAS, 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 will be presented by SOC and PT within each SOC as well as by PT. The summary tables will be 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). If the frequencies of SOC/PTs are the same, alphabetical order will be used.

Subjects with two or more occurrences of the same disease (as qualified by the same PT) are counted only once.

Separate listings of medical history and specific medical history [Section 5.2.4.1] are provided.

10.4.5 Previous and concomitant therapies

Number and percentages of subjects having taken at least one therapy will be presented by ATC class (level 4) and PT within each ATC class. All summaries will be tabulated by ATC class, and individual PTs within each ATC class, for the FAS. ATC classes will be sorted by descending order of frequency in the macitentan treatment group. If the frequencies of ATC class are the same, alphabetical order will be used. The same rule applies for PTs within ATC class.

Previous therapies [Section 5.2.5.1], study-treatment concomitant therapies [Section 5.2.5.4], and post-study treatment therapies [Section 5.2.5.5] will be summarized separately, by treatment group and overall, using descriptive statistics for categorical data.

All therapies will be reported in a subject listing for all subjects flagging previous, study-treatment concomitant and post-study treatment therapies accordingly.

10.4.6 Baseline therapies of special interest

Baseline therapies of special interest [Sections 5.2.5.2 and 5.2.5.6] will be summarized separately, by treatment group and overall, using descriptive statistics for categorical data as above in Section 10.4.5.

All therapies of special interest will be reported separately in a subject listing flagging baseline therapies accordingly.

10.4.7 Changes in Diuretics

Details for calculating TDD and deriving changes in diuretics intake are described in Section 5.2.5.7. The number and percentage of subjects with a change (increase/ decrease) in diuretic therapy: during placebo run-in period will be summarized on PRI set; during macitentan run-in period will be summarized on MRI set, while those with a change during double-blind treatment period will be summarized on FAS.

Summaries of increase in diuretic therapy will be presented overall and by category (increase TDD, new diuretic, IV) [Section 5.2.5.7].

Diuretics [Section 5.2.5.7] will be listed in a separate listing including description of changes (increases and decreases) when applicable.

10.4.8 Other subject characteristics

10.4.8.1 Contraceptive methods

For female of childbearing potential only, contraceptive methods [as defined in Section 5.2.6] together with serum pregnancy tests data (i.e., choriogonadotropin beta) will be provided in separate subject listings.

10.5 Analysis of study treatment exposure and compliance

Exposure data as well as drug accountability will be provided in subject listings.

10.5.1 Exposure

The duration of study treatment (including interruptions, see definition in Section 5.3.1) as well as the study treatment exposure (excluding interruptions) and the cumulative exposure (patient years) for the double-blind treatment period is summarized per treatment group using descriptive statistics for the FAS and SAF. The study treatment duration is also summarized as a categorical variable, presenting the cumulative distribution of treatment duration by class interval (i.e., at least 4 weeks, at least 8 weeks, at least 12 weeks, at least 16 weeks, at least 20 weeks, at least 24 weeks, at least 36 weeks and at least 52 weeks) and displaying counts and percentages of subjects in each class interval.

Study treatment exposure (excluding interruptions) up to Week 24 is summarized per treatment group using descriptive statistics for the FAS and PPS.

The total exposure to macitentan treatment (see definition in Section 5.3.1) is summarized using descriptive statistics for the MRI set.

A listing of exposure and study treatment interruptions is provided for the PRI.

10.5.2 Compliance with study treatment

Double-blind study-treatment compliance [see definition in Section 5.3.2] is summarized overall per treatment group on the FAS, displaying counts and percentages of subjects with compliance < 80%, 80%–120% and > 120%. In the same manner, also the double-blind study-treatment compliance up to Week 24 is

summarized per treatment group on the FAS and PPS. Compliance during macitentan run-in period is summarized on the MRI set.

A listing of compliance for each visit and compliance during macitentan run-in is also provided, containing the compliance values directly collected in "Study Drug Dispensing & Accountability" form in CRF.

10.5.3 Study treatment discontinuation

A summary table is provided on the SAF to display, for each treatment group, and overall, counts and percentages of subjects who prematurely discontinued study treatment during double-blind treatment period together with reason associated with study treatment discontinuation [see definition in Section 5.3.3].

A subject listing with all study treatment premature discontinuations and related reasons is provided.

A summary table to display counts and percentages of subjects who prematurely discontinued study treatment during the two run-in treatment periods is provided on the PRI set and MRI set respectively, together with reasons for discontinuation.

10.5.4 Study treatment adjustments or interruptions

Subjects with study treatment interruption and corresponding reasons for interruption(s) [defined in Section 5.3.4] will be summarized per treatment group for the double-blind study treatment period using the number and percentages of subjects with at least one interruption on the FAS. The numbers in the categories may sum to a number greater than the number of subjects with an interruption as a single subject could have multiple interruptions for multiple reasons.

Multiple interruptions of the same type (reason) are only counted once per subject.

All subject treatment interruptions are presented in the listing of exposure described in Section 10.5.1 above.

10.6 Analysis of the primary efficacy variable(s)

Percent of baseline in NT-pro-BNP at Week 24 [defined in Section 5.5.1] will be analyzed using ANCOVA as described in Section 9.1. NT-pro-BNP is assumed to follow a log-normal distribution. Therefore, data is log transformed before analysis.

The following covariate will also be used for the statistical analyses of the primary endpoint:

- log of the ratio between baseline NT-pro-BNP and the NT-pro-BNP value as per IRT stratification.

The main analysis will be performed on subjects of the FAS, according to the intent-to-treat principle. The PPS analysis will be used to perform specific sensitivity analysis on the primary efficacy variable.

10.6.1 Hypothesis and statistical model

Due to the underlying log-normal distribution of the primary efficacy variable, an analysis of covariance will be applied on log-transformed data.

$$y_{ij} = \alpha + \beta_1 TRT_i + \beta_2 x_{ij} + e_{ij}$$

The null hypothesis is that macitentan and placebo effects on NT-pro-BNP are the same. The alternative hypothesis is that the effect of macitentan on NT-pro-BNP differs from placebo effect.

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

H_1: $M_{\text{macitentan 10 mg}} \neq M_{\text{placebo}}$

The null hypothesis will be tested by a 2-sided Wald test with a 2-sided significance level of 0.10.

Assuming normal distribution of the primary variable, the null hypothesis will be tested by geometrical means of an ANCOVA model on percent of baseline NT-pro-BNP assessed at Week 24.

The ANCOVA model as described in Section 9.1 will be applied where y_{ij} denotes log-transformed percent of baseline in NT-pro-BNP at Week 24 and x_{ij} denotes the log-transformed ratio for subject j under treatment i, respectively.

Log-transformed percent of baseline in NT-pro-BNP at Week 24 will be used to test the null hypothesis of equality of treatment effects. According to the model notation, null and alternative hypotheses above conform to H_0: $\beta_1 = 0$ versus H_1: $\beta_1 \neq 0$.

10.6.2 Handling of missing data

The primary efficacy variable is the percent of baseline NT-pro-BNP at Week 24 (Visit 11). Baseline is defined as the last non missing value among all measures collected during placebo and macitentan run-ins, up to and including the day of double-blind treatment start (Visit 6). Therefore, no subject will have a missing value for baseline because the subject cannot be randomized via IRT in case of missing values at both pre-macitentan run-in (Visit 4) & screening (Visit 1).

During the conduct of the study, all efforts will be made to avoid missing values in NT-pro-BNP at Week 24. For Week 24, the time window interval [see Section 11.1] is applied.

Furthermore, if a subject prematurely discontinues study treatment the subject will be asked to return for a visit within 7 days of EOT where NT-pro-BNP assessments is performed. If treatment is discontinued prior to Week 24, and the subject enters the PTOP, it is requested that the subject return for a NT-pro-BNP assessment at Week 24. Any unscheduled assessment or local laboratory assessment will also be mapped to a time window.

However, if missing values still occur in NT-pro-BNP at Week 24 the following approach for imputation will be applied:

Imputation for Main Analysis

For randomized subjects without available NT-pro-BNP value at Week 24, the last available value observed before Week 24 (including baseline) is carried forward and considered for the main analysis.

This approach is chosen as it is commonly used for this specific parameter [Solomon 2012] and in order to facilitate the clinical interpretation across study publications in the same indication.

Imputation for Sensitivity analysis

A sensitivity analysis will be performed by **imputing** missing NT-pro-BNP values at Week 24 **according to the reason for drop-out**. In particular, randomized subjects who died or were hospitalized for HF before Week 24 will be assigned a value significantly higher than the one assigned to subjects who have a missing value at Week 24 due to technical/administrative reason. In details the following rules will be applied:

i) **Subjects who died or were hospitalized for HF before Week 24**: for each subject, their highest available value (worst case) observed from all visits from Visit 6 onwards (up to Week 24) will be imputed.

ii) **Subjects who are known to be alive, were not hospitalized for HF and with a missing value at Week 24**: the mean of their available values from Visit 6 onwards (up to Week 24) will be imputed.

iii) **Subjects Lost to follow-up:** their highest available value (worst case) observed from Visit 6 onwards (up to Week 24) will be imputed.

10.6.3 Main analysis

The main analysis will be performed on subjects in the FAS, according to the intent-to-treat principle.

The main analysis of the primary efficacy variable will be carried out after log-transformation using an ANCOVA adjusting for the log of the ratio: baseline NT-pro-BNP over NT-pro-BNP value as per IRT with an overall type I error of 10% 2-sided.

The treatment effect expressed as geometric means ratio and its associated 90% 2-sided CI will be estimated based on the same model by inversely transforming using the exponential function the least squares mean and 90% CIs obtained in log scale.

A listing of NT-pro-BNP values overtime will be provided for the FAS.

10.6.4 Supportive/sensitivity analyses

The following supportive analyses are planned:

- Main analysis on the PPS,
- Analysis of variance (ANOVA) on the FAS without adjustment for the NT-pro-BNP evolution observed during the macitentan run-in,
- Main analysis imputing missing NT-pro-BNP values at Week 24 according to the reason for drop-out.

10.6.5 Subgroup analyses

Subgroup analyses will be performed with a separate analysis for each subgroup variable [see Section 8] using a two factor ANOVA model with treatment, subgroup and treatment-by-subgroup interaction terms.

The test of interaction is performed with 'n-1' degrees of freedom, where 'n' is the number of subgroup categories. The treatment effect, measured as a geometric means ratio, will be estimated within each level of the subgroup variable based on the ANOVA including the interaction term. Significance of the interaction terms will be tested at a 0.01.

Treatment effect geometric means ratio and corresponding 90% CIs for the different levels of each subgroup will be presented in a forest-plot. The forest-plot will be prepared as described in Cuzick [Cuzick 2005] with a vertical reference line displayed at the level of the overall treatment effect for macitentan versus placebo and a vertical reference line at geometric means ratio = 1. The p-value for the interaction test is displayed on the plot for each subgroup along with the number of patients in macitentan and the number of patients in placebo within each subgroup level.

Some of the pre-specified subgroup analyses will not be conducted or categories will be combined when it is determined that there is insufficient number of subjects in a subgroup category to produce a valid result.

Further subgroups will be added for analysis if deemed appropriate.

No multiplicity adjustment is introduced; the subgroup analysis is descriptive in nature.

10.6.6 Additional supportive analyses

- NT-pro-BNP vs MR-proANP

A scatter plot of NT-pro-BNP observed values versus MR-proANP concentrations at Week 24 will be presented including a linear regression fitted line and 90% CLs for the fitted line. If a linear model does not fit the data, then alternative models may be explored (including a log-transformation of both variables).

- NT-pro-BNP at Week 24 (as observed) vs increase in diuretics

NT-pro-BNP is summarized partitioned by those with increase in diuretics (Yes/ No) and by type of increase new drug, increase TDD and IV.

10.7 Analysis of the secondary efficacy variables

The analysis of secondary efficacy variables will be performed on subjects in the FAS.

