A randomized parallel-group, placebo-controlled, double-blind, multicenter dose finding phase II trial exploring the pharmacodynamic effects, safety and tolerability, and pharmacokinetics of four dose regimens of the oral sGC stimulator BAY 1021189 over 12 weeks in patients with worsening heart failure and reduced ejection fraction (HFrEF) - SOLuble guanylate Cyclase stimulatoR in heArT failurE patientS with REDUCED EF (SOCRATES-REDUCED)

Bayer study drug  BAY 1021189 / Vericiguat / sGC stimulator
Study purpose:  dose finding, safety, efficacy, pharmacokinetics
Clinical study phase:  IIb  Date:  01 July 2015
Study No.:  15371  Version:  3.0
Author:  Katharina Müller

Confidential
The information provided in this document is strictly confidential and is intended solely for the guidance of the clinical investigation. Reproduction or disclosure of this document, whether in part or in full, to parties not associated with the clinical investigation or its use for any other purpose without the prior written consent of the sponsor is not permitted.

Throughout this document, symbols indicating proprietary names (®, TM) are not displayed. Hence, the appearance of product names without these symbols does not imply that these names are not protected.

This Statistical Analysis Plan is produced on a word-processing system and bears no signatures. The approval of the Statistical Analysis Plan is documented in a separate Signature Document.
# Table of Contents

1. Introduction ............................................................................................................... 6  
2. Study Objectives ......................................................................................................... 6  
3. Study Design ................................................................................................................ 7  
4. General Statistical Considerations ........................................................................... 9  
   4.1 General Principles ................................................................................................ 9  
   4.2 Handling of Dropouts ....................................................................................... 9  
   4.3 Handling of Missing Data ................................................................................. 10  
   4.4 Interim Analyses and Data Monitoring ............................................................ 10  
   4.5 Data Rules ......................................................................................................... 11  
      4.5.1 Compliance ............................................................................................... 11  
      4.5.2 Baseline and Change from Baseline ......................................................... 11  
      4.5.3 Repeated Measures .................................................................................. 12  
      4.5.4 Laboratory Data Handling ...................................................................... 12  
      4.5.5 Subgroup analyses ................................................................................... 12  
      4.5.6 Maintenance Dose for Analysis of Actual Dose taken ............................ 14  
      4.5.7 Exposure Groups ..................................................................................... 14  
   4.6 Validity Review .................................................................................................... 14  
5. Analysis Sets ............................................................................................................... 15  
   5.1 Assignment of analysis sets .............................................................................. 15  
      5.1.1 Full analysis set (FAS) ............................................................................ 15  
      5.1.2 Safety analysis set (SAF) ......................................................................... 15  
      5.1.3 Per protocol set (PPS) .............................................................................. 15  
      5.1.4 Pharmacokinetics analysis set (PKS) ....................................................... 15  
6. Statistical Methodology ............................................................................................ 16  
   6.1 Population characteristics ............................................................................... 16  
      6.1.1 Disposition ............................................................................................... 16  
      6.1.2 Demography and Baseline Characteristics .............................................. 16  
      6.1.3 Medical History ...................................................................................... 17  
      6.1.4 Concomitant Medication ....................................................................... 18  
      6.1.5 Treatment Duration and Exposure ......................................................... 22  
      6.1.6 Treatment Compliance ............................................................................ 22  
      6.1.7 Titration Status ....................................................................................... 22  
   6.2 Efficacy ................................................................................................................ 22  
      6.2.1 Primary Efficacy Variable ...................................................................... 22  
      6.2.2 Analyses of Exploratory Efficacy Variables ........................................... 29  
   6.3 Pharmacokinetics / pharmacodynamics ............................................................ 35  
   6.4 Safety ................................................................................................................... 35  
      6.4.1 Adverse events ....................................................................................... 35  
      6.4.2 Further safety parameters ...................................................................... 37  
7. Document history and changes in the planned statistical analysis ......................... 38  
   7.1 Document history .............................................................................................. 38  
   7.2 Changes in the planned statistical analysis ....................................................... 39  

Reference Number: BPD-SOP-060  
Supplement Version: 5
7.2.1 Changes from the protocol ................................................................. 39
7.2.2 Changes from SAP version 1.0 to version 2.0................................. 39
7.2.3 Changes from SAP version 2.0 to version 3.0................................. 40

8. References .......................................................................................... 40

9. Appendix ............................................................................................ 42

9.1 KCCQ Scoring ................................................................................. 42

9.1.1 Physical Limitation ....................................................................... 42
9.1.2 Symptom Stability ........................................................................ 42
9.1.3 Symptom Frequency .................................................................... 43
9.1.4 Symptom Burden ......................................................................... 43
9.1.5 Self-Efficacy ................................................................................. 44
9.1.6 Quality of Life ............................................................................... 44
9.1.7 Social Limitation .......................................................................... 45
9.1.8 Total Symptom Score ................................................................... 45
9.1.9 Overall Summary Score ............................................................... 45
9.1.10 Clinical Summary Score ............................................................. 46

9.2 EQ-5D-3L Scoring ........................................................................... 46

9.2.1 UK Time Trade-Off (TTO) value set .......................................... 46
9.2.2 US TTO value set ......................................................................... 47

9.3 Composite congestion score ............................................................. 48

List of Tables

Table 6-1: Equivalence doses for different diuretics ................................. 19
Table 6-2: Recommended maximum doses ............................................ 21
Table 6-3: Definition of observation period and censoring rules ............ 32
Table 6-4: Scoring of Signs and Symptoms of Congestion .................... 34
Table 9-1: UK TTO terms for individual dimensions .............................. 46
Table 9-2: UK TTO terms for special variables ..................................... 46
Table 9-3: US TTO terms for individual dimensions .............................. 47
Table 9-4: US TTO terms for special variables ..................................... 48

List of Figures

Figure 3-1: Study design ...................................................................... 8
Abbreviations

ACE Angiotensin converting enzyme
AE Adverse event
ALT Alanine aminotransferase
AN(C)OVA Analysis of (Co-)Variance
ARB Angiotensin receptor blocker
ATC Anatomical Therapeutic Chemical
BMI Body mass index
CAD Coronary artery disease
cGMP Cyclic guanosine monophosphate
CKD Chronic kidney disease
COPD Chronic obstructive pulmonary disease
CRT(-D) Cardiac Resynchronization Therapy (Defibrillator)
CSR Clinical Study Report
CV Cardiovascular
DSMC Data Safety Monitoring Committee
DBP Diastolic blood pressure
ECG Electrocardiogram
eCRF Electronic Case Report Form
EF Ejection fraction
eGFR Estimated glomerular filtration rate
e.g. exempli gratia, for example
FAS Full analysis set
GCP Good Clinical Practice
HEOR Health Economics, Outcomes & Reimbursement
HF Heart failure
HFrEF Heart failure with reduced ejection fraction
HR Heart rate
ICD Implantable Cardiac Defibrillator
ICH International Committee on Harmonization
i.e. id est, that is
IV Intravenous
KCCQ Kansas City Cardiomyopathy Questionnaire
LAV Left atrial volume
LLOQ Lower limit of quantification
LOCF Last observation carried forward
LVEF Left ventricular ejection fraction
MACE Major adverse cardiovascular event
MedDRA Medical Dictionary for Regulatory Activities
MRA Mineralocorticoid receptor antagonist
NO Nitric oxide
NT-proBNP N-terminal pro-brain natriuretic peptide
NYHA New York Heart Association
o.d. omni die, once daily
PD Pharmacodynamics
PDD Protocol deviation document
PDE5i Phosphodiesterase type 5 inhibitor
PK Pharmacokinetics
PKS Pharmacokinetic analysis set
PPS Per protocol set
PT Preferred term
SAC Statistical Analysis Center
SAE Serious adverse event
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAF</td>
<td>Safety analysis set</td>
</tr>
<tr>
<td>SAP</td>
<td>Statistical Analysis Plan</td>
</tr>
<tr>
<td>SBP</td>
<td>Systolic blood pressure</td>
</tr>
<tr>
<td>sGC</td>
<td>Soluble guanylate cyclase</td>
</tr>
<tr>
<td>SMQ</td>
<td>Standard MedDRA query</td>
</tr>
<tr>
<td>SOC</td>
<td>System organ class</td>
</tr>
<tr>
<td>TEAE</td>
<td>Treatment emergent adverse event</td>
</tr>
<tr>
<td>TLF</td>
<td>Tables, listings, and figures</td>
</tr>
<tr>
<td>ULN</td>
<td>Upper limit of normal</td>
</tr>
<tr>
<td>VAS</td>
<td>Visual analogue scale</td>
</tr>
<tr>
<td>WCHF</td>
<td>Worsening chronic heart failure</td>
</tr>
<tr>
<td>WHO-DD</td>
<td>World Health Organization Drug Dictionary</td>
</tr>
</tbody>
</table>
1. Introduction

Heart failure (HF) is the leading cause of cardiovascular morbidity and mortality constituting the major public health problem worldwide. HF is associated with a wide spectrum of left ventricular (LV) functional abnormalities, ranging from patients with normal LV size and preserved ejection fraction (EF) to those with severe dilatation and/or markedly reduced EF. Previously, LVEF has been widely used to define systolic function, assess prognosis, and select patients for therapeutic interventions [1]. This concept resulted in the entity systolic HF, or more recently designated as HF with reduced EF (HFrEF).

Patients with worsening symptoms and/or signs of HF who require hospitalization represent an important subpopulation of patients with chronic HF. In these patients, the requirement of hospitalization for HF indicates more severe clinical presentation, and substantially worsened prognosis compared to outpatients not hospitalized for HF [2]. While angiotensin converting enzyme (ACE) inhibitors, angiotensin II receptor blockers (ARB), β-blockers, and mineralocorticoid-receptor antagonists (MRA) were shown to improve outcomes in chronic heart failure, the prognosis of patients who required hospitalization for HF remains unfavorable [3]. However, the number of annual hospitalizations for heart failure (HF) and the mortality rates among patients hospitalized for HF remains unacceptably high despite these treatments.