10.7.1 Quality of life: KCCQ Clinical Summary score at Week 24

10.7.1.1 Hypothesis and statistical model

The change from baseline to Week 24 in the clinical summary score will be analyzed by means of an ANCOVA adjusting for the score baseline value and for the log of the ratio: baseline NT-pro-BNP over NT-pro-BNP value per IRT.

The statistical hypotheses for this secondary endpoint is similar to statistical hypotheses of the primary endpoint.

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

H_1: $M_{\text{macitentan 10 mg}} \neq M_{\text{placebo}}$

Least Squares Estimates within each treatment group and between treatment groups will be displayed along with 2-sided 90% CIs and p-values.

More details regarding the model and SAS code are presented in the DPS Part 1 document.

10.7.1.2 Handling of missing data

The last available value observed before Week 24 (including baseline) is carried forward and considered for the analysis.

10.7.1.3 Statistical analysis

An ANCOVA on the change from baseline to Week 24 in the clinical summary score, adjusting for the score baseline value and for the log of the ratio: baseline NT-pro-BNP over NT-pro-BNP value per IRT will be applied on the FAS. The treatment effect and its associated 90% 2-sided CI will be estimated based on the model.

An individual subject listing of the KCCQ scores overtime will be provided on the FAS.

10.7.1.4 Supportive/sensitivity analyses

Not applicable.

10.7.1.5 Subgroup analyses

Not applicable.

10.7.2 Accelerometer-assessed light to vigorous physical activity at Week 24

10.7.2.1 Hypothesis and statistical model

The change from baseline to Week 24 in the proportion of time spent in light to vigorous (including very vigorous) physical activity will be analyzed by means of an ANCOVA adjusting for the variable baseline value and for the log of the ratio: baseline NT-pro-BNP over NT-pro-BNP value per IRT.

The statistical hypotheses for this secondary endpoint is similar to statistical hypotheses of the primary endpoint.

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

H_1: $M_{\text{macitentan 10 mg}} \neq M_{\text{placebo}}$

Least Squares Estimates within each treatment group and between treatment groups will be displayed along with 90% 2-sided CIs and p-values.

More details regarding the model and SAS code are presented in the DPS Part 1 document.

10.7.2.2 Handling of missing data

To be considered evaluable, physical activity should have been measured for at least 4 complete days at a specific time point of assessment. A complete day is defined as a record of at least 7 hours of data (after excluding the periods when the device was apparently not worn).

Subjects received a reminder from the site staff (e.g., telephone call, text message) to ensure that they were wearing the accelerometer during interest periods.

If the baseline value measured after Visit 3 is missing or not evaluable then the data collected following Visit 2 (training visit) or Visit 5 may be used [see Section 5.5.2.2]. No imputation of missing values is planned.

10.7.2.3 Statistical analysis

An ANCOVA on the change from baseline to Week 24 in the proportion of time spent in light to vigorous (including very vigorous) physical activity, adjusting for the variable baseline value and for the log of the ratio: baseline NT-pro-BNP over NT-pro-BNP value per IRT will be applied on the FAS. The treatment effect and its associated 90% 2-sided CI will be estimated based on the model.

10.7.2.4 Supportive/sensitivity analyses

Not applicable.

10.7.2.5 Subgroup analyses

Not applicable.

10.7.3 Time to first occurrence of worsening heart failure over 52 weeks (up to EOS / OLE)

10.7.3.1 Hypothesis and statistical model

The analysis of time to first occurrence of WHF will be carried out using a proportional hazards Cox model adjusting for the log of the ratio: baseline NT-pro-BNP over NT-pro-BNP value per IRT.

Subjects without any worsening of heart failure event up to EOS will be right-censored at their time of EOS. Estimate of Hazard Ratio and its associated 90% CI and p-value will be displayed.

KM estimates will be calculated with 2-sided 90% CIs at time points (16, 24, 36, and 52 weeks) for each treatment group and displayed in both a graphical (where the number of subjects at risk is at least 10% of the total number of subjects in the analysis

set) and a tabular form. In addition, the number of subjects at risk, the number of subjects censored and the number of subjects with event will be computed at each time point and for each treatment group.

Time to first occurrence of WHF will be presented in a subject listing on the FAS.

10.8 Analysis of other efficacy variables

Exploratory analysis of other efficacy variables will be conducted on the FAS. Missing values will not be imputed.

10.8.1 Number of days alive and out of the hospital

For both DAOH and %DAOH [see Section 5.5.3.1], the treatment effect will be assessed by means of an ANCOVA adjusting for the log of the ratio: baseline NT-pro-BNP over NT-pro-BNP value per IRT.

Least Squares Estimates within each treatment group and between treatment groups will be displayed along with 90% 2-sided CIs.

More details regarding the model and SAS code are presented in the DPS Part 1 document.

Data regarding DAOH and %DAOH will be reported in an individual subject listing on the FAS.

10.8.2 Time to first occurrence of HF death or HF hospitalization over 52 weeks (up to EOS/ OLE)

The analysis of time to first occurrence of HF death or HF hospitalization will be carried out using a proportional hazards Cox model adjusting for the log of the ratio: baseline NT-pro-BNP over NT-pro-BNP value per IRT.

Estimate of Hazard Ratio and its associated 90% CI will be displayed.

KM estimates will be calculated with 2-sided 90% CIs at time points (12, 24, 36, and 52 weeks) for each treatment group and displayed in both a graphical (where the number of subjects at risk is at least 10% of the total number of subjects in the analysis set) and a tabular form. In addition, the number of subjects at risk, the number of subjects censored and the number of subjects with event will be computed at each time point and for each treatment group.

More details regarding the model and SAS code are presented in the DPS Part 1 document.

Time to first occurrence of HF death or hospitalization will be presented in a subject listing on the FAS.

10.8.3 Time to first occurrence of CV death or CV hospitalization over 52 weeks (up to EOS/ OLE)

The analysis of time to first occurrence of CV death or CV hospitalization will be carried out using a proportional hazards Cox model adjusting for the log of the ratio: baseline NT-pro-BNP over NT-pro-BNP value per IRT.

Estimate of Hazard Ratio and its associated 90% CI will be displayed.

KM estimates will be calculated with 2-sided 90% CIs at time points (12, 24, 36, and 52 weeks) for each treatment group and displayed in both a graphical (where the number of subjects at risk is at least 10% of the total number of subjects in the analysis set) and a tabular form. In addition, the number of subjects at risk, the number of subjects censored and the number of subjects with event will be computed at each time point and for each treatment group.

More details regarding the model and SAS code are presented in the DPS Part 1 document.

Time to first occurrence of CV death or hospitalization will be presented in a subject listing on the FAS.

10.8.4 Number of HF hospital admissions over 52 weeks

The number of recurrent HF hospital admissions will be analyzed using a Negative binomial regression model adjusting for the log of the ratio: baseline NT-pro-BNP over NT-pro-BNP value per IRT.

Adjusted Estimates within each treatment group and between treatment groups will be displayed along with 2-sided 90% CIs, after inverse transformation using the exponential function.

More details regarding the model and SAS code are presented in the DPS Part 1 document.

10.8.5 NYHA FC (improved/worsened/stable) at each post-baseline assessment

The proportion of subjects having improved, having worsened or being stable will be calculated at each post-baseline assessment by treatment group. The relative risk (macitentan over placebo) will be displayed with 2-sided 90% CIs.

To calculate the relative risk, "stable" and "improved" categories will be put together as a single category "not worsened".

More details regarding the model and SAS code are presented in the DPS Part 1 document.

A listing of NYHA FC change overtime will be presented on the FAS.

10.8.6 Clinical composite outcome ('worsened', 'unchanged', 'improved')

The proportion of subjects having improved, having worsened or being unchanged will be calculated for each period of time by treatment group. The relative risk (macitentan over placebo) will be displayed with 2-sided 90% CIs.

To calculate the relative risk, "unchanged" and "improved" categories will be put together as a single category "not worsened".

More details regarding the model and SAS code are presented in the DPS Part 1 document.

A listing of clinical composite outcome will be presented on the FAS.

10.8.7 Quality of life KCCQ scores over time

For each variable:

- Overall summary score
- Clinical summary score
- Physical limitations score

the change over time is evaluated, assuming a linear pattern over time, with the use of a random coefficient regression model (with random slopes and intercepts) that includes treatment, time and time-by-treatment as factors and the score baseline value as a covariate.

The random variables for this model will be the subject-intercept and the subject-by-time interaction. The error term is assumed to follow a Normal distribution and the vector of the two correlated random effects is assumed to follow a bivariate Normal distribution, each with mean 0 and unstructured covariance matrix.

The treatment effect is determined by using estimated slopes for each study treatment group on the basis of the time-by-treatment interaction term. Estimates within each treatment group and between treatment groups will be displayed overall and for each study visit along with their corresponding 90% 2-sided CIs.

No imputation of missing values is planned.

Modelling adjustments, such as addition of a quadratic term (both in the model and in the random effect) or a different covariance pattern within subject, will be made in case of meaningful conflict between the predicted and the observed data (via a review of the predicted and observed residuals plot).

More details regarding the model and SAS code are presented in the DPS Part 1 document.

A listing of KCCQ scores and changes overtime will be presented on the FAS.

10.8.8 Percent of baseline NT-pro-BNP over time

The percent of baseline NT-pro-BNP over time is evaluated after log-transformation with the use of a random coefficient regression model (with random slopes and intercepts) that includes treatment, time and time-by-treatment as factors and the log of the ratio: baseline NT-pro-BNP over NT-pro-BNP value per IRT as a covariate.

The random variables for this model will be the subject-intercept and the subject-by-time interaction. The error term is assumed to follow a Normal distribution and the vector of the two correlated random effects is assumed to follow a bivariate Normal distribution, each with mean 0 and unstructured covariance matrix.

The treatment effect expressed as geometric means ratio and its associated 90% 2-sided CIs will be estimated based on the same model by inversely transforming using the exponential function the least squares means and 90% CIs obtained in log scale. Estimates within each treatment group and between treatment groups will be displayed overall and for each study visit along with their corresponding 90% 2-sided CIs.

No imputation of missing values is planned.

Modelling adjustments, such as addition of a quadratic term or a different covariance pattern within subject, will be made in case of meaningful conflict between the predicted and the observed data.

More details regarding the model and SAS code are presented in the DPS Part 1 document.

A listing of NT-pro-BNP values overtime will be presented on the FAS.

10.8.9 Change from baseline in accelerometer-assessed physical activity over time

For each variable as described in Sections 5.5.2.2 and 5.5.3.9:

- Proportion of time spent in light to vigorous physical activity based on a threshold of > 100 activity counts per minute
- Mean daily number of episodes of activity over 100 activity counts per minute of at least 1 minute duration
- Mean count per minute of daily activity
- Mean daily number of episodes of activity over 100 activity counts per minute of at least 5 minutes duration
- Mean daily accelerometer units

the change over time is evaluated with the use of a random coefficient regression model (with random slopes and intercepts) as described in Section 9.5.

In addition, the proportion of time spent in:

- Sedentary physical activity
- Light physical activity
- Moderate physical activity
- Vigorous physical activity

will be described at each time point of assessment by calculating for each treatment group the mean proportion of time spent in each category and will be displayed in both a graphical (bar plots) and a tabular form.

More details regarding the model and SAS code are presented in the DPS Part 1 document.

A listing of accelerometer-assessed physical activity data will be presented on the FAS.

10.8.10 Change from baseline in echocardiography left and right heart function at Week 24 and Week 52

For each variable listed in Section 5.5.3.11, the change over time is summarized.

A listing of echocardiography data will be presented on the FAS.

10.8.11 Percent of baseline MR-proANP over time

The percent of baseline MR-proANP over time is analyzed by means of an ANCOVA at Week 24 adjusting for the log of the variable baseline value.

A listing of MR-proANP values overtime will be presented on the FAS.

10.8.12 NT-pro-BNP over time

Summary statistics of NT-pro-BNP over time will be presented per treatment group for the FAS for all assessments from screening, through run-in and double-blind periods. In addition, summaries will be presented on the MRI and PRI sets.

A listing of NT-pro-BNP values overtime will be presented on the FAS.

10.8.13 Patient global assessment

Summary statistics of PGA 7-point scale over time will be presented per treatment group for the FAS.

The proportion of subjects having favorable, unchanged or unfavorable ratings will be calculated at each post-baseline assessment by treatment group.

The relative risk (macitentan over placebo) will be displayed at each time point with 2-sided 90% CIs. To calculate the relative risk, "unchanged" and "favorable changes" categories will be put together as a single category "not worsened".

A listing of all PGA data will be presented on the FAS.

10.9 Analysis of safety variables

10.9.1 Adverse events

All AEs collected from signing of informed consent up to EOS will be reported in the subject listings together with the assigned PT/SOC. The SAEs collected during PTOP will be included in these subject listings. Treatment-emergent AEs during the placebo run-in, macitentan run-in and double-blind treatment periods [see definition in Section 5.6.1] will be flagged accordingly. Missing intensity is reported in listings and is replaced with the worst case only for the summary tables.

AE summaries will be presented for the safety set (i.e., subjects received double-blind treatment) and by double-blind treatment group unless specified otherwise. A few select summary tables, will be presented for the PRI and MRI sets as in Table 5 and detailed in the sections below:

Table 5 Overview of summary tables for AEs

Summary	Period *	Analysis set
Overview of AEs	Study treatment	SAF
	Double-blind	SAF
TEAEs	Study treatment	SAF
	PRI, MRI, double-blind	SAF
	Double-blind, related to study treatment	SAF
	Double-blind, by maximum intensity	SAF
	Placebo run-in (overall)	PRI
	Macitentan run-in (overall)	MRI
	Combined macitentan (overall)	MRI
	Combined macitentan, related to study treatment (overall)	MRI
	Combined macitentan, by maximum intensity (overall)	MRI
Serious TEAEs	Study treatment	SAF
	PRI, MRI, double-blind	SAF
	Placebo run-in (overall)	PRI
	Macitentan run-in (overall)	MRI
	Combined macitentan (overall)	MRI
AE leading to treatment discontinuation **	Double-blind	SAF
	Double-blind TEAE	SAF
	Placebo run-in (overall)	PRI
	Macitentan run-in (overall)	MRI
AE with fatal outcome **	Double-blind	SAF
	Combined macitentan (overall)	MRI
Deaths	Double-blind	SAF
TEAEs of special interest	Double-blind	SAF
	Combined macitentan (overall)	MRI

*presented by treatment group unless specified otherwise as overall
** Period of discontinuation/fatality may include AEs started and worsened in previous periods, therefore a separate table is included for double blind treatment-emergent AEs leading to discontinuation.
AE = adverse event; MRI = Macitentan Run-In; PRI = Placebo Run-In; SAF = Safety ; TEAE = Treatment-Emergent Adverse Event.

A summary table with an overview of TEAEs is provided displaying, for each treatment group, counts and percentages of subjects having experienced at least a TEAE, a severe AE, a study-treatment related AE, an AE leading to study treatment discontinuation, a non-serious frequent AE [see Section 10.9.2.5], a serious AE, a study-treatment related serious AE, a fatal SAE. The overview is also repeated for double-blind TEAEs.

10.9.1.1 Treatment emergent adverse events

Treatment-emergent AEs will be summarized presenting counts and percentages of subjects having experienced at least one TEAE by SOC and PT within each SOC as well as by PT. The summary tables will be presented in descending order according to the incidence in double-blind macitentan treatment group (e.g., SOC and PT within each SOC with the highest number of occurrences appears first). If frequencies of SOC/PTs are the same, alphabetical order will be used. The summaries will be presented split by treatment emergent period (PRI, MRI and for the double-blind TEAEs) within each treatment group.

Double-blind TEAEs related to study treatment will be summarized similarly by SOC/ PT. Double-blind TEAEs by maximum intensity will be summarized similarly by PT, only.

A graphical display of the relative risk (macitentan over placebo) together with the 2-sided 95% CLs for the most common double-blind TEAEs in either treatment groups (i.e., in at least 5 subjects [~ 3%] in either treatment groups) is provided by PT. The graph is presented in descending order according to the incidence in the macitentan treatment group. The choice to use 95% CLs instead of 90% CLs (as done for the efficacy analysis) is to be in line with FDA guidance for QT/QTc studies.

Placebo run-in TEAEs and macitentan run-in TEAEs and combined macitentan TEAEs will be summarized for the placebo run-in set and macitentan run-in set, respectively. Number and percentages of subjects having experienced at least a TEAE will be presented by SOC and PT within each SOC. The summary tables will be presented in descending order (e.g., SOC and PT within each SOC with the highest number of occurrences appears first). If frequencies of SOC/PTs are the same, alphabetical order will be used.

Combined macitentan TEAEs related to study treatment will be summarized similarly by SOC and PT. Macitentan TEAEs by maximum intensity will be summarized similarly by PT, only.

10.9.2 Deaths, other serious adverse events

10.9.2.1 Deaths

All-cause death cases will be reported in the subject listing. Treatment-emergent deaths during the placebo run-in, macitentan run-in, and double-blind treatment periods [see definition in Section 5.6.2] will be flagged accordingly.

A summary table reporting the incidence of deaths in the double-blind treatment period will be provided for the safety set. Number and percentages of deaths will be summarized together with the primary reason for death in descending order according to the incidence in double-blind macitentan treatment group.

10.9.2.2 Serious adverse events

All SAEs will be reported in a separate subject listing. Treatment-emergent SAEs during the placebo run-in, macitentan run-in and double-blind treatment periods [see definition in Section 5.6.3] will be flagged accordingly. SAEs reported during PTOP will also be included in the listing.

Treatment-emergent SAEs will be summarized by SOC and PT within each SOC similarly to AEs [see Section 10.9.1]. The summaries will be presented split by treatment emergent period (PRI, MRI and for the double-blind TEAEs) within treatment group.

Placebo run-in treatment-emergent SAEs, macitentan run-in treatment-emergent SAEs and combined macitentan treatment-emergent SAEs will be summarized for the placebo run-in set and macitentan run-in set, respectively. Counts and percentages of subjects having experienced at least one treatment-emergent SAE will be presented by SOC and PT within each SOC.

10.9.2.3 Adverse events leading to study treatment discontinuation

A separate subject listing is provided for all AEs leading to discontinuation of study treatment indicating the treatment period discontinued.

AEs leading to premature discontinuation of study treatment will be summarized similarly to TEAEs [see Section 10.9.1.1], by SOC and PT within each SOC, on the double-blind safety set (SAF)*.

**Note the above summary will include all AEs leading to discontinuation of double-blind treatment and by definition may include AEs that were emergent and worsened during the run-in periods. This is considered an appropriate conservative approach in comparison of double-blind treatment groups. In this summary of AEs, the action of withdrawing the treatment is considered as a worsening in the double-blind treatment period in the absence of any worsening of intensity for AEs that were emergent in the run-in period.*

The summary of AEs leading to discontinuation is repeated for the double-blind TEAEs on the SAF. In addition, AEs leading to discontinuation of study treatment during the run-in periods are presented on the PRI and MRI analyses sets respectively.

10.9.2.4 Other significant adverse events

Treatment-emergent AEs with fatal outcome will be summarized for the safety set and for the combined macitentan period on the MRI set by SOC and PT within each SOC. *Note this summary will include any AEs that are emergent or worsened in the run-in periods but with fatal outcome occurring after double-blind treatment start for these subjects in the safety set.*

A separate subject listing is provided for all subject AEs with fatal outcome. Deaths occurring during the run-in periods will be included in the listing.

For each area of clinical interest, double-blind TEAEs and macitentan TEAEs of special interest [see definition in Section 5.6.5.2] will be summarized for the safety set and MRI set respectively, presenting counts and percentages of subjects having experienced at least one TEAE of special interest by frequency of PT within each AESI group. Separate subject listings will be provided for all AEs of special interest.

10.9.2.5 Eudract and CTgov summaries

For the disclosure of the results to EudraCT and ClinicalTrials.gov (and not for the purpose of the CSR), separate similar summary tables are provided by SOC and PT within each SOC for:

- treatment-emergent SAEs related to study treatment during the double-blind treatment period
- treatment-emergent SAEs with fatal outcome during the double-blind treatment period
- treatment-emergent SAEs related to study treatment with fatal outcome during double-blind treatment period.

In addition, for the disclosure of the results to VICTOR, treatment-emergent SAEs during double-blind treatment period are summarized displaying, for each treatment group, counts and percentages of subjects with at least a treatment-emergent SAE plus the number of events (counted exactly the number of times they occurred also within a subject) by SOC and PT. The summary table is presented in descending order according to the incidence in the macitentan treatment group (i.e., SOC and PT within each SOC with the highest number of occurrences appears first). Equal frequency of different PTs is sorted in alphabetical order of the PT.

Double-blind treatment-emergent non-serious AEs (any AE not classed as serious) with frequencies ≥ 5% in at least one treatment arm (referred to as non-serious frequent AEs) are summarized displaying, for each treatment group, counts and percentages of subjects with at least a treatment-emergent non-serious frequent AE plus the number of events (counted exactly the number of times they occurred also within a subject) by SOC and individual preferred. The summary table is presented in descending order according to the number of events in the macitentan treatment group (i.e., SOC and individual PT within each SOC with the highest number of occurrences appears first). Equal frequency of different individual PTs is sorted in alphabetical order of the individual PT.

10.9.3 Electrocardiogram

The count and percentage of subjects with presence of treatment emergent qualitative finding [Section 5.6.7, atrial fibrillation/atrial flutter] during the double-blind study treatment period will be summarized in a shift table, for the safety set. These will be presented as shift from macitentan run-in baseline.

In addition, shift for ECG findings will be presented by history of atrial fibrillation or atrial flutter (yes/no), as reported in the specific medical history CRF [see Section 5.2.2].

10.9.4 Laboratory tests

All hematology and chemistry parameters [see Section 5.6.8] provided by the central and local laboratory will be displayed in subject listings, including those from unscheduled visits. MLAs will be flagged accordingly.

For Week 4, Week 8, Week 16, Week 24, Week 36, Week 52, the selected laboratory test parameters [Section 10.9.4.2] will be summarized by treatment group displaying descriptive statistics for:

- observed value at each visit specified
- absolute change from macitentan run-in baseline to each visit.

In each evaluation, will be included only subjects who had both the assessments at macitentan run-in baseline and post baseline visit.

10.9.4.1 Fasted state laboratory values

Laboratory values collected at screening / re-screening in fasted state [Section 5.6.8.1] will be summarized on FAS. Only assessments from central lab will be considered.

10.9.4.2 Marked laboratory abnormalities

For the evaluation of MLAs, separate summaries will be presented for hematology and chemistry laboratory tests.

A dictionary listing of definitions of MLA (LL, LLL, HH, HHH) is provided for each parameter.

For each category (i.e., LL, LLL, HH, HHH, HHHH), double-blind treatment-emergent MLAs [see definition in Section 5.6.8.2] will be summarized by treatment group, displaying counts and percentages of subjects with at least one treatment-emergent MLA for each parameter for which the marked abnormality is defined. Percentages will be calculated as number of subjects with at least one treatment-emergent MLA for the parameter under consideration divided by the number of subjects with any post-baseline laboratory measurement. The summary will be partitioned to include macitentan run-in emergent MLAs and placebo run-in emergent MLAs occurring in the same safety set.

Macitentan run-in emergent MLAs will be summarized in the same way on MRI set.

Shifts from macitentan run-in baseline to worst post-baseline abnormality category (during double-blind period) up to EOT + 30 days will be summarized with frequencies of subjects and percentages per category of shift. Categories considered are LLL, LL, L (L = below LLN), normal (between LLN and ULN), H (H = above ULN), HH, HHH. For baseline the category 'Missing' is also considered. Subjects experiencing a worsening from baseline in two different directions (e.g., going from normal at baseline once to H and once to L post-baseline) are counted in both directions.

The denominator for percentages is the number of subjects with at least one post-baseline assessment.