The nitric oxide-soluble guanylate cyclase-cyclic guanosine monophosphate (NO-sGC-cGMP) pathway is a relevant mechanism in HF that is not targeted by neurohumoral antagonists. A novel class of sGC stimulators directly stimulates the NO receptor sGC with a dual mode of action. They sensitize sGC to endogenous NO by stabilizing the NO-sGC binding and also directly stimulate sGC via a different binding site, independently of NO.

Previous oral sGC stimulators have shown improvement in systematic and pulmonary hemodynamics with systolic HF and elevated pulmonary capillary wedge pressure and mean pulmonary artery pressure (BAY 60-4552) and improvement in cardiac index, pulmonary vascular resistance, systemic vascular resistance, and quality of life in subjects with left ventricular systolic dysfunction associated with pulmonary hypertension (BAY 63-2521).

In this study (15371), four different dose regimens and one placebo arm will be compared to find the optimal dose of vericiguat (BAY 1021189) for a phase III trial in subjects with worsening chronic HF with reduced EF.

This statistical analysis plan (SAP) v3.0 is an amendment to the original SAP v1.0 dated 24 February 2015 and SAP amendment v2.0 dated 12 June 2015 and is based on the study protocol (Version 1.0 dated 16 July 2013; Version 2.0 dated 06 July 2014) and describes the final analysis of study 15371.

2. Study Objectives

Objective of the study is to find the optimal dose of the oral sGC stimulator vericiguat for Phase III that can be given in addition to standard therapy for heart failure (HF) with reduced ejection fraction (HFrEF) by characterizing the safety, tolerability, pharmacodynamic effects,
and pharmacokinetics, and detecting a significant dose-response relationship in the primary endpoint change in N-terminal pro-brain natriuretic peptide (NT-ProBNP) at 12 weeks in patients with worsening chronic HFrEF.

3. **Study Design**

This is a prospective, randomized, placebo-controlled, double blind, 5 parallel arms, global multi-center dose finding phase II trial in subjects stabilized after hospitalization or IV diuretic treatment for worsening chronic HF with reduced ejection fraction (EF).

**Screening**

Screening may be initiated any time after admission of a patient to the hospital for HF. As an equivalent to hospitalization, use of IV diuretic treatment for HF without hospitalization is also accepted as indicator of worsening chronic HF.

Candidates for screening are required to have a history of chronic HF (New York Heart Association [NYHA] functional class II – IV) and treatment with standard HF therapy for ≥30 days prior to hospitalization or equivalent. Furthermore, suitable candidates need to fulfill the following criteria:

- history of LVEF <45%
- symptoms and signs of congestion [clinical or radiographic signs in routine chest x-ray demonstrating pleural effusion, pulmonary congestion (redistribution, interstitial or alveolar edema), or cardiomegaly] at hospitalization or equivalent
- NT-ProBNP or BNP above the cut-off given in the inclusion criteria at any time upon hospitalization or equivalent

These three screening criteria are collected from available clinical routine and do not represent study procedures.

**Randomization**

Patients who meet the screening criteria are eligible for randomization upon clinical stabilization. Randomization should occur when the LVEF <45% is confirmed in local baseline echocardiography by the investigator who will determine if the inclusion/exclusion criteria are met. Assuming a 20% drop-out and invalidity rate, approximately 410 subjects will need to be randomized to one of the 5 equally sized treatment groups.

**Treatment**

Patients will take vericiguat and/or placebo once daily in the morning with food. The starting dose will be 1.25 mg or 2.5 mg of vericiguat or matching placebo.

Patients will be randomized to either

1. Placebo and sham titration
2. 1.25 mg and sham titration
3. 2.5 mg and sham titration
4. 2.5 mg, which will be up-titrated to 5 mg after 14 days, and sham titration after 28 days
5. 2.5 mg, which will be up-titrated to 5 mg after 14 days and to 10 mg after 28 days

Titration steps at 14 and 28 days include sham titrations in groups 1, 2, and 3, and after 28 days in group 4. The maximal dose of vericiguat is 10 mg o.d. in the 10 mg treatment group, 5 mg o.d. in the 5 mg treatment group, 2.5 mg o.d. in the 2.5 mg group, and 1.25 mg o.d. in the 1.25 mg group (Figure 3-1).

Titration in all groups will depend on systolic blood pressure (SBP) before intake of the dose:
- Double dose: If SBP ≥100 mmHg
- Maintain dose: If SBP between 90 and <100 mmHg
- Halve the dose: if SBP <90 mmHg without symptoms of hypotension

Dose halving is possible at any time if the investigator feels this is justified for safety reasons. If a second dose halving is required this will result in discontinuation of study drug treatment, and down-titration attempts from a dose of 1.25 mg will result in discontinuation.

**End-of-treatment**

End of treatment is reached at visit 5 after 12 weeks of treatment, when echocardiography, laboratory, clinical, and safety read-outs are collected.
Follow-up
A safety follow-up visit is scheduled at 30±5 days after end of treatment or discontinuation of study drug.
In addition, the vital status will be collected at 114±5 days after randomization for all subjects who prematurely discontinue study drug.

4. General Statistical Considerations

4.1 General Principles
All variables will be analyzed by descriptive statistical methods. The number of data available, number of subjects with missing information, mean, standard deviation, minimum, median, quartiles, and maximum will be calculated for continuous data. Frequency tables will be generated for categorical data.
All subjects will be analyzed according to the assigned treatment group if not specified otherwise. Each treatment group will comprise all subjects within one randomized titration scheme regardless of the individual subject’s actual titration, i.e. in case a subject could not be uptitrated or was downtitrated during the study.
The three highest dose arms 2.5 mg, 5 mg, and 10 mg will be pooled for the primary efficacy analysis. Selected tables will also be given by this pooled treatment group where indicated.
A measurement is considered to be "on-treatment", if it was taken during treatment or up to 5 days after end of treatment with study medication.
If measurements are to be analyzed at a certain visit, the actual visit date is used, if not otherwise specified.
The statistical evaluation will be performed using the software package SAS release 9.2 or higher (SAS Institute Inc., Cary, NC, USA)[4].
The analysis will be based on the Global Standard Tables (Version 2.0) and the Clinical Pharmacology Standards (CLIPS) (Version 1.2) where appropriate.

4.2 Handling of Dropouts
A subject who discontinues study participation prematurely for any reason is defined as a "dropout" if the subject has been randomized (even if no study drug has been taken). Dropouts will not be replaced. Data from subjects who prematurely terminated the study will be used to the maximum extent possible. The number and percentage of subjects who did not complete the treatment phase or the follow-up phase, as well as the primary reason for discontinuation, will be displayed by treatment group and overall as described in Section 6.1.1.
All subjects that prematurely stopped study medication for any reason will be followed-up for vital status at 114±5 days after randomization.

4.3 Handling of Missing Data

All missing or partial data will be presented in the subject data listings as they are recorded on the electronic Case Report Form (eCRF).

For the primary analysis of the primary endpoint, a subject needs to have a valid (non-missing) NT-proBNP measurement at baseline and at visit 5 (week 12). Subjects who drop-out during the course of the study and hence do not have an NT-proBNP measurement at week 12 will not be considered for the primary analysis (i.e., per protocol analysis on completers).

Sensitivity analyses using imputation methods for subjects with missing NT-proBNP measurement at visit 5 will be described in Section 6.2.1.3.1.

In case of (partially) missing end dates for study medication intake the following approach will be applied to impute end of study medication intake:

- In case the subject stopped treatment due to death: impute partially missing end dates by the 'worst study medication end date' defined as the maximal possible date (i.e. last month of the year, and last day of the month, respectively). Take the minimum of death date and, if available, imputed worst study medication end date as the imputed study medication end date.

- In case the subject completed treatment or stopped due to reasons other than death: impute partially missing end dates by the 'minimum study medication end date' defined as the minimal possible date (i.e. first day of month and first month of year, respectively). Take the maximum of the last available study medication intake date (i.e. largest non-missing study medication start or stop date) and, if available, imputed minimum study medication end date as the imputed study medication end date.

In case of (partially) missing start dates for clinical events, the earliest possible date after first drug intake will be imputed.

4.4 Interim Analyses and Data Monitoring

An independent Data Safety Monitoring Committee (DSMC) has been established for this study together with the parallel phase IIb study in patients with preserved ejection fraction (BAY 1021189/Study 15829), who will review safety and tolerability regularly during the course of the two studies. Members of the Committee will receive unblinded safety data. The involvement of an external Statistical Analysis Center in this process will ensure that unblinded information is not available for third parties. The study team will be kept blinded in this process.

No formal interim analysis is planned. There are no predefined stopping conditions for the ongoing safety monitoring of this trial. However, the DSMC may recommend termination, temporary suspension of the study, and intervention of treatment arm or modification of the
study. Details on DSMC responsibilities and decisions, as well as a detailed plan for DSMC meetings is described in the DSMC charter (Version 1.0, dated 10 December 2013). The statistical analyses for the DSMC meetings are detailed under a separate DSMC SAP (Version 1.0, dated 4 March 2014; Version 2.0, dated 29 August 2014) and DSMC tables, listings, and figures (TLF) specification (Version 1.0, dated 4 March 2014; Version 2.0, dated 29 August 2014).

If an update of the final statistical analysis is required due to study modifications recommended by the DSMC, an amendment or supplement to this SAP will be provided.

In case of emergency or any finding that requires unblinding, the investigator will be able to break the blind for an individual subject via IxRS according to the unblinding procedure implemented. This will allow breaking the blind for an individual subject without impairing the study as a whole.

4.5 Data Rules

General data rules are described in this section, further data rules for specific parameters or analyses are specified in the respective subsections of Section 6.

4.5.1 Compliance

The compliance is calculated as percent of planned tablet intake:

\[
100 \times \frac{\text{Number of taken tablets}}{\text{Number of planned tablets}}.
\]

The number of planned tablets is calculated as: Treatment duration (days) * Number of planned tablets per day. The number of planned tablets per day is two.

To derive treatment duration, the start and stop dates of tablet intake are derived from the collected study drug exposure (dataset EX). Times with temporary discontinuation from study drug will be excluded from the treatment duration. The number of taken tablets is derived from the drug accountability (dataset DA).

For subjects who withdraw prematurely from the study drug, compliance will be calculated up to the time of last dose.