In-text tables will be provided for the following parameters:

- Hematology:
  - Hemoglobin
  - Hematocrit
  - Erythrocytes
- Chemistry:
  - ALT
  - AST
  - Alkaline phosphatase
  - Total bilirubin

For the above selected parameters, values over time from Screening visit will be plotted.

10.9.4.3 Additional liver laboratory abnormalities

For the evaluation of liver laboratory abnormalities, measurements from unscheduled visits will be included. All subject with at least one treatment-emergent LTA will be provided in a separate subject listing.

Double-blind treatment-emergent LTAs [see definition in Section 5.6.8.3] will be summarized similarly to the MLAs [see Section 10.9.4.2].

A similar summary is created also for single-blind macitentan run-in emergent liver abnormalities.

10.9.5 Vital signs and body weight

Vital signs (blood pressure measurements, pulse rate) and body weight [see Section 5.6.6] will be reported in a subject listing.

SBP, DBP, pulse rate, and body weight will be summarized, by treatment group, displaying descriptive statistics for:

- observed value at each visit up to Week 52
- absolute change from macitentan run-in baseline to each visit up to Week 52.

Vitals signs parameters will be analyzed regardless of the measurement position.

The number and percent of subjects with a decrease from baseline during double-blind period in SBP [see definition in Section 5.6.6.1] will be summarized by treatment group on the safety set, by displaying frequency and percentages of subjects together with 2-sided 90% CLs. Similar summary is presented for decrease in SBP during single-blind macitentan run-in period on the MRI set.

Vital signs and body weight will be described over time up to EOT + 30 days by means of a random coefficient regression model as described in Section 9.5.

The change in SBP and DBP, pulse rate and body weight over time is evaluated by means of a random coefficient (9.5) regression model (with random slopes and intercepts). The dependent variable of the model is the change from baseline. Randomized study treatment visit (i.e., Week 4/Week 8/Week 16/Week 24/Week 36/Week 52), treatment-by-visit interaction terms, and baseline value will be included as fixed effects in the model. The random variables for this model will be both the subject-intercept and the subject-by-time interaction. The error term is assumed to follow a normal distribution and the vector of the two correlated random effects is

assumed to follow a bivariate normal distribution, each with mean 0 and unstructured covariance matrix. This statistical model allows for missing data, assuming they are missing at random. Therefore, no imputation of missing values is planned.

The treatment effect is determined by using estimated slopes for each study treatment group on the basis of the treatment-by-time interaction term from the mixed model.

The results will be presented in a summary table displaying:

- For each visit, the adjusted least squares means and 2-sided 90% CLs for each treatment group as well as the adjusted least squares means difference macitentan vs placebo with 2-sided 90% CLs.

10.9.6 Other safety variables

10.9.6.1 Number of all-cause hospital admissions up to 30 days after study treatment discontinuation

A negative binomial model as described in Section 9.3 is used for the statistical analysis. The null hypothesis of no treatment effect is tested by a 2-sided Wald-type test. 2-sided 90% Wald CLs is calculated for the relative reduction in mean annualized hospitalization rate for macitentan compared to placebo.

10.9.6.2 Glomerular Filtration Rate

Estimated GFR [see definition in Section 11.9] is summarized on the SAF, by treatment group, displaying descriptive statistics for:

- observed value at each visit up to Week 52
- absolute change from macitentan run-in baseline to each post baseline visit.

GFR will be described over time up to 30 days after study treatment discontinuation by means of a random coefficient regression model as described in Section 9.5.

10.9.6.3 Physical examination

Physical examinations [as defined in Section 5.6.9.3] performed during the course of the study will be reported in a subject listing.

10.9.6.4 Heart failure signs and symptoms of special interest

Heart failure signs and symptoms of special interest [as defined in Section 5.6.9.4] assessed during the course of the study will be reported in a subject listing.

10.10 Analysis of quality of life variables

The three scores derived from KCCQ will be summarized on the FAS, by treatment group, displaying descriptive statistics for:

- observed value at each visit
- absolute change from baseline to each visit.

10.11 Analysis of sub-study variables

The 6-MWD will be summarized on the SSAS by treatment group, displaying descriptive statistics for:

- observed value at each visit
- absolute change from baseline to each visit.

Sub-study data including 6-MWD and Borg Dyspnea Index will be listed.

11 GENERAL DEFINITIONS AND DERIVATIONS

11.1 Time windows

To allow analysis of data at the relevant planned (scheduled) visits during the double-blind period, recorded assessments, including PTOP and unscheduled ones, are re-assigned to the most appropriate visit according to the best fitting time window for that visit [see Table 6 and Table 7]. Note that the visit windows are contiguous in order to retain all values in the analysis by time point.

Table 6 Visit time windows NT-pro-BNP

Visits	Target day*	Min. day	Max. day
Baseline	1	No limit	1
Week 4	28	2	70
Week 16	112	71	140
Week 24	168	141	267
Week 52	365	268	No upper limit

* Number of days from start of double-blind treatment.

Table 7 Safety Visit time windows and efficacy with similar schedule

Visits	Target day*	Min. day	Max. day
Week 4^	28	2	42
Week 8	56	43	84
Week 16	112	85	140
Week 24	168	141	210
Week 36	252	211	308
Week 52	365	309	No upper limit

* Number of days from start of double-blind treatment.

^ For KCCQ / PGA assessments this visit is not present and Visit 9 - Week 8 min. day = 2.

Echocardiographic measures, MR-ProANP presented at Week 24 and Week 52 only.

KCCQ = Kansas City Cardiomyopathy Questionnaire; MR-proANP = Mid-Regional pro-Atrial Natriuretic Peptide; PGA = Patient global assessment

In the event that there is more than one value within the same time window, the value closest to the planned assessment date will be taken. In the event of equidistant values from the planned time point, the last assessment will be considered for the analyses. Where multiple assessments fall on the same day the latest is then used.

For laboratory values the central laboratory values are always selected in favor of any local laboratory value. If more than one value falls on the same date and time (and laboratory) then the one with the last sequential number in SDTM will be used.

11.2 Dates

11.2.1 Signed informed consent

It is the date of signed informed consent collected in the "Informed Consent" CRF.

11.2.2 Screening date

It is the date of screening visit collected in the "Visit Summary" CRF. If the subject is screened once, it is the visit date associated with the first screening attempt. For subjects screened more than once, the date of screening is the visit date associated with the latest re-screening attempt, regardless if the subject is entered into the run-in.

11.2.3 Randomization date

It is the date of subject randomization in the study from IVRS.

11.2.4 Study treatment start date

This is the "Treatment start date" from the first interval, in chronological order, recorded in the study drug log CRF and is equivalent to start date of Single-blind placebo run-in.

11.2.5 Study treatment end (EOT) date

Also known as the date of end of treatment or treatment discontinuation, this is the "Treatment end date" from the last interval, in chronological order, recorded in the study drug log CRF. If missing or incomplete, rules in Section 12 are followed.

11.2.6 Single-blind placebo run-in start date

It is the first day of intake of single-blind placebo. It is derived as the first treatment start date (in chronological order) in the "Study Drug Log" CRF if 'Treatment period' is flagged as 'Single-blind placebo run-in'.

11.2.7 Single-blind placebo run-in end date

It is the last day of intake of single-blind placebo. It is derived as the last treatment end date (in chronological order) in the "Study Drug Log" CRF if 'Treatment period' is flagged as 'Single-blind placebo run-in'.

11.2.8 Single-blind Macitentan run-in start date

It is the first day of intake of single-blind macitentan. It is derived as the first treatment start date (in chronological order) in the "Study Drug Log" CRF if 'Treatment period' is flagged as 'Single-blind Macitentan run-in'.

11.2.9 Single-blind Macitentan run-in end date

It is the last day of intake of single-blind macitentan. It is derived as the last treatment end date (in chronological order) in the "Study Drug Log" CRF if 'Treatment period' is flagged as 'Single-blind Macitentan run-in'.

11.2.10 Double-blind treatment start date

It is the first day of intake of double-blind treatment. It is derived as the first treatment start date (in chronological order) in the "Study Drug Log" CRF if 'Treatment period' is flagged as 'Double-blind treatment'.

11.2.11 Double-blind treatment end date

It is the last day of intake of double-blind treatment. It is derived as the last treatment end date (in chronological order) in the "Study Drug Log" CRF if 'Treatment period' is flagged as 'Double-blind treatment' and the reason for treatment end is ≠ "TEMPORARILY INTERRUPTED DUE TO AN AE" or "TEMPORARILY INTERRUPTED NOT DUE TO AN AE".

Rules for handling incomplete/missing dates are detailed in Section 12.

11.2.12 End of study (EOS) date

For subjects not entering the placebo run-in period, EOS date is the "Date of subject's End of Study" collected in "Eligibility for Placebo Run-in" CRF form.

For subjects who died during the study, the EOS date is "Date of death" collected in "Death" form.

Subjects completing the EOS visit (Visit 14 or PTOP4), the date of the visit is the EOS date. Note subjects who have not completed Week 24 at the time of global protocol version 6 approval, will end PTOP at 24 weeks (PTOP2). Subjects that already completed PTOP2 or PTOP 3 visits at the time of global protocol version 6 approval will end PTOP at the next PTOP visit.

For all subjects who enter the SERENADE OL, the EOS date is the 'Date of first dose in OL study' - 1 (if the date is available) or the 'Date of enrollment in OL study' - 1, as collected in the "Open-label enrollment" CRF.

Otherwise for all other subjects, the EOS date is collected in "Study discontinuation" CRF form according to the reason for discontinuation as follows:

- Lost to follow-up = date of last successful contact.
- Subject decision/ withdrawal of consent = date of subject decision
- Physician decision = date of physician decision
- Sponsor decision = Date subject was informed of sponsor decision

11.3 Baseline

Except where otherwise specified, the baseline for a given measurement is the last value assessed ≤ the start date of double-blind treatment. If unscheduled/re-test visits are performed on the day of double-blind treatment start, the available value of the last unscheduled/re-test visit on double-blind treatment start date is considered as baseline.

11.3.1 Macitentan run-in Baseline

The macitentan run-in baseline for a given measurement is the last value assessed ≤ the start date of single-blind macitentan run-in treatment [Section 11.2.8].

11.4 Study day

The study day (for the double-blind period) is the number of days elapsed since the day of first double-blind treatment plus 1 (start of double-blind treatment is considered Day 1). For dates prior to Day 1, study day is the negative number of days between the date under consideration and the date of first double-blind treatment. Therefore, the study day is always different from zero.

11.5 Treatment emergent periods

Note: the periods below refer to the treatment emergent periods for reporting safety and other variables as appropriate.

11.5.1 Study treatment period

Start date is the date of first dose of single-blind placebo run-in treatment and end date is the EOT date + 30 days.

11.5.2 Double-blind treatment period

Start date is the date of first dose of double-blind treatment and end date is the last dose date + 30 days.

11.5.3 Placebo run-in period

Start date is the first dose of single-blind placebo treatment and end date is the day prior to first dose of macitentan in the run-in period. Subjects that entered this period and who do not receive macitentan in the single-blind macitentan run-in period, the end date is the last dose date + 30 days.

11.5.4 Macitentan run-in period

Start date is the first dose of single-blind macitentan treatment and end date is the day prior to first dose of treatment in the double-blind treatment period. Subjects that entered this period and who do not receive double-blind treatment, the end date is the last dose date + 30 days.

11.5.5 Combined macitentan treatment period

Start date is the first dose of single-blind macitentan run-in treatment and end date is the date of the last dose of any macitentan treatment (single-blind or double-blind)+ 30 days.

11.5.6 Safety follow-up period

The safety follow-up period starts the day after the EOT date [see definition in Section 11.2.5] and ends 30 days thereafter with the End-of-Study visit, or PTOP1 visit for those subjects who prematurely discontinue double-blind study treatment and enter PTOP.

Duration of FU (days) is defined as the time elapsing between date of last study drug intake + 1 day and the safety follow-up (FU) visit + 1 (in days).

11.5.7 PTOP period

The PTOP period starts the day of visit PTOP1 and ends at day of the last PTOP visit (or EOS date if discontinued early - see Section 11.2.12).