4.5.2 Baseline and Change from Baseline

Baseline is defined as the last non-missing measurement before first intake of study treatment. Generally, this would be the measurements taken on visit 1 before randomization. In case the last observation available prior to treatment is from the screening visit, this would be used as the baseline value, except for data that need to be evaluated centrally on visit 1, like biomarkers, laboratory values and echo parameters. These parameters are only measured locally at screening and will not be used for imputation of baseline value. If more than one measurement was planned for a scheduled time point (i.e. blood pressure measurements and
heart rate), the average value of the last set of measurements per time point prior to treatment will be used as the baseline value.

Change from baseline for vital signs or laboratory parameters will in general be displayed as the difference to baseline defined as:

\[ \text{Change} = \text{Post baseline value} - \text{baseline value}. \]

NT-proBNP is assumed to be log-normally distributed. The change from baseline of NT-proBNP will therefore be calculated on the log-transformed values. Results will be displayed on the original scale, which is equivalent to the ratio of geometric means:

\[ \text{ChangeRatio} = \frac{\text{Post baseline NT-proBNP}}{\text{baseline NT-proBNP}}. \]

### 4.5.3 Repeated Measures

At all post-treatment visits and if not stated otherwise, only the values at scheduled measurements will be used for analysis. Measurements taken at unscheduled visits will be displayed in subject data listings.

In case of repeated measurements for pre-treatment visits and the randomization visit, the closest non-missing measurement prior to the treatment start will be used for analysis instead of the scheduled measurements.

#### 4.5.3.1 Vital Signs

Three measurements of vital signs parameters will be taken at time intervals of about 2 minutes. Averages of non-missing values of these three measurements will be calculated and used for the statistical analysis. If only one of the planned measurements is available, this value will be used.

### 4.5.4 Laboratory Data Handling

Only the data provided by the central laboratory will be used for analysis, values from local laboratories will not be used. An exception will be made for the screening lab, which will be only measured locally. These values will be presented in the analysis including a note that screening is measured in a local lab.

For values which are below the lower limit of quantification (LLOQ), half the value of the LLOQ will be used for analysis. Differences between two values of below the LLOQ will be assigned values of 0.

### 4.5.5 Subgroup analyses

The following subgroups will be considered for descriptive and explorative analyses:

- Age (years): <65, 65-75, >75
• Gender: male vs. female
• Race: White, Black, Hispanic, Asian, American Indian or Alaska Native
• Region: Eastern Europe, Western Europe, Asia/Pacific, North America
• Body Mass Index (BMI; kg/m²): ≤30 vs. >30
• Estimated Glomerular Filtration Rate (eGFR; mL/min/1.73m²) at baseline: ≤45, >45-60, >60
• Heart rhythm from central ECG assessment: Atrial fibrillation (defined as subjects with a finding of 'Atrial Fibrillation' or 'Atrial Flutter') vs. non-Atrial fibrillation (defined as subjects with any other heart rhythm different from 'Atrial Fibrillation' or 'Atrial Flutter')
• Initial presentation for worsening chronic heart failure (WCHF): hospitalization vs. equivalent
• LVEF (%) at baseline from echo core lab: ≤35 vs. >35
• NT-proBNP (pg/mL) at baseline: <median vs. ≥median
• Left atrial volume (LAV; mL) at baseline: <median vs. ≥median
• Time (days) from clinical stabilization to randomization: <median vs. ≥median
• SBP (mmHg) at baseline: <120 vs. ≥120
• New York Heart Association (NYHA) class at baseline: I/II vs. III/IV
• Diabetes mellitus: diabetes vs. no diabetes
• Coronary artery disease (CAD): present vs. not present
• Arterial hypertension: yes vs. no

Clinical stabilization is defined as end of prior IV diuretic treatment. If a subject did not take any IV diuretic treatment prior to randomization, clinical stabilization is defined as the end of initial hospitalization, if the subject was randomized after hospital discharge, or as the start of initial hospitalization, if the subject was randomized during initial hospitalization, respectively.

For subgroups split by median value, the full analysis set (FAS) defined in section 5.1.1 will be used to calculate the median.

If the total number of subjects in a subgroup category is below 35, the respective subgroup category will be either omitted from the analysis or combined with other categories, if a logical combination to another subgroup category is possible.
4.5.6 Maintenance Dose for Analysis of Actual Dose taken

In general, subjects will be analyzed by their assigned treatment group. Additional sensitivity analyses will be performed analyzing subjects according to their actual taken dose. As the dosing regimen comprises several (sham) titration steps, the actual taken dose may vary throughout the study. In order to assess the actual taken dose for each subject, the concept of maintenance dose is therefore used: each subject will be assigned to a maintenance dose (placebo, 1.25mg, 2.5mg, 5mg, or 10mg) group. Maintenance dose is defined as the dose reached at the first maintenance visit (visit 4), at which point the subject is assumed to generally be on his/her individual maximal tolerated dose. In case a subject is still on treatment at the planned visit 4 but misses the actual visit, the dose at the planned visit date will be used as maintenance dose. Subjects not reaching visit 4, e.g. due to drop-out, will be reported as a separate category "Maintenance dose not reached".

4.5.7 Exposure Groups

In addition to the maintenance dose analyses, sensitivity analyses per exposure will also be performed as an alternative derivation of actual taken dose. Exposure will both be used as continuous variable, and will be grouped to mimic the original defined treatment groups. If all medication is taken according to protocol and uptitration is always possible, the expected total exposure is expected to be 0mg, 105mg, 210mg, 385mg, and 665 mg in the respective treatment groups placebo, 1.25 mg, 2.5 mg, 5 mg, and 10 mg. Using cutpoints approximately in the middle between expected exposure of groups, the following exposure groups will be defined:

- Actual Placebo: total exposure < 50 mg
- Actual 1.25 mg: 50mg ≤ total exposure < 155mg
- Actual 2.5 mg: 155 ≤ total exposure < 290
- Actual 5 mg: 290 ≤ total exposure < 525
- Actual 10 mg: 525 ≤ total exposure

4.6 Validity Review

A validity review meeting will be held before final database closure where subject validity for the individual analysis sets will be decided. The results of the validity review meeting will be documented in the validity review report and may comprise decisions and details relevant for statistical evaluation. Any changes to the statistical analysis prompted by the results of the validity review meeting will be documented in an amendment or in a supplement to this SAP, as applicable.
5. Analysis Sets

5.1 Assignment of analysis sets

Final decisions regarding the assignment of subjects to analysis sets will be made during the validity review meeting and documented in the validity review report (see section 4.6).

Data from all subjects who signed informed consent, regardless of their assignment to certain analysis sets, will be presented in individual subject data listings. In addition, the following analysis sets will be used for this study:

5.1.1 Full analysis set (FAS)

The FAS includes all subjects randomized to treatment. The FAS will be used to display baseline characteristics and to display efficacy analyses.

5.1.2 Safety analysis set (SAF)

The SAF includes all subjects from FAS that have at least one dose of study drug administered. The SAF will be used to display baseline characteristics and safety analyses.

5.1.3 Per protocol set (PPS)

The PPS includes all subjects randomized to treatment that have a valid measurement of NT-proBNP at baseline and at week 12 (visit 5) and showed no major protocol deviations.

5.1.3.1 Major protocol deviations

Major protocol deviations include:

- Overall compliance with study drug intake of <80% or >120%.
- Study medication was interrupted for more than 10 consecutive days
- Study medication was not taken for more than 14 days in total before visit 5
- No history of chronic heart failure.

A complete list of all major protocol deviations leading to exclusion from the PPS will be specified in the Protocol Deviation Document (PDD). The PPS will be the primary analysis set for the primary efficacy analysis and will be used for further efficacy analyses as well as baseline characteristics.

5.1.4 Pharmacokinetics analysis set (PKS)

The PKS includes all subjects randomized to treatment that have at least one valid set of pharmacokinetics (PK) / pharmacodynamics (PD) measurements.
6. Statistical Methodology

6.1 Population characteristics

Population characteristics will be summarized overall and by randomized treatment group. Analyses will be performed in the FAS, PPS, and SAF, if the SAF differs from the FAS, if not stated otherwise.

6.1.1 Disposition

The number of subjects enrolled, randomized, and valid for the FAS, SAF, PPS and PKS will be summarized overall and by treatment groups, country and investigator. The number of subjects discontinuing the screening phase together with the primary reason for discontinuation will be presented overall. The number of subjects discontinuing the treatment and follow-up phases together with the primary reason for discontinuation will be presented by treatment groups and overall in separate tables. The number of subjects with survival status available at end of study will be presented. In addition, the number of subjects with major and minor protocol deviations will be presented overall and by investigator and country for each treatment group and in total. The frequencies of each major protocol deviation will be presented by treatment group and total.

6.1.2 Demography and Baseline Characteristics

Demographic variables and baseline characteristics will be summarized by treatment group and overall. Summary statistics will be presented for metric variables. Frequency tables will be presented for categorical variables.

Demography includes age, gender, race, ethnicity, region (Eastern Europe, Western Europe, Asia/Pacific, North America), body height, body weight, body mass index (BMI), smoking history, and alcohol consumption. Age and BMI will each be given as continuous variable and categorized with the following categories:

- Age (years): <65, 65-75, >75
- BMI (kg/m^2): ≤30, >30.

The following additional baseline characteristics will be analyzed:

- LVEF (%) at baseline from echo core lab (continuous)
- NT-proBNP (pg/mL) at baseline (continuous)
- eGFR (mL/min/1.73m^2) at baseline (continuous)
- eGFR (mL/min/1.73m^2) at baseline: ≤45, >45-60, >60
• Heart rhythm from central ECG assessment: Atrial fibrillation (defined as subjects with a finding of 'Atrial Fibrillation' or 'Atrial Flutter') vs. non-Atrial fibrillation (defined as subjects with any other heart rhythm different from 'Atrial Fibrillation' or 'Atrial Flutter')
• Heart rhythm from investigator assessment: Atrial fibrillation vs. non-Atrial fibrillation
• Initial presentation for WCHF: hospitalization vs. equivalent
• Time (days) from clinical stabilization to randomization (continuous)
• Previous HF hospitalization prior to the index event: yes vs. no
• SBP (mmHg) at baseline (continuous)
• SBP (mmHg) at baseline: <120, ≥120
• Diastolic blood pressure (DBP, mmHg) at baseline (continuous)
• Heart rate (HR, beats/min) at baseline (continuous)
• NYHA class at baseline: I, II, III, IV
• Cardiac device use at baseline: no, ICD, CRT, CRT-D, pacemaker
• Medical history of coronary artery disease (CAD): present vs. not present
• Medical history of diabetes mellitus: diabetes vs. no diabetes
• Medical history of Cardiomyopathy: ischemic vs. non-ischemic
• Medical history of atrial fibrillation: yes vs. no
• Medical history of arterial hypertension: present vs. not present
• Medical history of chronic kidney disease (CKD): present vs. not present
• Medical history of chronic obstructive pulmonary disease (COPD): present vs. not present
• Medical history of anemia: present vs. not present

Demographics and baseline characteristics tables will also be presented by the subgroups defined in section 4.5.5 in the FAS and PPS.

6.1.3 Medical History
Medical history findings will be coded by Medical Dictionary for Regulatory Activities (MedDRA) codes. Medical history will be presented for each MedDRA Primary System Organ Class (SOC) and Preferred Term (PT) by treatment group and overall.
6.1.4 Concomitant Medication

Prior and concomitant medications will be coded by Anatomical Therapeutic Chemical (ATC) classification system according to the World Health Organization Drug Dictionary (WHO-DD).

Medications with missing start and stop date but flagged as being ongoing at end of study will be considered to have started prior to study medication start and end after stop of study medication.

6.1.4.1 Selected Concomitant Medication of Special Interest

Concomitant medication of special interest, derived via ATC codes, Bayer Drug Grouping (BDG), or combinations of both ATCs and BDGs, will be summarized by class and corrected generic name. Of special interest are the drug groups:

- diuretics
- ACE inhibitors
- ARBs
- β-blockers
- MRAs

Diuretics will be categorized into the subcategories loop diuretics, thiazides, and potassium-sparing diuretics. In addition the combined drug group ACE inhibitor and/or ARB will be analyzed.

In addition, the following concomitant medication combinations are of special interest:

- Diuretics + ACE inhibitors and/or ARBs + β-blockers + MRAs
- Diuretics + ACE inhibitors and/or ARBs + β-blockers
- Diuretics + ACE inhibitors and/or ARBs

The number of subjects taking a medication in a drug group of special interest or combination of concomitant medications of special interest at any time during the study will be given. In addition, the number of subjects taking a medication in a drug group of special interest at baseline (visit 1) and visit 5 (week 12) will be given. Medication start and stop date information will be used to assess, if a medication was taken at the respective visits. In case of (partially) missing date information, a medication will be considered for the respective visit (i.e. missing start date will be imputed with the minimal possible date; missing stop date will be imputed with maximal possible date). In addition, the tables with number of subjects taking a medication in a drug group of special interest at baseline (visit 1) will be provided for the subgroups specified in section 4.5.5.
A cross-tabulation comparing the number of subjects taking a combination of concomitant medications of special interest at visit 1 and the number of subjects taking individual drug groups of interest at visit 5 will be presented.

**Analysis of diuretics**

An analysis of percentage of equivalence dose taken will be performed. Equivalence doses are given in Table 6-2. Percentage of equivalence dose is calculated individually per subject and medication as:

\[
\text{Percentage of equivalence dose} = \frac{\text{actual dose}}{\text{equivalence dose}} \times 100
\]

In case of missing dose and/or frequency information, the respective medication will be omitted from the analysis. Percentage of equivalence dose will be calculated for medications taken at baseline (visit 1) and at visit 5 (week 12).

In case a subject takes several medications from one diuretics subcategory at a visit, the sum of the percentages of equivalence dose of the different medications will be used for the subject for the respective visit.

For each diuretics subcategory, summary statistics will be provided for percentages of equivalence dose as well as changes from baseline in percentages of equivalence dose by treatment group and overall.

In addition, subjects will be grouped into 'below 50% of equivalence dose', '50% to 100% of equivalence dose', 'more than 100% of equivalence dose' and 'no respective medication intake'. For each diuretics subcategory, number and percentage of subjects within each percentage group will be given by treatment group and overall at baseline (visit 1) and visit 5.

Changes from baseline to visit 5 will be categorized into the categories 'dose increased', 'dose not changed', 'dose reduced', 'new drug started', 'drug intake stopped'.

<table>
<thead>
<tr>
<th>Table 6-1: Equivalence doses for different diuretics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diuretic (ingredient)</strong></td>
</tr>
<tr>
<td>---------------------------</td>
</tr>
<tr>
<td><strong>Loop diuretics</strong></td>
</tr>
<tr>
<td>Bumetanide</td>
</tr>
<tr>
<td>Furosemide</td>
</tr>
<tr>
<td>Torasemide</td>
</tr>
<tr>
<td><strong>Thiazides</strong></td>
</tr>
<tr>
<td>Bendroflumethiazide</td>
</tr>
<tr>
<td>Hydrochlorothiazide</td>
</tr>
</tbody>
</table>

<sup>*</sup> Equivalence doses are given in Table 6-2.
### Diuretic (ingredient) | Daily Equivalence* Dose (mg)
---|---
Indapamide | 2.5
Metolazone | 2.5

### Potassium-sparing diuretics

<table>
<thead>
<tr>
<th>Diuretic</th>
<th>Daily Dose (mg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amiloride</td>
<td>5# or 10†</td>
</tr>
<tr>
<td>Spironolactone/epi renone</td>
<td>50# or 100†</td>
</tr>
<tr>
<td>Triamterene</td>
<td>100# or 200†</td>
</tr>
</tbody>
</table>

*Equivalence does not mean bio-equivalence.

Daily equivalence doses are taken from table 16 in the ESC HF Guideline 2012 [5].

#### Analysis of maximum dose

For the drug groups ACE inhibitors, ARBs, β-blockers, and MRAs, an analysis of percentage of recommended maximum dose taken will be performed. Recommended maximum doses are given in Table 6-2. Percentage of recommended doses is calculated individually per subject and medication as:

\[
\text{Percentage of recommended maximum dose} = \frac{\text{actual dose}}{\text{recommended maximum dose}} \times 100
\]

In case of missing dose and/or frequency information, the respective medication will be omitted from the analysis. Percentage of recommended maximum dose will be calculated for medications taken at baseline (visit 1) and at visit 5 (week 12).

In case a subject takes several medications from one drug group at a visit, the sum of the percentages of recommended maximum dose of the different medications will be used for the subject for the respective visit.

For each drug group, summary statistics will be provided for percentages of recommended maximum dose as well as changes from baseline in percentages of recommended maximum dose by treatment group and overall.

In addition, subjects will be grouped into 'below 50% of recommended maximum dose', 'at least 50% of recommended maximum dose', and 'no respective medication intake'. For each drug group, as well as for the drug group combinations, number and percentage of subjects within each percentage group will be given by treatment group and overall at baseline (visit 1) and visit 5. A cross-tabulation showing the number of subjects in each category at baseline in comparison to visit 5 (week 12) will be provided by treatment group and overall.
### Table 6-2: Recommended maximum doses

<table>
<thead>
<tr>
<th>Medication (Ingredient)</th>
<th>Maximal daily dose (mg)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ACE inhibitor</strong></td>
<td></td>
</tr>
<tr>
<td>Captopril</td>
<td>150</td>
</tr>
<tr>
<td>Enalapril</td>
<td>40</td>
</tr>
<tr>
<td>Fosinopril</td>
<td>40</td>
</tr>
<tr>
<td>Lisinopril</td>
<td>40</td>
</tr>
<tr>
<td>Perindopril</td>
<td>16</td>
</tr>
<tr>
<td>Quinapril</td>
<td>40</td>
</tr>
<tr>
<td>Ramipril</td>
<td>10</td>
</tr>
<tr>
<td>Trandolapril</td>
<td>4</td>
</tr>
<tr>
<td><strong>ARB</strong></td>
<td></td>
</tr>
<tr>
<td>Candesartan</td>
<td>32</td>
</tr>
<tr>
<td>Losartan</td>
<td>150</td>
</tr>
<tr>
<td>Valsartan</td>
<td>320</td>
</tr>
<tr>
<td><strong>Beta-blocker</strong></td>
<td></td>
</tr>
<tr>
<td>Bisoprolol</td>
<td>10</td>
</tr>
<tr>
<td>Carvedilol</td>
<td>100</td>
</tr>
<tr>
<td>Metoprolol succinate (CR/XL)</td>
<td>200</td>
</tr>
<tr>
<td>Nebivolol</td>
<td>10</td>
</tr>
<tr>
<td><strong>MRA</strong></td>
<td></td>
</tr>
<tr>
<td>Eplerenone</td>
<td>50</td>
</tr>
<tr>
<td>Spironolactone</td>
<td>50</td>
</tr>
</tbody>
</table>

ACE = Angiotensin converting enzyme; ARB = Angiotensin receptor blocker; ATC = Anatomical Therapeutic Chemical; MRA = Mineralocorticoid receptor antagonist;

Maximal doses are taken from table 14 of the ESC HF Guideline 2012 [5] and table 15 of the ACCF/AHA HF Guideline 2013 [6].

6.1.4.2 Co-administration of Prohibited Concomitant Medication

The number of subject with co-administration of
• nitrates
• Phosphodiesterase type 5 inhibitors (PDE5i)

will be summarized by treatment group and overall. Co-administration is defined as intake of study medication and concomitant medication on the same day.

6.1.5 Treatment Duration and Exposure

Treatment duration (number of days with study drug intake, not including any gaps) will be summarized using descriptive statistics by treatment group and overall. In addition, treatment duration will be categorized to ≤7 days, >7-21 days, >21-35 days, >35-63 days, and >63 days, and presented with the corresponding number and percentage of subjects by treatment group and overall. A table will be presented with the absolute and relative frequencies of subjects still in the study at each visit. Kaplan-Meier plots for time-to end of study treatment will be provided for each treatment group.

The extent of exposure to study drug (total amount of intake in mg) will be summarized using descriptive statistics and by 50mg intervals up to 500mg, and 100mg intervals thereafter by treatment group and overall.

6.1.6 Treatment Compliance

The compliance will be summarized descriptively by treatment group and overall. In addition, compliance will be categorized into three groups (<80%, 80-120%, >120%) and summarized by treatment group and overall.