11.5.8 End of Study (EOS) visit

The EOS is reached when the safety follow-up period or, if applicable, PTOP have been completed. For an individual subject, the study is completed with the EOS visit, which is either Visit 14 (safety follow-up visit) for subjects who completed the treatment period as per protocol, or PTOP4 for subjects entering PTOP (or PTOP1 depending on time point of premature treatment discontinuation). Subjects who are past PTOP2 at the time of consenting to global protocol version 6 will return for an EOS visit within 60 days but no later than Week 52.

Hence EOS visit and safety follow-up visit may differ when subject enters PTOP.

11.5.9 Duration of time on study

Duration on study (weeks) is defined as the time elapsing between date of Screening / re-screening and the EOS visit + 1 (in days) / 7.

11.5.10 Duration of PTOP

Duration of PTOP (weeks) is defined as the time elapsing between date of PTOP1 and the last PTOP visit (or study discontinuation) + 1 (in days) / 7.

11.6 Subjects in Run-in

A subject is considered as entered in placebo run-in if answer ='Yes' to the question 'Did the subject enter the single-blind run-in period?' in the "Eligibility for Placebo Run-in" CRF. A subject is considered as entered in macitentan run-in if answer ='Yes' to the question 'Did the subject enter the single-blind run-in period?' in the "Eligibility - Macitentan Run-in" CRF.

11.7 Run-in failures

Includes all subjects defined as placebo run-in failures [see Section 5.1.2] or macitentan run-in failures [see Section 5.1.3].

11.8 Randomized subjects

A subject is randomized into the study when he/she is assigned with a randomization number in the IVRS.

11.9 Conversion rules

- BMI (kg/m^2) = weight (kg) / (height (cm) / 100)^2

In case height is measured in "inch" and body weight in "lbs" the following conversion rules are applied in CRF:

  - Height (cm) = height (inch) × 2.54
  - Weight (kg) = weight (lbs) × 0.4536

- PVR (wood units) = [(mPAP – PAWP)-CO]*80, if PAWP is missing then substitute with LVDEP
- DPG (mmHg) = diastolic PAP – PAWP
- eGFR =175 x $(S_{Cr})^{-1.154}$ x$(age)^{-0.203}$ x0.742 [if female] x1.212 [if Black]

  where Scr (standardized serum creatinine) in mg/dL, age in years.

12 HANDLING OF MISSING/INCOMPLETE DATE AND TIME FIELDS

This section describes some general principles to be followed in the case of missing or incomplete dates/times.

The dates in these types that are missing or incomplete are derived as follows:

- Dates are split in 3 parts: year, month and day. Year is the top-level, month is medium level and day is low level. If a part expected to contain a number is numeric but the value is outside a valid range, the complete date is handled as missing. For example, if date = 44Nov2000 the whole date is considered to be missing.
- If a Part is missing, all other parts of a lower level are considered to be missing. This means that a ddmmyy date '21--99' is considered as '----99'.
- Missing parts are changed into acceptable non-missing values in a way depending on the type of date to be replaced.

Missing parts for specific dates/times are changed into acceptable non-missing values as described in the table below.

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 dates refer to the first or last date, respectively, when ordered in sequence. Other sources should be considered for possible dates including death date and date of withdrawal from study when determining possible dates.

Type of date/time	Date/time is incomplete	Date/time is missing
Single-blind placebo run-in end date	Use the earliest date between the: start of macitentan run-in minus 1 day (when applicable) upper limit date of Visit 13 EOS date of death date of last successful contact	Use the earliest date between the: start of macitentan run-in minus 1 day (when applicable) date of Visit 13 EOS date of death date of last successful contact

Type of date/time	Date/time is incomplete	Date/time is missing
Single-blind macitentan run-in end date	Use the earliest date between the: start of double-blind treatment minus 1 day (when applicable) upper limit date of Visit 13 EOS date of death date of last successful contact	Use the earliest date between the: start of double-blind treatment minus 1 day (when applicable) date of Visit 13 EOS date of death date of last successful contact
Double-blind treatment start date	- Day missing: Replaced by day of the randomization date. In case the day of randomization is in the previous month, replace the day by 1 - Day and month are missing: replace entirely by the randomization date	Replace entirely by the randomization date, if evidence the subject was treated in the double-blind. Otherwise is considered not treated. Evidence includes Visit > Visit 6 other than study discontinuation. Double-blind treatment end date present. Other information may be available during the programming should be considered and updated in the programming specifications.
EOT	Use the earliest date between the: upper limit date of Visit 13 EOS date of death date of last successful contact	Use the earliest date between the: Date of Visit 13 EOS date of death date of last successful contact
AE resolution date	The upper limit	No approximation, the AE is considered as ongoing in the analysis

Type of date/time	Date/time is incomplete	Date/time is missing
AE onset date	**If the resolution date of the AE is before the single-blind placebo run-in start date:** the lower limit. **If the resolution date of the AE is on or after the single-blind placebo run-in start date and before the single-blind macitentan run-in start date (or macitentan run-in start date is missing):** if the single-blind placebo run-in start falls in the range of possible dates, this date is used. Otherwise, the lower limit is used. **If the resolution date of the AE is on or after the single-blind macitentan run-in start date and before the double-blind treatment start date (or double-blind treatment start date is missing):** if the single-blind macitentan run-in start falls in the range of possible dates, this date is used. Otherwise, the lower limit is used. **If the resolution date of the AE is on or after the double-blind treatment start date:** if the double-blind treatment start falls in the range of possible dates, this date is used. Otherwise, the lower limit is used. **If the resolution date of the AE is missing:** if the start date of the last period accessed by the subject (i.e., double-blind, macitentan run-in, placebo run-in) falls in the range of possible dates, this date is used. Otherwise, the lower limit is used.	**If the end date of the AE is before the single-blind placebo run-in start date:** the date of AE resolution. **If the end date of the AE is on or after the single-blind placebo run-in start date and before the single-blind macitentan run-in start date:** the single-blind placebo run-in start. **If the end date of the AE is on or after the single-blind macitentan run-in start date and before the double-blind treatment start date:** the single-blind macitentan run-in start **If the end date of the AE is on or after the double-blind treatment start date:** the double-blind treatment start **If the resolution date of the AE is missing:** the start date of the last period accessed by the subject (i.e., double-blind, macitentan run-in, placebo run-in) is used.
AE intensity change date	If the start date of the next period following the AE onset period / accessed by the subject (i.e., double-blind, macitentan run-in, placebo run-in) falls in the range of possible dates, this date is used. Otherwise, the lower limit is used.	Start date of the next period following the AE onset period / accessed by the subject

Type of date/time	Date/time is incomplete	Date/time is missing
Concomitant medication end date	The upper limit unless: the medication started prior to placebo run-in start date and - upper limit is after the placebo run-in start and 'Ongoing at start of treatment?' is ticked 'No', in which case it is 1 day before the placebo run-in start date.	No replacement.
Concomitant medication start date	If the end date of the medication is not before the informed consent signature date and if the informed consent date falls in the range of possible dates, the informed consent date is used. If the end date of the medication is not before the placebo run-in start date and if the placebo run-in start date falls in the range of possible dates, the placebo run-in start date is used. In all the other cases, the lower limit is used.	No replacement, the medication is considered to have started before the date of consent.
EOS	Upper limit	Data extract date (after dblock).
Death date (Investigator assessment)	Use the lower limit	No replacement.
WHF/HF/CV hospitalization date (Investigator assessment)	If the onset date of the AE leading to hospitalization falls in the range of possible dates, it is the onset date of the AE. In all the other cases, it is the lower limit.	The onset date of the AE leading to hospitalization.

13 LIST OF SUMMARY TABLES, LISTINGS AND FIGURES

This list and the mock layouts are provided in the DPS Part 1 document.

14 REFERENCES

[Brookmeyer 1982] Brookmeyer R, Crowley JA. CI for the median survival time. Biometrics. 1982;38:29-41.

[Cuzick 2005] Cuzick J. Forest plots and the interpretation of subgroups. Lancet. 2005 Apr 9-15;365(9467):1308.

[Dmitrienko 2010] Dmitrienko A, Tamhane AC, Bretz F. Multiple testing problems in pharmaceutical statistics. CRC Press;2010.

[FDA 2010] Guidance for Industry: Pharmacokinetics in Patients with Impaired Renal Function — Study Design, Data Analysis, and Impact on Dosing and Labeling. March 2010.

[Freedson 1998] Freedson PS, Melanson E, Sirard J. Calibration of the computer science and applications, Inc. Accelerometer. Med Sci Sports Exerc. 1998, 30;777-781.

[ICH 1995] ICH harmonized tripartite guideline E3: Structure and content of Clinical study reports CPMP/ICH/137/95 November 1995.

[ICH 1998] ICH harmonized tripartite guideline E9: statistical principles for clinical trials CPMP/ICH/363/96 September 1998.

[Packer 2001] Packer M. Proposal for a new clinical end point to evaluate the efficacy of drugs and devices in the treatment of chronic heart failure. J Cardiac Fail. 2001;7(2):176-182.

[Pocock 2002] Pocock SJ, Clayton TC, Altman DG. Survival plots of time-to-event outcomes in clinical trials: good practice and pitfalls. The Lancet 2002;359:1686-9.

[Pocock 2012] Pocock S, Ariti C, Collier T, Wang D. The Win Ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. European Heart J. 2012;33:176-182.

[Quanjer 1993] Quanjer PH, et al. Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. Eur Respir J Suppl. 1993 Mar;16:5-40.

[Robertson 2011] Robertson W, Stewart-Brown S, Wilcock E, Oldfield M, Thorogood M. Utility of Accelerometers to Measure Physical Activity in Children Attending an Obesity Treatment Intervention. J Obes. 2011:1-8.

[Solomon 2012] Solomon SD, Zile M, Pieske B, et al. The angiotensin receptor neprilysin inhibitor LCZ696 in heart failure with preserved ejection fraction: a phase 2 double-blind randomised controlled trial. Lancet. 2012;380:1387-95.

[Stokes 2000] Stokes ME, Davis CS, Koch GG. Categorical data analysis using the SAS® system, 2nd Edition. SAS Institute Inc.

[Troiano 2008] Troiano RP, et al. Physical Activity in the United States Measured by Accelerometer. American College of Sports Medicine. 2008.

15 APPENDICES

Appendix A Protocol synopsis, and schedule from Protocol version 6 (6 February 2020)

TITLE	A multi-center, double-blind, placebo-controlled Phase 2b study to evaluate the efficacy and safety of macitentan in subjects with heart failure with preserved ejection fraction and pulmonary vascular disease
ACRONYM	SERENADE Macitentan in heart failure with pre**SER**ved ej**E**ction fractio**N** and pulmon**A**ry vascular **D**iseas**E**
OBJECTIVES	**Primary objective** To evaluate whether macitentan 10 mg reduces n-terminal pro-brain natriuretic peptide (NT-pro-BNP) versus placebo at Week 24 in subjects with heart failure with preserved ejection fraction (HFpEF) and pulmonary vascular disease. **Secondary objectives** To evaluate the effect of macitentan 10 mg as compared to placebo on: • Quality of life • Daily physical activity • Worsening of heart failure **Other objectives** Other objectives are described in Section 2.3.
DESIGN	A prospective, multi-center, double-blind, randomized, placebo-controlled, parallel-group design Phase 2b study. Subjects will be randomized in a 1:1 ratio to either macitentan or placebo. Treatment allocation will be stratified by NT-proBNP level (< 1000 pg/mL and ≥ 1000 pg/mL) at macitentan run-in entry.
PERIODS	**Screening period:** Lasts up to 30 days; starts with the signature of the Informed Consent Form at Visit 1 and ends prior to administration of the first dose of study treatment at Visit 2.