6.1.7 Titration Status

Dose titration by visit, and dose titration sequence will be summarized by treatment group and overall. Separate tables will be given for blinded dose (including sham titration) and actual dose received. Number of and reasons for up- and down-titration by visit and dose will be summarized using frequency counts.

6.2 Efficacy

6.2.1 Primary Efficacy Variable

The primary efficacy variable is change from baseline (visit 1) to visit 5 (week 12) in log-transformed NT-proBNP:

\[ y = \log (NT-proBNP \text{ at week 12}) - \log (\text{baseline NT-proBNP}) \]
This is equivalent to the log-transformed ratio of NT-proBNP at week 12 to NT-proBNP at baseline:

\[ y = \log \left( \frac{\text{NT-proBNP at week 12}}{\text{baseline NT-proBNP}} \right) \]

### 6.2.1.1 Primary Analysis of the Primary Efficacy Variable

The primary analysis will be performed in the PPS. This analysis set comprises only subjects that have a valid measurement of NT-proBNP at baseline (visit 1) and at visit 5 (week 12). Hence, no imputation of missing data in the primary efficacy variable is required for the primary analysis. Sensitivity analyses in the FAS with respective missing value imputations are described in Section 6.2.1.3.1.

The goal of the primary statistical analysis is to detect a significant relationship between the oral sGC stimulator vericiguat dose and the change in log-transformed NT-proBNP from baseline to visit 5 (week 12).

The assigned treatment groups of the three highest dose arms (2.5 mg, 5 mg, and 10 mg) will be pooled and compared to the assigned placebo treatment group with a one-sided two-sample t-test at the significance level \( \alpha = 5\% \). The 1.25 mg dose treatment group is assumed to have no or at most a very minimal effect and is hence not included in the pool of assumed effective treatment groups. The test problem is defined by:

\[ H_0^{pool} : \mu_{pool} \geq \mu_{placebo} \quad \text{vs.} \quad H_1^{pool} : \mu_{pool} < \mu_{placebo} \]

where \( \mu_{pool} \) is the mean change from baseline to week 12 in the population of the pooled treatment groups of the three highest doses 2.5 mg, 5 mg, and 10 mg, respectively. \( \mu_{placebo} \) is the mean change from baseline to visit 5 (week 12) in the population of the placebo treatment group.

To test the null hypothesis, the test statistic

\[ T = \frac{\bar{y}_{pool} - \bar{y}_{placebo}}{s \sqrt{1/n_{pool} + 1/n_{placebo}}} \]

will be compared to the \( \alpha \)-quantile of a t-distribution with \( n_{pool} + n_{placebo} - 2 \) degrees of freedom. \( \bar{y}_{pool} \) is the observed mean change from baseline to visit 5 in the pooled treatment groups of the three highest doses 2.5mg, 5mg, and 10mg, respectively. \( \bar{y}_{placebo} \) is the observed mean change from baseline to visit 5 in the placebo treatment group. \( n_{pool} \) and \( n_{placebo} \) are the sample sizes of the respective treatment groups and \( s \) is the observed standard deviation.

Results of the analysis will be back-transformed to be displayed on the original scale. Geometric means of the ratio of NT-proBNP at visit 5 (week 12) to baseline NT-proBNP in the pooled treatment groups, the placebo treatment group, and the ratio of these geometric
means of pooled treatment groups compared to placebo treatment group will be displayed along with two-sided 90% confidence intervals (CI).

6.2.1.2 Secondary Analyses of the Primary Efficacy Variable

The secondary analyses of the primary efficacy variable will be performed in the assigned treatment groups in the PPS.

6.2.1.2.1 Individual Pairwise Comparisons

If the null hypothesis of the primary analysis is rejected, the individual dose arms will each be compared to placebo with a one-sided two-sample t-test at the significance level \( \alpha = 5\% \) in a sequential order (i.e., if all the previous null hypotheses have been rejected, the next null hypothesis will be tested):

- the 10 mg treatment group vs. placebo treatment group:
  \[ H_0^{10} : \mu_{10} \geq \mu_{\text{placebo}} \quad \text{vs.} \quad H_1^{10} : \mu_{10} < \mu_{\text{placebo}} \]
- the 5 mg treatment group vs. placebo treatment group:
  \[ H_0^{5} : \mu_{5} \geq \mu_{\text{placebo}} \quad \text{vs.} \quad H_1^{5} : \mu_{5} < \mu_{\text{placebo}} \]
- the 2.5 mg treatment group vs. placebo treatment group:
  \[ H_0^{2.5} : \mu_{2.5} \geq \mu_{\text{placebo}} \quad \text{vs.} \quad H_1^{2.5} : \mu_{2.5} < \mu_{\text{placebo}} \]
- the 1.25 mg treatment group vs. placebo treatment group:
  \[ H_0^{1.25} : \mu_{1.25} \geq \mu_{\text{placebo}} \quad \text{vs.} \quad H_1^{1.25} : \mu_{1.25} < \mu_{\text{placebo}} \]

where \( \mu_{10}, \mu_{5}, \mu_{2.5}, \mu_{1.25}, \) and \( \mu_{\text{placebo}} \) are the mean changes from baseline to visit 5 (week 12) in the populations of the individual treatment groups 10 mg, 5 mg, 2.5 mg, 1.25 mg, and placebo, respectively.

Results of each analysis will be back-transformed to be displayed on the original scale. For each test, the geometric means of the ratio of NT-proBNP at visit 5 (week 12) to baseline NT-proBNP in the respective active treatment group, the placebo treatment group, and the ratio of these geometric means of active treatment group compared to placebo treatment group will be displayed along with two-sided 90% CIs.

In case the primary analysis has not been successful, the comparisons of individual active treatment groups to the placebo treatment group will be exploratory only.

6.2.1.2.2 Linear Regression

To obtain further insight on the dose-response relationship in the primary efficacy variable a linear regression model will be analyzed in an explorative manner:

\[ y_i = \beta_0 + \beta_1 x_i + e_i, \]

where
\[ y_i \] primary efficacy variable for subject \( i \)
\[ \beta_0 \] intercept (effect of placebo/no dose)
\[ \beta_1 \] slope parameter (dose-response relationship)
\[ x_i \] dose of subject \( i \)
\[ e_i \] random error for subject \( i \) (assumed to be iid \( \sim N(0,\sigma^2_{ei}) \))

The slope of this model will be assessed to analyze the dose-response relationship. The test problem is defined by:

\[ H_0 : \beta_1 \geq 0 \quad \text{vs.} \quad H_1 : \beta_1 < 0 \]

The test will be performed at the one-sided significance level \( \alpha = 5\% \).

6.2.1.3 Exploratory and Sensitivity Analyses of the Primary Efficacy Variable

6.2.1.3.1 Sensitivity Analysis in FAS Population

The primary and secondary analyses of the primary efficacy variable will be repeated in the FAS. This analysis set comprises also subjects that have missing measurements of NT-proBNP at baseline and/or at visit 5 (week 12). Subjects with a missing NT-proBNP measurement at baseline will be omitted from this analysis. The following single imputation methods will be used to impute missing values at visit 5 (week 12):

- a completers analysis taking only subjects with available NT-proBNP measurement at visit 5 (week 12) into account (FAS on completers; only in case the population differs from PPS)
- an on-treatment last observation carried forward (LOCF) approach will be applied, i.e. the last available on-treatment NT-proBNP measurement before visit 5 (week 12) will be imputed for visit 5. In case the baseline is the only available NT-proBNP measurement, this will be imputed for visit 5.
- an observed cases analysis taking only subjects with an available NT-proBNP measurement at visit 5 or subjects who died prior to visit 5 or prematurely dropped out due to adverse event (AE). For subjects dropping out due to AE or death, a worst value will be imputed as the maximum NT-proBNP measurement observed for that patient from baseline until dropout.

6.2.1.3.2 Delta Method for Data Missing not at Random

The primary analysis of the primary efficacy variable will be repeated using a multiple imputation approach in the FAS. In the context of a pattern mixture framework, this is also called the delta method. This analysis assumes data is missing not at random (MNAR).
Subjects' NT-proBNP values are assumed to worsen after drop-out. The following scenarios will be investigated, where different penalties are assigned to the different treatment groups.

Fixed penalties after drop-out:

- placebo: 0 log(pg/mL) vs. pooled dose groups: 0.075 log(pg/mL)
- placebo: 0 log(pg/mL) vs. pooled dose groups: 0.15 log(pg/mL)
- placebo: 0.075 log(pg/mL) vs. pooled dose groups: 0.075 log(pg/mL)
- placebo: 0.075 log(pg/mL) vs. pooled dose groups: 0.15 log(pg/mL)
- placebo: 0.15 log(pg/mL) vs. pooled dose groups: 0.15 log(pg/mL)

Decreasing slope (per visit after drop-out):

- placebo: 0 log(pg/mL) vs. pooled dose groups: 0.025 log(pg/mL)
- placebo: 0 log(pg/mL) vs. pooled dose groups: 0.05 log(pg/mL)
- placebo: 0.025 log(pg/mL) vs. pooled dose groups: 0.025 log(pg/mL)
- placebo: 0.025 log(pg/mL) vs. pooled dose groups: 0.05 log(pg/mL)
- placebo: 0.05 log(pg/mL) vs. pooled dose groups: 0.05 log(pg/mL)

To properly account for the incompleteness of the data, multiple imputation will be used to draw sets of completed data that will then be modified according to the scenarios given above. Multiple imputation will be done using SAS PROC MI using the following generic code:

```
proc mi data=change_miss seed=999 nimpute=50 out=change_mi;
  by treat;
  MCMC niter=500 nbiter=500;
  var NTproBNPchange_Visit2 - NTproBNPchange_Visit5;
run;
```

After modifying the completed data sets according to the scenarios, the t-test of the main analysis will be performed at visit 5 for each completed data set. The results are then combined using SAS PROC MIANALYSE.

For each scenario, the treatment difference at visit 5 will be given with a 90%-confidence interval and a p-value. In addition, the minimum and maximum of observed treatment differences over the imputed data sets will be presented.

### 6.2.1.3.3 Sensitivity Analysis for highest Dose Arms

The primary efficacy variable will be analyzed in the PPS and FAS on completers by pooling the actual treatment groups of the two highest dose arms (5 mg, and 10 mg) and comparing them to the placebo treatment group with a two-sample t-test at the one-sided significance level $\alpha = 5\%$.
6.2.1.3.4 Sensitivity Analyses with Analysis of Covariance (ANCOVA) Models

The primary efficacy variable will be analyzed by ANCOVA models including log-transformed baseline NT-proBNP value as covariate in the PPS and FAS on completers. One model will be performed as an extension to the primary analysis, including placebo vs. the pooled three highest dose arms as the main effect. A second ANCOVA model will include all treatment groups separately.

For each model, the treatment effect will be tested at a two-sided significance level of 5%. For the second model, pairwise differences between each vericiguat treatment group and the placebo treatment group will be calculated and corresponding two-sided 95% confidence intervals will be computed.

These analyses will be repeated including in addition atrial fibrillation vs. non-atrial fibrillation (from central ECG assessment) and hospitalization vs. equivalent as strata in the model.

In case of discrepant demographics or baseline characteristics, ANCOVA models including the baseline characteristic as covariate and placebo vs. the pooled three highest dose arms as the main effect will be performed in the PPS and FAS on completers.

To check if demographics or baseline characteristics are different between placebo and the pooled three highest dose arms, an t-test will be performed for continuous variables and a chi-square test (with SAS option: normal) will be performed for categorical variables, both at the two-sided significance level of 10%. The per-protocol set will be used for these analyses. In case both a continuous and categorical variable is defined for the same endpoint, only the continuous variable is checked and included in the ANCOVA model.

6.2.1.3.5 Analysis per Maintenance Dose

The primary and secondary analyses of the primary endpoint will be repeated in the PPS and FAS on completers using the maintenance doses; i.e., subjects will be categorized according to the maintenance doses they reached instead of the assigned treatment groups.

6.2.1.3.6 Analysis per Exposure

The primary analysis of the primary endpoint and the comparisons of the secondary analysis of the individual pairwise comparisons will be repeated in the PPS and FAS on completers using the exposure groups; i.e. subjects will be categorized according to the exposure groups instead of the assigned treatment groups.

In addition, the linear regression of the primary endpoint will be repeated using the total exposure as continuous variable instead of dose.

6.2.1.3.7 Area under the Curve (AUC)

AUC until visit 5 of change from baseline in log-transformed NT-proBNP will be summarized by treatment group and overall for the PPS and FAS on completers. In addition, figures will be provided for each treatment group and overall. The planned visit dates will be
used for this analysis. Linear interpolation will be used between the visits. In case a subject missed a visit, the subsequent visit will be used for the linear interpolation.

6.2.1.3.8 Exploratory Subgroup Analyses

The primary efficacy variable will be summarized descriptively for the subgroups described in section 4.5.5. In addition, 90% confidence intervals for the ratio of pooled treatment groups to placebo treatment group will be given for each subgroup. These will be obtained from an Analysis of Variance (ANOVA) model fit to the subgroup, containing treatment group (pooled three highest dose arms vs. placebo) as main effect. An additional ANOVA model will include all treatment groups separately. The comparison of treatment effect of the pooled three highest dose arms, as well as each of the individual dose arms to placebo will be displayed with forest plots for all subgroups.

Interaction of each subgroup with treatment will be assessed by separate ANOVA models including the main effects treatment group (pooled three highest dose arms vs. placebo), subgroup and the interaction term subgroup*treatment.

6.2.1.3.9 Exploratory Fitting of Dose-Response Model

Several models will be fit to model the dose response in the primary efficacy variable:

- Saturation model: \( (E_0 + E_{max} d^h)/(g + d^h) \)
- Exponential: \( E_0 + E_{max}[\exp(d/\delta) - 1] \)
- \( E_{max}: E_0 + E_{max} d/(ED_{50} + d) \)
- Sigmoidal \( E_{max} \) model: \( E_0 + E_{max} d^h/(ED_{50}^h + d^h) \),

where \( d \) is the dose. The saturation model was originally assumed for sample size calculation. The other models will be fit in addition to explore if other shapes of dose-response models would fit the data better. Model fitting will be performed by nonlinear least squares using the iterative Gauss-Newton method. For each model, parameter estimates with corresponding approximate standard errors and 95% confidence intervals will be displayed. A plot of each fitted model will be provided.

For the model fitting algorithm, starting values are required. The following initial starting values will be used for the above models:

- \( E_0 = \text{mean(change from baseline in log-transformed NT-proBNP in placebo group)} \)
- \( E_{max} = \max(\text{mean change in Placebo group, mean change in 1.25mg treatment group, mean change in 2.5mg treatment group, mean change in 5mg treatment group, mean change in 10mg treatment group}) \)
- \( ED_{50} = \text{mean(dose with least difference to mean drug effect} (= (E_0 + E_{max})/2)) \)
- \( g = 0 \)
- \( \delta = 1 \)
- \( E_1 = 1 \)
- \( h = 1 \)

In case the fitting algorithm does not converge for a model with the above starting values, other starting values will be assessed to allow model convergence.
6.2.1.3.10 Converter Analyses
NT-proBNP cut-offs have previously been studied for goal-directed therapy:

- 1000 pg/mL [7,8]
- 1692 pg/mL (converted from 200 pmol/L) [9]
- 2200 pg/mL [10].

Clinical outcomes were improved when using these cut-offs to direct therapy. The number of subjects with NT-proBNP below each of these cut-offs at visit 5 among all subjects with NT-proBNP above the respective cut-off at baseline as well as the number of subjects with NT-proBNP above the cut-offs at visit 5 among all subjects with NT-proBNP below the respective cut-off at baseline will be summarized by treatment group and overall.

6.2.2 Analyses of Exploratory Efficacy Variables
All further exploratory efficacy variables will be analyzed in the FAS and PPS if not otherwise stated.

6.2.2.1 Echocardiography
All echocardiography parameters described in the protocol and their changes from baseline will be summarized descriptively by assigned treatment group and by maintenance dose, respectively, and by visit.

The change in echocardiography parameters from baseline to visit 5 (week 12) will be analyzed by two-sided two-sample t-tests comparing each individual vericiguat treatment group with the placebo treatment group at the 5% level in an explorative setting.

In case further parameters are derived via post-processing of stored echo images, a separate SAP will be created describing any further analyses of such additional parameters.

6.2.2.2 Efficacy Biomarkers
All efficacy biomarkers, i.e. NT-proBNP, Galectin-3 (Gal-3), growth differentiation factor 15 (GDF-15), osteopontin (OPN), pro-collagen III peptide (PIIINP), soluble suppression of tumorigenicity 2 (sST2), tissue metallopeptidase inhibitor 4 (TIMP-4), and cyclic guanosine monophosphate (cGMP), and their absolute changes from baseline will be summarized descriptively by assigned treatment group and by maintenance dose, respectively, and by visit.

The change in efficacy biomarkers from baseline to visit 5 (week 12) will be analyzed by two-sided two-sample t-tests comparing each individual vericiguat treatment group with the placebo treatment group at the 5% level in an explorative setting.

Additional analyses of efficacy biomarkers and their results will be provided in a separate report.
6.2.2.3 Health-related Quality of Life

The 23 questions of the Kansas City Cardiomyopathy Questionnaire (KCCQ) comprise seven domains:

- Physical Limitation (questions 1a to 1f), referred to as ‘functional status’ in the protocol
- Symptom Stability (question 2), referred to as ‘change in symptoms’ in the protocol
- Symptom Frequency (questions 3, 5, 7, 9), referred to as a part of ‘symptoms/signs’ in the protocol
- Symptom Burden (question 4, 6, 8), referred to as a part of ‘symptoms/signs’ in the protocol
- Self-Efficacy (question 10, 11), referred to as ‘self-efficacy and knowledge’ in the protocol
- Quality of Life (questions 12, 13, 14), referred to as ‘health perceptions’ in the protocol
- Social limitations (questions 15a to 15d), referred to under the same name in the protocol.

For each domain, a score will be calculated according to the KCCQ scoring instruction [11]. The individual domain scores will additionally be summarized to three summary scores:

- The Total Symptom Score, referred to as ‘symptoms/signs’ in the protocol, is the mean of the Symptom Burden and the Symptom Frequency scores.
- The Clinical Summary Score is the mean of the Total Symptom score and the Physical Limitation score.
- The Overall Summary Score is the mean of the Total Symptom score, and the Physical Limitation, Quality of Life, and Social Limitation scores.

For the EQ-5D-3L, two summary scores will be calculated out of the five dimensions according to the EQ-5D-3L User Guide [12]. The value sets for UK and US described in the EQ-5D Value Sets [13] will be used to derive the scores.

For the KCCQ and EQ-5D-3L, frequencies of answers to individual questions will be displayed by treatment group and overall by visit. In addition, changes from baseline for the single questions will be categorized into the categories improvement / no change / worsening. Frequencies of the different categories will be displayed by assigned treatment group and overall by visit.

The summary scores of the KCCQ (Total Symptom Score, Clinical Summary Score and Overall Summary Score) and EQ-5D-3L (UK and US) as well as the EQ Visual Analogue Scale (VAS) and the scores for the seven distinct domains of the KCCQ and the corresponding changes from baseline will be described by treatment group and visit.
6.2.2.4 Clinical Efficacy Variables (Hospitalization and mortality)

Blinded adjudication of all hospitalizations and deaths will be performed by a central clinical event committee. Details about the pre-specified categories can be found in the protocol and in the adjudication charter. In addition, all reported SAEs will also be adjudicated by the clinical event committee to assess, if they belong to the category of emergency presentation for WCHF. When death is collected during vital status follow-up and is not related to an AE, no AE information can be used for adjudication. In this case when no sufficient information for adjudication is available, death will be adjudicated as CV death by definition, to apply a conservative approach. Only the adjudication results will be used in the analyses of hospitalization, mortality and emergency presentation for WCHF.

The absolute and relative frequency of subjects with the following events until day 84 (planned Visit 5), on treatment (until end of study drug + 5 days) and until end of follow-up will be displayed overall and by treatment group.