	Run-in period: 9-week run-in period consisting of a single-blind placebo run-in of 4 weeks, followed by a single-blind macitentan run-in of 5 weeks. The placebo run-in starts with administration of the first dose of placebo at Visit 2 and ends on the day before Visit 4. The macitentan run-in starts with the administration of the first dose of macitentan at Visit 4 and ends with the subject's randomization at Visit 6. **Double-blind treatment period:** The double-blind treatment period consists of 2 phases: *Core phase:* The double-blind core phase will last for 24 weeks. It starts with the administration of the first dose of double-blind study treatment and ends on the day before Visit 11 (Week 24). Subjects who are still in the core phase at the time of global protocol Version 6 approval will end treatment at 24 weeks and will not proceed to the extension phase [see Section 3.1.1.8]. *Extension phase:* The double-blind extension phase will last for 28 weeks. It starts on the day of Visit 11 (Week 24) and ends on the day of last study treatment intake. Subjects who are in the extension phase at the global protocol Version 6 approval will return for an EOT visit within 60 days, but no later than Week 52. **Post-treatment observation period:** Subjects who prematurely discontinue double-blind study treatment (core or extension phase) will be asked to enter a post-treatment observation period (PTOP), which ends 52 weeks after randomization. The PTOP starts with visit PTOP1, which corresponds to the safety follow-up visit. Thereafter, visits are scheduled at Week 24 (PTOP2), Week 36 (PTOP3) and Week 52 (PTOP4), depending on time point of premature discontinuation. Subjects who have not completed Week 24 at the time of global protocol Version 6 approval, will end PTOP at 24 weeks (PTOP2) [see Section 3.1.1.8]. **Safety follow-up period:** The safety follow-up period starts on the day after the last dose of study treatment and

	ends 30 days thereafter with the End-of-Study (EOS) visit, or PTOP1 visit for those subjects who prematurely discontinue study treatment. **End-of-Study:** EOS is reached when the safety follow-up period or, if applicable, PTOP have been completed. For an individual subject, the study is completed with the EOS visit, which is either Visit 14 (safety follow-up visit) for subjects who completed the treatment period as per protocol, or PTOP4 for subjects who prematurely discontinued. **Transition to the SERENADE OL extension study (AC-055G203):** Subjects who remain in the SERENADE study for 52 weeks after randomization may transition to the SERENADE OL (AC-055G203) study if they meet the eligibility criteria defined in the SERENADE OL protocol. Eligible subjects who are consented to global protocol Version 6 will be allowed to transition to the SERENADE OL study if they remained in the SERENADE study for 24 weeks after randomization [see Section 3.1.1.8].
PLANNED DURATION	Approximately 2.5 years from first subject, first visit to last subject, last visit.
SITE(S) / COUNTRY(IES)	77 sites in 17 countries.
SUBJECTS / GROUPS	It was planned to enroll 300 subjects in 2 groups; randomized in a 1:1 ratio (150 subjects per group) by an interactive Voice/Web System to macitentan or placebo using the NT-proBNP value observed before entry in macitentan run-in as a stratification factor. Following closure of Screening on 27 December 2019, it is expected that approximately 140 subjects will be randomized.
INCLUSION CRITERIA	1. Signed informed consent prior to any study-mandated procedure 2. Male or female subjects ≥ 18 years of age

	3. Signs or symptoms of heart failure (HF) requiring treatment with at least one oral diuretic (any type) 4. Left ventricular ejection fraction (LVEF) ≥ 40% (by echocardiography at Screening) 5. New York Heart Association (NYHA) functional class (FC) II to III 6. Criterion removed (see footnote [1]) 7. Patients with HFpEF defined as either one of the following by echocardiography at Screening: a. Left atrial (LA) enlargement: i. Left atrial volume > 58 mL (male) / > 52 mL (female) *or* ii. Left atrial volume index ≥ 28 mL/m^2, *or* iii. LA area > 20 cm^2, *or* iv. LA diameter > 4.0 cm (male) / > 3.8 cm (female) b. Left ventricular septal thickness or posterior wall thickness ≥ 1.1 cm 8. Elevated NT-proBNP / BNP ≥ 200 / 60 pg/mL for subjects in sinus rhythm or ≥ 500 / 150 pg/mL for subjects with atrial fibrillation (AF) at any time within 3 months prior to Screening or at Screening 9. Pulmonary vascular disease or right ventricular (RV) dysfunction meeting <u>at least one</u> of the following for echocardiographic (at Screening) and/or right heart catheterization (RHC) parameters (at Screening or from any RHC previously performed): a. Echocardiographic peak TR velocity > 2.8 m/s *or* invasive mean pulmonary artery pressure ≥ 25 mmHg (RHC) *or* PASP > 40 mmHg **and** evidence of RV dysfunction (TAPSE < 17 mm *or* RV fractional area change < 35% *or* RV tissue Doppler s' velocity < 9.5 cm/s)

[1] Inclusion criterion 6 (HF hospitalization within 12 months prior to Screening or RHC within 6 months prior to Screening showing PAWP/LVEDP > 15mmHg) was removed in global protocol v.4.

	b. Diastolic Pulmonary Vascular Pressure Gradient (DPG) > 5 mmHg (RHC) c. Pulmonary vascular resistance (PVR) > 3 Wood Units (RHC) 10. A woman of childbearing potential must have a negative pre-treatment serum pregnancy test, agree to use reliable contraception from at least 30 days prior to Visit 2 up to at least 30 days after study treatment discontinuation, and agree to undertake monthly pregnancy tests from Screening up to at least 30 days after study treatment discontinuation.
EXCLUSION CRITERIA	**Disease-related** 1. Any prior valid measurement of LVEF < 40% **Cardiovascular comorbidities:** 2. Significant unrepaired structural valvular heart disease (i.e., greater than mild aortic or mitral stenosis, and greater than moderate aortic or mitral regurgitation) 3. Hypertrophic, restrictive, and infiltrative cardiomyopathies 4. Pericardial disease 5. Acute coronary syndrome, including ST-segment elevation myocardial infarction (STEMI), non-ST-segment elevation myocardial infarction (NSTEMI) or unstable coronary artery disease or has undergone coronary artery bypass graft (CABG) or percutaneous coronary intervention (PCI) within 3 months of screening. 6. Known indication for PCI or CABG 7. Uncontrolled heart rate (HR) from atrial fibrillation or atrial flutter (> 110 beats per minute) as assessed by ECG 8. History of serious life-threatening or hemodynamically significant arrhythmias, including symptomatic or sustained ventricular

	tachycardia or defibrillator shock within 12 month(s) of Screening 9. History of or anticipated heart transplant or anticipated/implanted ventricular assist device 10. Transient ischemic attack (TIA) or stroke within 3 months of Screening 11. Systolic blood pressure (SBP) ≥ 180 mmHg or diastolic blood pressure (DBP) ≥ 110 mmHg **Other causes of right heart failure not associated with left ventricular dysfunction:** 12. Significant parenchymal lung disease fulfilling any of the following: a. Forced expiratory volume in 1 second / forced vital capacity (FEV_1/FVC ratio) < 0.7 associated with FEV_1 < 50% of predicted value after bronchodilator administration in subjects with a known or suspected history of significant lung disease. b. Known moderate or severe restrictive lung disease, e.g., total lung capacity (TLC) < 60% (predicted) c. Clinical suspicion of diffuse interstitial fibrosis or alveolitis, unless excluded by high resolution computed tomography (CT) d. Clinical suspicion of pulmonary thromboembolism within 12 months prior to Screening, unless excluded by ventilation/perfusion (V/Q) scan or computed tomography angiography (CTA) 13. Criterion removed (see footnote[2]) 14. Clinically significant congenital heart disease that could be the cause of the patient's symptoms and signs of HF. **Criteria related to macitentan use:** 15. Administration of pulmonary arterial hypertension-specific therapy (i.e., endothelin

[2] Exclusion criterion 13 (BMI ≥ 45 kg/m^2) was removed in global protocol v.4.

	receptor antagonists, prostanoids, phosphodiesterase-5 [PDE-5] inhibitors, guanylate cyclase stimulators) within 1 month prior to Screening
	16. Hypotension, i.e., SBP < 90 mmHg or DBP < 50 mmHg
	17. Severe renal dysfunction with an estimated Glomerular Filtration Rate (eGFR) < 30 mL/min per 1.73 m^2 using the Modification of Diet in Renal Disease (MDRD) formula
	18. Known and documented severe hepatic impairment, e.g., Child-Pugh Class C
	19. Serum aspartate aminotransferase (AST) and/or alanine aminotransferase (ALT) ≥ 3 × the upper limit of normal at Screening
	20. Hemoglobin < 100 g/L (< 10 g/dl) at Screening
	21. Plan to become pregnant or lactating
	22. Treatment with strong cytochrome P-450 3A4 (CYP3A4) inducers such as rifabutin, rifampin, rifampicin, rifapentin, carbamazepine, phenobarbital, phenytoin, or St. John's wort within 1 month prior to Screening
	23. Treatment with strong CYP3A4 inhibitors such as ketoconazole, itraconazole, voriconazole, clarithromycin, telithromycin, nefazodone, ritonavir, and saquinavir within 1 month prior to Screening
	24. Criterion removed (see footnote[3])
	25. Known hypersensitivity to macitentan or drugs of the same class, or any of the excipients (e.g., soy lecithin, lactose)

[3] Exclusion criterion 24 (treatment with BCRP substrates) was removed in global protocol v.4.

	General criteria: 26. Planned or current treatment with another investigational treatment within 2 months prior to screening 27. Inadequate control of comorbidities according to current standards of care, as per judgment of the investigator 28. Any known factor or disease that might interfere with treatment compliance, study conduct, or interpretation of the results, such as drug or alcohol dependence or psychiatric disease 29. Known concomitant life-threatening disease with a life expectancy < 12 months.
RUN-IN FAILURE CRITERIA	A subject must be discontinued from the study in case any of the following criteria is fulfilled at any time during the run-in period: 1. Study treatment compliance < 80% 2. Central laboratory results showing a decrease in hemoglobin by > 50 g/L from Screening or hemoglobin < 80 g/L or need for transfusion 3. Significant fluid retention / worsening of HF as evidenced by one of the following: a. Administration of i.v. diuretics due to fluid retention b. Addition of high potency thiazide diuretic (metolazone, indapamide), ≥ 100% increase in loop diuretic to a total oral dose ≥ 120 mg of furosemide equivalents/day [see Section 4.5]. c. Increase in body weight by ≥ 5% or ≥ 5 kg from the value at the start of the corresponding run-in period (i.e., Visit 2 and Visit 4 for the placebo and macitentan run-in, respectively) due to fluid overload d. Worsening in NYHA FC e. Hospitalization for worsening of HF

	4. Any adverse events (AEs) that preclude continuation based on the investigator's judgment.
STUDY TREATMENTS	**Investigational treatment** Macitentan oral tablet, 10 mg once daily. **Comparator** Matching placebo, once daily.
AUXILIARY MEDICINAL PRODUCTS	All subjects must be on oral diuretic therapy (any type).
ENDPOINTS	**Primary efficacy endpoint** • Percent of baseline NT-proBNP assessed at Week 24 **Secondary efficacy endpoints** • Change from baseline to Week 24 in the clinical summary score (as assessed by the Kansas City Cardiomyopathy Questionnaire [KCCQ]) • Change from baseline to Week 24 in accelerometer-assessed proportion of time spent in light to vigorous physical activity based on a threshold of > 100 activity counts per minute • Time to worsening heart failure (WHF) event over 52 weeks. • **Other efficacy endpoints** Other efficacy endpoints are described in Section 6.1.3. **Safety endpoints** • All-cause death up to 30 days after study treatment discontinuation • Number of all-cause hospital admissions up to 30 days after study treatment discontinuation • Treatment-emergent AEs and serious adverse events (SAEs) up to 30 days after study treatment discontinuation • AEs leading to premature discontinuation of study treatment

	• Change in vital signs (systolic and diastolic arterial blood pressure and pulse rate) and body weight up to 30 days after study treatment discontinuation • Treatment-emergent marked laboratory abnormalities up to 30 days after study treatment discontinuation • Change from baseline in eGFR up to 30 days after study treatment discontinuation • Decrease from baseline in SBP of $\geq 5\%$ and SBP < 100 mmHg up to 30 days after study treatment discontinuation.
ASSESSMENTS	Refer to the schedule of assessments in Table 1, Table 2, Table 3, and Table 4.
STATISTICAL METHODOLOGY	**Analysis sets** The Screened Analysis Set includes all subjects who are screened and have a subject identification number. The placebo run-in Set includes all screened subjects who enter the placebo run-in and receive at least one dose of study treatment in single-blind placebo run-in. The macitentan run-in Set includes all subjects who have completed the placebo run-in, enter the macitentan run-in and receive at least one dose of study treatment in single-blind macitentan run-in. The Full Analysis Set (FAS) includes all subjects randomized to double-blind study treatment. The Per-protocol Analysis Set comprises all randomized subjects who received double-blind study treatment and who complied with the protocol sufficiently to allow a reliable assessment of the treatment effect on the primary efficacy endpoint. The Safety Set includes all subjects who received at least one dose of double-blind study treatment. **Primary efficacy variable** The primary efficacy variable is the percent of baseline in NT-proBNP at Week 24.

	For subjects without available NT-proBNP value at Week 24, the last available value observed before Week 24 will be carried forward and considered for the main analysis. Percent of baseline is calculated as the ratio of the Week 24 NT-proBNP value over baseline value, expressed in percentage. **Null and alternative hypotheses** Due to the underlying log-normal distribution of the primary efficacy variable, an analysis of covariance (ANCOVA) will be applied on log-transformed data. The null hypothesis is that macitentan and placebo effects on NT-proBNP are the same. The alternative hypothesis is that the effect of macitentan on NT-proBNP differs from placebo effect. The null hypothesis will be tested by a 2-sided Wald test with a 2-sided significance level of 0.10. **Primary statistical analysis** The main analysis will be performed on subjects of the FAS, according to the intent-to-treat principle. The main analysis of the primary efficacy variable will be carried out after log-transformation using ANCOVA, adjusting for the log of the ratio: baseline NT-proBNP over NT-proBNP value as per Interactive Response Technology system [IRT] with an overall type I error of 10% 2-sided. The treatment effect expressed as geometric means ratio and its associated 90% 2-sided confidence interval (CI) will be estimated based on the same model by inversely transforming, using the exponential function, the Least Squares Mean and 90% CI obtained in log scale. **Key secondary efficacy variables** The three key secondary efficacy variables will be tested using a hierarchical testing approach to address multiplicity concern. The change from baseline to Week 24 in KCCQ clinical summary score and accelerometer-assessed physical activity will be analyzed by means of an ANCOVA adjusting for the variable baseline value and for the log of

	the ratio: baseline NT-proBNP over NT-proBNP value as per IRT. Least Squares Estimates within each treatment group and between treatment groups will be displayed along with 90% 2-sided CIs and p-values. The analysis of time to first occurrence of WHF event will be explored using a proportional hazards Cox model adjusting for the log of the ratio: baseline NT-proBNP over NT-proBNP value as per IRT. Estimate of HR and its associated 90% CI and p-value will be displayed. Kaplan-Meier estimates will be calculated with 2-sided 90% CIs at relevant time points for each treatment group and displayed in both a graphical and a tabular form. **Safety variables** The number and percentage of: • All-cause deaths occurring up to 30 days after study treatment discontinuation • Subjects with at least one treatment-emergent AE occurring up to 30 days after study treatment discontinuation • Subjects with at least one SAE occurring up to 30 days after study treatment discontinuation • Subjects with at least one AE leading to premature discontinuation of study treatment will be tabulated by treatment group and by System Organ Class and Preferred Term. The number of recurrent all-cause hospital admissions will be estimated using a Negative binomial regression model. Marked laboratory abnormalities will be summarized for each laboratory variable by treatment group providing their incidence and frequency. For selected laboratory variables (erythrocytes, hemoglobin, hematocrit, AST, ALT, total bilirubin and alkaline phosphatase), absolute values and changes from baseline will be summarized over time up to 30 days after study treatment discontinuation. Vital signs, body weight and GFR will be described over time up to 30 days after study treatment discontinuation by means of random coefficient regression models.

	Decreases in SBP will be summarized by treatment group providing the number and percentage of subjects with, for at least one post-baseline assessment and up to 30 days after study treatment discontinuation, a percent change from baseline $\leq -5\%$ and an SBP < 100 mmHg at the time of assessment. **Sample size** A number of 300 randomized subjects is adequate to detect a geometric means ratio of 0.75 (macitentan over placebo, −0.288 in log scale) corresponding to a 25% improvement with a power of 80% and a type I error of 0.10 2-sided, when considering a standard deviation of 1 in log scale and using the Wald test. The critical value expressed as geometric means ratio is 0.83. Following closure of Screening on 27 December 2019, it is expected that approximately 140 subjects will be randomized. With the same assumptions as above for the reduced sample size of 140, the critical value expressed as geometric means ratio is 0.76.
STUDY COMMITTEES	An Independent Data Monitoring Committee (IDMC) has overall responsibility for safeguarding the interests of subjects by monitoring benefit-risk ratio and making appropriate recommendations based on all the reported data and thus ensuring that the study is being conducted with the highest scientific and ethical standards. The IDMC will be fully operational prior to enrollment of the first subject into the study. The composition and operation of the IDMC is described in the IDMC charter. Both a Scientific Advisory Board and a Steering Committee are involved in the study design and will be consulted prior to and during the study for relevant medical issues and study publications. An Independent Liver Safety Data Review Board (ILSDRB, an external expert committee of hepatologists) has been appointed to provide ongoing assessment and advice regarding serious hepatic AEs of special interest

	that require further evaluation during the study as per the ILSDRB charter.

Table 1 Visit and assessment schedule during the run-in period

PERIODS		SCREEN-ING	SINGLE-BLIND RUN-IN						Rando--mization [2]
			Placebo run-in			Macitentan run-in			
Duration		Up to 30 d	4 w			5 w			
VISITS[1]	Number	1	2	3	3a ☎	4	5	5a ☎	6
	Time		D1-R	W1-R (±2 d)	W-2R (±2 d)	W4-R (±4 d)	W5-R (±2 d)	W6-R (±2 d)	W9-R (±4 d)
Informed consent		X							
Eligibility		X	X[3]	X[3]	X[3]	X[3]	X[3]	X[3]	X[3]
Medical history		X							
Concomitant therapy		X	X	X	X	X	X	X	X
Physical examination		X	X	X		X	X		X
Vital signs (BP, HR)		X	X	X		X	X		X
Body weight, height[4]		X	X	X		X	X		X
Home body weight monitoring[5]			←			Daily			→
PFTs[6]		X							
12-lead ECG[7]		X				X			X
Echocardiography		X							
RHC (optional)		X							
Accelerometry			X[8]	X[9]			X[9]		
NYHA FC		X	X	X		X	X		X
WHF									X
PGA / KCCQ									X
Study trt dispensing			X			X	X		X
SAEs/AEs[10]		X	X	X	X	X	X	X	X
Signs indicative of a fluid retention /HF event[11]					X			X	

[1] Unscheduled visits may be performed at any time during the study. Body weight and vital signs (BP, HR) must be performed at each unscheduled visit. Other assessments are performed at the discretion of the investigator.

[2] Subjects who are not randomized (i.e., run-in failures) must perform a safety follow-up visit 30 (+5) days after intake of the last dose of study treatment.

[3] Run-in eligibility criteria defined in Section 4.5 of the protocol.

[4] Body weight is measured at each visit. Height is only measured at Screening.

[5] Subjects will be instructed to monitor their body weight at home on a daily basis during the run-in period, and to contact the study site if they notice any unusual weight increase (i.e., ≥ 2 kg / 4.4 lbs) after start of study treatment.

[6] Historical PFTs accepted if performed within 1 year prior to Screening and judged reliable by the investigator, provided the subject's pulmonary status remained unchanged during this time. Only necessary for subjects with a known or suspected history of significant lung disease.

[7] ECGs are read locally only.

[8] Accelerometry will be performed from Visit 2 until Visit 3.

[9] Accelerometry will be performed during 9 days following the visit.

[10] All AEs and SAEs that occur after signing the Informed Consent Form and up to 30 days after study treatment discontinuation must be reported [see also Section 9 of the protocol].

[11] During the phone calls, the investigator must inquire about signs and symptoms that may be indicative of a fluid retention- or HF event. Outcome of the phone call may trigger unscheduled visits per investigator's discretion.

AE = adverse event; BP = blood pressure; D = Day; d = days; ECG = electrocardiogram; FC = functional class; HR = heart rate; KCCQ = Kansas City Cardiomyopathy Questionnaire; NYHA = New York Heart Association; PGA = patient's global assessment;

PFT = pulmonary function test; R = Run-in; RHC = right heart catheterization; SAE = serious adverse event; trt = treatment; W = Week; w = weeks; WHF = worsening heart failure.

Table 2 Visit and assessment schedule during the double-blind treatment period

PERIODS		DOUBLE-BLIND TREATMENT								FOLLOW-UP
		Rando-mization	Core phase					Extension phase		
Duration			24 w					28 w		30 d
VISITS[1]	Number	6	7	8	9	10	11	12	13 (EOT)[8]	14 (EOS)
	Time	D 1 (±4 d)	W1 ☎[2] (±2 d)	W4 (±4 d)	W8 (±4 d)	W 16 (±8 d)	W 24 (±8 d)	W 36 (±8 d)	W 52 (±8 d)	30 (+5) d after last dose
Concomitant therapy		X	X	X	X	X	X	X	X	X
Physical examination		X		X	X	X	X	X	X	X
Vital signs (BP, HR)		X		X	X	X	X	X	X	X
Body weight		X		X	X	X	X	X	X	X
Home body weight monitoring[3]		←			Weekly				→	
12-lead ECG[4]		X		X		X	X		X	
Echocardiography							X		X	
Accelerometry[5]				X		X	X			
NYHA FC		X		X	X	X	X	X	X	X
WHF		X		X	X	X	X	X	X	X
PGA / KCCQ		X			X	X	X	X	X	
Study trt dispensing		X		X	X	X	X	X		
SAEs/AEs[6]		X	X	X	X	X	X	X	X	X

[1] Unscheduled visits may be performed at any time during the study. Body weight and vital signs (BP, HR) must be performed at each unscheduled visit. Other assessments are performed at the discretion of the investigator.
[2] Telephone call only.
[3] Subjects will be instructed to monitor their body weight at home on a weekly basis during the double-blind treatment period, and to contact the study site if they notice any unusual weight increase (i.e., ≥ 2 kg / 4.4 lbs) after start of study treatment.
[4] ECGs are read locally only.
[5] Accelerometry will be performed during 9 days following Visits 8 and 10.
[6] All AEs and SAEs that occur after signing the Informed Consent Form and up to 30 days after study treatment discontinuation must be reported [see also Section 9 of the protocol].
[7] Under global protocol Version 6, accelerometry will be performed during the 9 days prior to Visit 11 (Week 24 visit).
[8] Subjects who have passed Week 24 at the time of approval of global protocol Version 6 will have their EOT visit within 60 days but no later than Week 52.

AE = adverse event; BP = blood pressure; D = Day; d = days; ECG = electrocardiogram; EOS = end-of-study; EOT = end-of-treatment; FC = functional class; HR = heart rate; KCCQ = Kansas City Cardiomyopathy Questionnaire; NYHA = New York Heart Association; PGA = patient's global assessment; PFT = pulmonary function test; R = Run-in; RHC = right heart catheterization; SAE = serious adverse event; trt = treatment; W = Week; w = weeks; WHF = worsening heart failure; V = Visit.