- First event of the composite endpoint of death from any cause, cardiovascular (CV) hospitalization, or emergency presentation for WCHF
- First event of the composite endpoint of CV death, CV hospitalization, or emergency presentation for WCHF
- First event of the composite endpoint of HF hospitalization and CV death
- First event of HF: composite of HF hospitalization and emergency presentation for WCHF
- First event of major adverse cardiovascular events (MACE): CV death, non-fatal acute myocardial infarction, non-fatal cerebrovascular accident (transient ischemic attack or stroke)
- CV hospitalizations
- HF hospitalizations
- Emergency presentation for WCHF
- Death (all cause)
- CV death

The number of individual events until day 84, on treatment, and until end of follow-up will be displayed overall and by treatment group for CV hospitalization and emergency presentation for WCHF.

Absolute and relative frequency of subjects with events of the sub-classifications of CV hospitalization and CV death will be presented as well as the number of individual events in the FAS, PPS, and SAF.

Summary statistics of the length of hospitalizations during the treatment period of the study per subject will be given both absolute and relative to treatment duration for each treatment
Treatment duration is defined as the period from start of study drug intake to end of study drug intake (including any gaps) for this analysis. Hospitalizations during the study are approximated by serious adverse events that have 'requires or prolongs hospitalization' as reason for seriousness. In case of several overlapping respective SAEs exist, the earliest start and latest stop date are taken to approximate length of corresponding hospital stay. In case start and/or stop date of the hospitalization is (partially) missing, the subject will not be considered for the analysis. ANOVA models with treatment group as main effect will be fit to compare each vericiguat treatment group, as well as the pooled three highest dose arms, to the placebo treatment group.

The number of days out of hospital and alive until day 84 and the days within hospital or lost due to death until day 84 will be summarized. For this analysis, (partially) missing stop dates of hospitalizations are imputed by the death date, in case a subject died until day 84. Subjects who did not die until day 84 and have a hospitalization with (partially) missing start and/or stop date, as well as subjects prematurely terminated the study treatment and did not die until day 84 will not be considered for the analysis. Number of days out of hospital and alive is defined as

\[
\text{Min}(\text{day 84, date of death}) - \text{Randomization date} - \text{length of hospitalization until day 84 + 1.}
\]

The number of days within hospital or lost due to death is defined as

\[
\text{Max}(\text{day 84 – date of death+1, 0}) + \text{length of hospitalization until day 84.}
\]

In addition, time-to-event analyses will be performed. Start of observation period and rules for censoring are displayed in Table 6-3.

### Table 6-3: Definition of observation period and censoring rules

<table>
<thead>
<tr>
<th>Timepoint</th>
<th>Analysis set</th>
<th>Start of period</th>
<th>Events until</th>
<th>Censoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 84</td>
<td>PPS</td>
<td>1st Treatment Intake</td>
<td>Day 84</td>
<td>Day 84 for completers; Minimum of Day of death and day 84 for subjects died (in case death is not event); Minimum of last visit* and day 84 for dropouts.</td>
</tr>
<tr>
<td>Day 84</td>
<td>FAS</td>
<td>Randomization</td>
<td>Day 84</td>
<td>Day 84 for completers; Minimum of Day of death and day 84 for subjects died (in case death is not event); Minimum of last visit* and day 84 for dropouts.</td>
</tr>
</tbody>
</table>
### Timepoint Analysis

<table>
<thead>
<tr>
<th>Timepoint</th>
<th>Analysis set</th>
<th>Start of period</th>
<th>Events until</th>
<th>Censoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>on treatment</td>
<td>PPS</td>
<td>1st Treatment Intake</td>
<td>End of study drug + 5 days</td>
<td>Minimum of end of study drug + 5 days and last visit* for completers and dropouts; Minimum of End of study drug + 5 days and Day of death for subjects died (in case death is not event).</td>
</tr>
<tr>
<td>on treatment</td>
<td>FAS</td>
<td>Randomization</td>
<td>End of study drug + 5 days</td>
<td>Minimum of end of study drug + 5 days and last visit* for completers and dropouts; Minimum of End of study drug + 5 days and Day of death for subjects died (in case death is not event).</td>
</tr>
<tr>
<td>end of follow-up</td>
<td>PPS</td>
<td>1st Treatment Intake</td>
<td>End of follow-up</td>
<td>Last visit* for completers and dropouts; Maximum of last visit and day of death (in case death is not event)</td>
</tr>
<tr>
<td>end of follow-up</td>
<td>FAS</td>
<td>Randomization</td>
<td>End of follow-up</td>
<td>Last visit* for completers and dropouts; Maximum of last visit and day of death (in case death is not event)</td>
</tr>
</tbody>
</table>

* for dropouts, last visit is defined as:
- last visit in SV for non-fatal events and composite endpoints containing non-fatal events
- last visit in DS for fatal events.

Kaplan-Meier estimates for each of the above event types will be presented by treatment group and overall considering all respective events until day 84, on treatment and until end of follow-up or until censoring, respectively. Kaplan-Meier curves will be presented for each vericiguat treatment group, as well as the pooled three highest dose arms, together with the placebo treatment group. The log-rank test will be applied in order to do an explorative test for differences between each vericiguat treatment group, as well as the pooled three highest dose arms, and the placebo treatment group at a two-sided significance level of 5%.

The event rate for CV death, first HF hospitalization and the composite of the two at 12 months will be approximated.

In the first scenario, the number of events per patient year will be displayed. The frequency of events per patient year for a treatment group, in the pooled treatment group and total
population will be calculated as the number of (first) events observed for a (pooled) treatment group or total population until end of treatment, divided by the sum of the days from start of study drug intake to min(stop of study drug intake, event date) for all subjects within a (pooled) treatment group or the total population and multiplied by 365.25.

In the second scenario, an exponential model will be fit to the observed event times for each treatment group, the pooled treatment group and the total population using the PROC LIFEREG procedure. The fitted model is then used to estimate the event rate at 12 months.

### 6.2.2.5 Further Clinical Efficacy Variables

#### 6.2.2.5.1 NYHA Class

Absolute and relative frequencies of subjects in the different NYHA classes will be presented by treatment group and overall by visit.

Absolute and relative frequencies of subjects with change in NYHA class from baseline (e.g. unchanged, improved by 1 category, deteriorated by 1 category, etc.) will be presented by treatment group and visit.

The number of subjects with transitions from baseline with respect to categories (class I, class II, class III, class IV) will be provided by baseline value, treatment group and visit.

#### 6.2.2.5.2 Signs and Symptoms of Congestion

Six signs and symptoms of congestion and their scoring are described in Table 6-4, as given by [14]. The composite congestion score is composed of the three components orthopnea, jugular venous distension, and edema. The composite congestion score is calculated as the sum of the scores of the respective three components, thus ranging from 0 to 9.

<table>
<thead>
<tr>
<th>Score</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signs and Symptoms</td>
<td>None</td>
<td>Seldom</td>
<td>Frequent</td>
<td>Continuous</td>
</tr>
<tr>
<td>Dyspnea</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orthopnea</td>
<td>None</td>
<td>Seldom</td>
<td>Frequent</td>
<td>Continuous</td>
</tr>
<tr>
<td>Fatigue</td>
<td>None</td>
<td>Seldom</td>
<td>Frequent</td>
<td>Continuous</td>
</tr>
<tr>
<td>Jugular venous distension (cmH\textsubscript{2}O)</td>
<td>≤ 6</td>
<td>6 - 9</td>
<td>10 - 15</td>
<td>≥ 15</td>
</tr>
<tr>
<td>Rales</td>
<td>None</td>
<td>Bases</td>
<td>to &lt;50%</td>
<td>to &gt;50%</td>
</tr>
<tr>
<td>Edema</td>
<td>Absent / trace</td>
<td>Slight</td>
<td>Moderate</td>
<td>Marked</td>
</tr>
</tbody>
</table>

For each of the six signs and symptoms of congestion, absolute and relative frequencies of score values and the change from baseline in score values will be presented by treatment

---

Reference Number: BPD-SOP-060
Supplement Version: 5
group and overall by visit. In addition, absolute and relative frequencies will be presented for the composite congestion score as well as change from baseline in composite congestion score by treatment group and overall by visit.

6.3 Pharmacokinetics / pharmacodynamics

The pharmacokinetic (PK) analysis will be performed in the PKS.

Vericiguat peak and trough plasma concentrations will be summarized per visit, separated according to assigned dose. The following statistics will be calculated for each of the sampling points: arithmetic mean, standard deviation and coefficient of variation, geometric mean, geometric standard deviation (re-transformed standard deviation of the logarithms) and geometric coefficient of variation, minimum, median, and maximum value and the number of measurements.

Means at any time will only be calculated if at least 2/3 of the individual data were measured and were above the lower limit of quantification (LLOQ). For the calculation of the mean value a data point below LLOQ will be substituted by one half of this limit. In tables showing mean values, where values below LLOQ are included in the calculation of mean values, these means will be marked.

In addition, PK and, if applicable, PK/PD modeling using population approaches to describe vericiguat pharmacokinetics including potential influence of relevant patient covariables (e.g. age, gender, body weight, etc.) or potentially to relate parameters of clinical safety and efficacy response with vericiguat plasma concentrations will be investigated under a separate detailed PK/PD evaluation plan.

6.4 Safety

Safety analyses will be performed in the SAF.

6.4.1 Adverse events

Adverse events (AEs) will be coded by MedDRA. The version number of MedDRA used for the analyses will be stored in the clinical database. A listing will be provided linking the original investigator terms and the coded terms.

AEs are considered to be treatment-emergent if they have started or worsened after first application of study medication up to 5 calendar days after end of treatment with study medication.

Acute renal failure (such as serum creatinine ×2 or GFR decreased >50% or to <15 mL/min/1.73m²) and newly developing signs and symptoms of WCHF requiring IV treatment with diuretics and/or positive inotropic agents have both been defined as AEs of special safety interest in the protocol (Section 7.5.1.6).
An overall summary of AEs and treatment-emergent (TE) AEs will be generated by treatment group and overall. The summaries will be repeated using the exposure groups and using the maintenance dose groups. The summaries will be repeated by the subgroups age and eGFR as defined in section 4.5.5.