Table 3 Laboratory assessments

PERIODS		SCREENING	SINGLE-BLIND RUN-IN				DOUBLE-BLIND TREATMENT							FOLLOW-UP	Anytime
			Placebo run-in		Macitentan run-in		Core phase					Extension phase			
Duration		Up to 30 d	4 w		5 w		24 w					28 w		30 d	NA
VISITS	Number	1	2	3	4	5	6	8	9	10	11	12	13 (EOT)[8]	14 (EOS)	Fluid retention/ WHF event (any visit)
	Time		D1-R	W1-R (±2 d)	W4-R (+/−4 d)	W5-R (±2 d)	W9-R / D 1 (±4 d)	W4 (±4 d)	W8, (±4 d)	W 16 (±8 d)	W 24 (±8 d)	W 36 (±8 d)	W 52 (±8 d)	30 (+5) d after last dose	
Central laboratory tests[1]		X[2]			X		X	X	X[3]	X[3]	X[3]	X[3]	X	X	
Local laboratory test							X[4]								
Serum Pregnancy test[5]		X			X		X	X	X	X	X	X	X	X	
Urine Pregnancy test[5]									W 12	W20		W 28, 32, 40, 44, 48[6]			
NT-proBNP		X			X		X	X		X	X		X		X
MR-proANP					X		X				X		X		X
Biomarker					X		X				X		X		X[7]

Note: Table does not display phone calls.

[1] Hematology and clinical chemistry.

[2] Laboratory test at Screening must be performed in fasted state. Includes measurement of glucose, lipid profile and TSH levels.

[3] Monthly AST/ALT monitoring recommended. Local laboratory can be used.

[4] Hemoglobin to be measured locally at randomization to confirm run-in failure criteria not met.

[5] For women of childbearing potential only.

[6] Urine pregnancy test will be performed by the subject. The site will follow-up on the results with a telephone call.

[7] For patients who gave their consent for this blood sample only. To be drawn at the time of a fluid retention - or WHF event (any visit, including unscheduled, as applicable).

[8] Applicable to subjects who have completed Week 24 prior to global protocol Version 6 approval.

ALT = alanine aminotransferase; AST = aspartate aminotransferase; D = Day; d = days; EOS = end-of-study; EOT = end-of-treatment; MR-proANP = mid-regional pro-atrial natriuretic peptide; NA = not applicable; NT-proBNP = n-terminal pro-brain natriuretic peptide; R = Run-in; TSH = thyroid stimulating hormone; W = Week; w = weeks; WHF = worsening heart failure; V = Visit.

Table 4 Visit and assessment schedule for subjects entering the PTOP

PERIODS	Name	POST-TREATMENT OBSERVATION PERIOD (PTOP)			
	Duration	Up to 52 Weeks			
VISITS	Number	PTOP1	PTOP2	PTOP3[7]	PTOP4[7]
	Name	Safety follow-up			EOS
	Time	30 (+5) d after last dose[1]	Week 24[2] (±14 d)	Week 36[2] (± 14 d)	Week 52[2] (± 14 d)
Concomitant therapy		X	X	X	X
Physical examination		X	X	X	X
Vital signs (BP, HR)		X	X	X	X
Body weight		X	X	X	X
12-lead ECG			X[3]		X
Echocardiography			X		X
Central laboratory tests		X[4]			
NT-proBNP / MR-proANP / Biomarker			X		X
NYHA functional class		X	X	X	X
WHF		X	X	X	X
PGA			X	X	X
KCCQ			X	X	X
Accelerometry			X[6]		
AEs		X			
SAEs		X	X	X	X

[1] If the safety follow-up visit falls within the time-window any of the PTOP visits, the visits can be combined.
[2] From randomization.
[3] ECGs are read locally only.
[4] Hematology and clinical chemistry, including serum pregnancy test for women of childbearing potential.

Appendix B Diuretics identified from Standard drug grouping*

	Ingredient / Diuretics*
High ceiling Sulfonamides (LOOP)	furosemide
	bumetanide
	piretanide
	torasemide
	etacrynic acid
	tienilic acid
	azosemide
	etozolin
Low ceiling, thiazides (THIAZIDES)	bendroflumethiazide
	hydroflumethiazide
	hydrochlorothiazide
	chlorothiazide
	polythiazide
	trichlormethiazide
	cyclopenthiazide
	methyclothiazide
	cyclothiazide
	mebutizide
	epitizide
	altizide
	bemetizide
	benzthiazide
	butizide
	benzylhydrochlorothiazide
Low ceiling, excl. thiazides	quinethazone
	clopamide
	chlortalidone
	mefruside
	clofenamide
	metolazone
	meticrane
	xipamide
	indapamide
	clorexolone
	fenquizone
	cicletanine
Potassium sparing	spironolactone
	potassium canrenoate
	canrenoic acid
	canrenone
	finerenone
	eplerenone
	amiloride

	triamterene
Vasopressin antagonists	tolvaptan
	conivaptan
	mozavaptan
	ribuvaptan
	lixivaptan
	satavaptan
Carbonic anhydrase inhibitors	acetazolamide
	chlorazanil
	demeclocycline
	diurgin
	esaxerenone
	ethiazide
	flumethiazide
	hydrobentizide
	methalthiazide
	metipamid
	paraflutizide
	penflutizide
	rolofylline
	teclothiazide
	tiamizide
	tripamide

* Identified diuretics starting with ATC code = C03

Appendix C The Kansas City Cardiomyopathy Questionnaire Scoring Instructions

There are 10 summary scores within the KCCQ, which are calculated as follows:

1. Physical limitations score

After coding the questions 1a-f collected in CRF as follows:

- Extremely limited = 1
- Quite a bit limited = 2
- Moderately limited = 3
- Slightly limited = 4
- Not at all limited = 5
- Limited for other reasons or did not do = <missing value>

If at least three of questions 1a-f are not missing:

Physical Limitation Score = 100*[(Mean of Questions 1a-f answered) -1]/4

where the mean is calculated on non-missing values.

2. Symptom Stability

Code the response to Question 2 as follows:

- Much worse = 1
- Slightly worse = 2
- Not changed = 3
- Slightly better = 4
- Much better = 5
- I've had no symptoms over the last 2 weeks = 3

If Question 2 is not missing, then compute

Symptom Stability Score = 100*[(Question 2) – 1]/4

3. Symptom Frequency score

After coding the questions 3, 5, 7 and 9 collected in CRF as follows:

<u>Question 3</u>

- Every morning = 1
- 3 or more times a week but not every day = 2
- 1-2 times a week = 3
- Less than once a week = 4
- Never over the past 2 weeks = 5

Questions 5 and 7

- All of the time = 1
- Several times a day = 2
- At least once a day = 3
- 3 or more times a week but not every day = 4
- 1-2 times a week = 5
- Less than once a week = 6
- Never over the past 2 weeks = 7

Question 9

- Every night = 1
- 3 or more times a week but not every day = 2
- 1-2 times a week = 3
- Less than once a week = 4
- Never over the past 2 weeks = 5

If at least two of Questions 3, 5, 7 and 9 are not missing, then compute:

- S3 = [(Question 3) – 1]/4
- S5 = [(Question 5) – 1]/6
- S7 = [(Question 7) – 1]/6
- S9 = [(Question 9) – 1]/4

Symptom Frequency Score = 100*(mean of S3, S5, S7 and S9)

4. Symptom Burden score

Code responses to each of questions 4, 6, and 8 as follows:

- Extremely bothersome = 1
- Quite a bit bothersome = 2
- Moderately bothersome = 3
- Slightly bothersome = 4
- Not at all bothersome = 5
- I've had no swelling/fatigue/shortness of breath = 5

If at least one of Questions 4, 6 and 8 is not missing, then compute:

Symptom Burden Score = 100*[(mean of Questions 4, 6 and 8)-1]/4

5. Total Symptom score

It is the mean of the following available summary scores:

- **Symptom Frequency score**
- **Symptom Burden score**

6. Self-Efficacy

Code responses to Questions 10 and 11 as follows:
Question 10

- Not at all sure = 1
- Not very sure = 2
- Somewhat sure = 3
- Mostly sure = 4
- Completely sure = 5

Question 11

- Do not understand at all = 1
- Do not understand very well = 2
- Somewhat understand = 3
- Mostly understand = 4
- Completely understand = 5

If at least one of Questions 10 and 11 is not missing, then compute:
Self-Efficacy Score = 100*[(mean of Questions 10 and 11 actually answered) – 1]/4

7. Quality of Life score

Code responses to the questions 12, 13 and 14 as follows:

Question 12

- It has extremely limited my enjoyment of life = 1
- It has limited my enjoyment of life quite a bit = 2
- It has moderately limited my enjoyment of life = 3
- It has slightly limited my enjoyment of life = 4
- It has not limited my enjoyment of life at all = 5

Question 13

- Not at all satisfied = 1
- Mostly dissatisfied = 2
- Somewhat satisfied = 3
- Mostly satisfied = 4
- Completely satisfied = 5

Question 14

- I felt that way all of the time = 1
- I felt that way most of the time = 2
- I occasionally felt that way = 3
- I rarely felt that way = 4
- I never felt that way = 5

If at least one of Questions 12, 13 and 14 is not missing, then compute:

Quality of Life Score = 100*[(mean of Questions 12, 13 and 14)-1]/4

8. Social Limitation score

After coding the questions 15a-d as follows:

- Severely limited = 1

- Limited quite a bit = 2
- Moderately limited = 3
- Slightly limited = 4
- Did not limit at all = 5
- Does not apply or did not do for other reasons = <missing value>

If at least two of Questions 15a-d is not missing, then compute:

Social Limitation Score = 100*[(mean of Questions 15a-d)-1]/4

9. Overall Summary score

It is the mean of the following available summary scores:

- **Physical limitations**
- **Total Symptom Score**
- **Quality of life Score**
- **Social Limitation Score**

10. Clinical Summary score

It is the mean of the following available summary scores

- **Physical limitations score**
- **Total Symptom score**

Appendix D PTs and SOCs for identification of HAESI

MedDRA PTs

Acute graft versus host disease in liver

Acute hepatic failure

Acute on chronic liver failure

Alanine aminotransferase

Alanine aminotransferase abnormal

Alanine aminotransferase increased

Allergic hepatitis

Alloimmune hepatitis

Aspartate aminotransferase

Aspartate aminotransferase abnormal

Aspartate aminotransferase increased

Asterixis

Autoimmune hepatitis

Biliary cirrhosis

Bilirubin conjugated

Bilirubin conjugated abnormal

Bilirubin conjugated increased

Blood bilirubin

Blood bilirubin abnormal

Blood bilirubin increased

Cardiohepatic syndrome

Cholestatic liver injury

Chronic graft versus host disease in liver

Chronic hepatic failure

Chronic hepatitis

Cirrhosis alcoholic

Coma hepatic

Cryptogenic cirrhosis

Drug-induced liver injury

Graft versus host disease in liver

Hepatic cirrhosis

Hepatic encephalopathy

Hepatic encephalopathy prophylaxis

Hepatic enzyme

Hepatic enzyme abnormal

Hepatic enzyme increased

Hepatic failure

Hepatic fibrosis

Hepatic function abnormal

Hepatic infiltration eosinophilic

Hepatic necrosis

Hepatic steato-fibrosis

Hepatitis

Hepatitis acute

Hepatitis alcoholic

Hepatitis cholestatic

Hepatitis chronic active

Hepatitis chronic persistent

Hepatitis fulminant

Hepatitis toxic

Hepatocellular injury

Hepatopulmonary syndrome

Hepatorenal failure

Hepatorenal syndrome

Hepatotoxicity

Hyperbilirubinaemia

Hyperbilirubinaemia

Hypertransaminasaemia

Increased liver stiffness

Ischaemic hepatitis

Jaundice

Jaundice cholestatic

Jaundice hepatocellular

Liver function test

Liver function test abnormal

Liver function test increased

Liver injury

Liver operation

Liver transplant

Lupoid hepatic cirrhosis

Lupus hepatitis

Minimal hepatic encephalopathy

Mixed liver injury

Multivisceral transplantation

Nodular regenerative hyperplasia

Non-alcoholic steatohepatitis

Ocular icterus

Portal fibrosis

Primary biliary cholangitis

Pseudocirrhosis

Radiation hepatitis

Renal and liver transplant

Reye's syndrome

Reynold's syndrome

Steatohepatitis

Subacute hepatic failure

Transaminases

Transaminases abnormal

Transaminases increased

Withdrawal hepatitis

Yellow skin

SOC terms

Hepatobiliary disorders

HLGT terms

Hepatobiliary investigations