Incidences of subjects with TEAEs, drug-related and/or serious TEAEs, TEAEs causing discontinuation of study drug, and TEAEs of special safety interest will be summarized by treatment group and overall grouped by MedDRA Primary System Organ Class (SOC) and Preferred Term (PT). In addition, the incidence of pre-treatment AEs and AEs during the follow-up (more than 5 calendar days after end of treatment with study medication) will be tabulated. The incidence table of subjects with TEAEs of special safety interest will be repeated by exposure groups and by maintenance dose groups, as well as by the subgroups age and eGFR as defined in section 4.5.5.

The incidence of subjects with TEAEs per patient year will be displayed. The incidence per patient year is derived by dividing the incidence of subjects with TEAEs observed for a treatment group by the sum of the days from start of study drug intake to stop of study drug intake + 5 days for all subjects within a treatment group and multiply this by 365.25.

Serious adverse events (SAEs), deaths, AEs leading to discontinuation and AEs of special safety interest will be listed. The date, relative day (to study medication) and phase of the study (pre-treatment, during treatment, post-treatment) will be included.

Further summaries of AEs by intensity and outcome will be provided, consistent with Bayer Global Medical Standards.

Adverse events which belong to hearing, vision, or neurological disorders will be described in separate tables. The events will be identified by Standard MedDRA Queries (SMQs; code refers to SMQ_VIEW of the given MedDRA version). Narrow search terms are flagged in SMQ_VIEW as CATEGORY="1A". The broad search covers all terms of a SMQ.

Hearing disorder: MedDRA SMQ "Hearing impairment (SMQ)" (code=20000171)
Visual disorder: MedDRA SMQ "Optic nerve disorders" (code=20000148)
Neurological signs: MedDRA SMQ "Parkinson-like events" (code=20000099)

For each of the three SMQs and both narrow and broad search strategy, the following tables will be provided: overall summary of TEAEs, incidences of subjects with TEAEs, and with serious TEAEs grouped by SOC and PT.

In addition, overall summary of TEAEs, incidences of subjects with TEAEs, and with serious TEAEs grouped by SOC and PT will be provided for hypotension adverse events:

Hypotension: MedDRA SMQ "Hypotension (Riociguat)" (code=SMQ_1388).

The overall summary of TE hypotension AEs will be repeated using the exposure groups and using the maintenance dose groups. The summary will be repeated by the subgroups age and eGFR as defined in section 4.5.5. A listing of all hypotension events will be provided,
including date, relative day (to study medication) and phase of the study (pre-treatment, during treatment, post-treatment).

Incidences of subjects with adverse events related to co-administration with nitrates or PDE5i will be summarized by treatment group and overall. Listings of subjects with AEs related to co-administration will be provided. An adverse event is defined as being related to co-administration if it occurs following up to 1 days after co-administration with nitrates or PDE5i, respectively.

An overview of all MedDRA SOC and PT belonging to the respective SMQs will be provided.

6.4.2 Further safety parameters

6.4.2.1 Safety biomarkers

All safety biomarkers, i.e. carboxyterminal cross-linking telopeptide (CTX) and bone-specific alkaline phosphatase (bAP), will be summarized descriptively by treatment group and visit including absolute changes from baseline.

Additional analyses of safety biomarkers and their results will be provided in a separate report.

6.4.2.2 Laboratory parameters

Summary statistics including changes to baseline will be calculated by treatment group and visit for all quantitative laboratory parameters, i.e. hematology, clinical chemistry and coagulation parameters. The summary statistics for troponin T and eGFR will also be provided by exposure groups and by maintenance dose groups.

The change in laboratory parameters troponin T, serum creatinine, eGFR, and bilirubin from baseline to visit 5 (week 12) will be analyzed by two-sided two-sample t-tests comparing each individual vericiguat treatment group with the placebo treatment group at the 5% level in an explorative setting.

The number of subjects with transitions from baseline with respect to reference range categories (low, normal, high) will be provided by treatment group and visit. In addition, the number of subjects with treatment-emergent abnormal laboratory values above or below the normal range will be tabulated by treatment group and overall.

To assess acute renal failure, the number and incidence of subjects with increase in serum creatinine >0.3 mg/dL, >0.5 mg/dL, and ≥×2 from baseline, as well as the number and incidence of subjects with decrease in eGFR >50% or to <15 mL/min/1.73m² will be presented by visit and treatment group.
The number and incidence of subjects with ALT > 3× upper limit of normal (ULN), and total bilirubin > 2×ULN will be presented by visit and treatment group.

AUC until visit 5 of troponin T, serum creatinine as well as eGFR will be summarized by treatment group and overall. In addition, figures will be provided for each treatment group and overall. The planned visit dates will be used for this analysis. Linear interpolation will be used between the visits. In case a subject missed a visit, the subsequent visit will be used for the linear interpolation. In case of dropouts, the last available value will be carried forward until visit 5.

6.4.2.3 Vital signs

The absolute values and the change from baseline values at each visit will be summarized by treatment group, exposure groups and by maintenance dose using descriptive statistics for systolic blood pressure (SBP), diastolic blood pressure (DBP), heart rate (HR), mean arterial pressure (MAP), and pulse pressure (PP).

The change in vital signs parameters from baseline to visit 5 (week 12) will be analyzed by two-sided two-sample t-tests comparing each individual vericiguat treatment group with the placebo treatment group at the 5% level in an explorative setting.

The number of subjects with systolic blood pressure below 90 mmHg, based on the individual measurements rather than the mean value, will be displayed by visit and by treatment group. A listing of all subjects with systolic blood pressure below 90 mmHg post baseline will be provided.

6.4.2.4 Electrocardiogram (ECG)

For ECGs, the status pre-treatment and post treatment-initiation will be tabulated. The incidence rates of treatment-emergent ECG abnormalities will be tabulated by treatment group. A descriptive analysis of continuous ECG parameters and their changes from baseline by visit and treatment group will also be presented. Prolongations of QT interval will be summarized by visit and treatment group.

7. Document history and changes in the planned statistical analysis

7.1 Document history

- SAP version 0.2 for team review 09 May 2014
- SAP version 0.3 for team review 30 January 2015
- SAP version 0.4 for team review 21 February 2015
- SAP version 1.0, final version 24 February 2015
- SAP amendment version 1.1 for team review 06 June 2015
7.2 Changes in the planned statistical analysis

7.2.1 Changes from the protocol

Health-related quality of life (KCCQ)

For the analysis of the KCCQ questionnaire the protocol states that it will be summarized to 6 domain scores and 2 summary scale scores. According to the Questionnaire manual one of the 6 mentioned domain scores (the symptom score) combines actually the domains Symptom frequency and Symptom burden. Therefore the two single domains Symptom frequency and Symptom burden will separately be described as domain scores and the Overall Symptom score as a third summary score.

Composite congestion score

The protocol lists six different signs and symptoms for the composite congestion score. However, only the three components orthopnea, jugular venous distension, and edema are used to derive the composite congestion score. The composite congestion score is calculated as the sum of the scores of the respective three components, thus ranging from 0 to 9.

7.2.2 Changes from SAP version 1.0 to version 2.0

Time from prior IV diuretic treatment to randomization

The variable is renamed and further information on its derivation are included. The variable is included as a baseline characteristic.

Region

Included for subgroup analyses.

Survival Follow-up

Display of survival follow-up included.

Concomitant medications

Rules for displaying medication with missing start date and ongoing at end of study are included.

Analyses of co-administration of nitrates and PDE5i are included.

ANCOVA models including discrepant baseline characteristics
Definitions of when baseline characteristics and demographics are assumed to be discrepant between treatment groups are included.

**Dose-response modeling**
Starting values for the model fitting algorithm are defined.

**Converter rates**
An analysis of converter rates is included.

**Clinical efficacy variables**
In addition to the pre-specified composite endpoints in the protocol, a further composite endpoint composed of first HF hospitalization and CV death is included, as this is the potential primary endpoint for the Phase 3 study. Correspondingly, the component HF hospitalization is included as a further endpoint of interest.

**Length of hospitalization**
Further details on the derivation of length of hospitalization are included.

**Censoring rules for clinical events**
Definition of last visit is clarified.

**Event rates at 12 months**
Two methods for estimating one-year event rates are included.

**Laboratory parameters**
Exploratory testing of selected laboratory parameters is included. Selection of laboratory parameters for AUC analysis is revised.

### 7.2.3 Changes from SAP version 2.0 to version 3.0

**Missing study medication end dates**
Missing study medication end dates have been observed in the database. An imputation rule to deal with these cases has been included.

**Censoring rules for clinical events**
Censoring rules updated to always include rule for non-completers.

### 8. References


5. McMurray JJV, Adamopoulos S, Anker SD, et al. ESC Guideline for the diagnosis and treatment of acute and chronic heart failure 2012. The Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC. European Heart Journal. 2012; 33:1787-1847.


9. Appendix

9.1 KCCQ Scoring

As described in the KCCQ Scoring instruction [7], the following derivations will be used. Generally only questions actually answered are used for derivation of the scores in the following way:

If there are \( n \) questions in a scale, and the subject must answer \( m \) to score the scale, but the subject answers only \( n-i \), where \( n-i \geq m \), calculate the mean of those questions as

\[
\text{mean} = \frac{(\text{sum of the responses to those } n-i \text{ questions})}{(n-i)}
\]

\[
\text{not} \quad \frac{(\text{sum of the responses to those } n-i \text{ questions})}{n}
\]

The 7 individual domain scores and 3 summary scores will be calculated as follows:

**9.1.1 Physical Limitation**

Code responses to each of Questions 1a-f 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, then compute

\[
\text{Physical Limitation Score} = 100 \times \frac{(\text{mean of Questions 1a-f actually answered}) - 1}{4}
\]

**9.1.2 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

\[
\text{Symptom Stability Score} = 100 \times \frac{(\text{Question 2}) - 1}{4}
\]
9.1.3 Symptom Frequency

Code responses to Questions 3, 5, 7 and 9 as follows:

**Question 3**
- 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:

\[
S_3 = \frac{\text{Question 3} - 1}{4} \\
S_5 = \frac{\text{Question 5} - 1}{6} \\
S_7 = \frac{\text{Question 7} - 1}{6} \\
S_9 = \frac{\text{Question 9} - 1}{4}
\]

Symptom Frequency Score = 100*(mean of S3, S5, S7 and S9)

9.1.4 Symptom Burden

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  
\[ \text{Symptom Burden Score} = 100 \times \left( \frac{\text{mean of Questions 4, 6 and 8 actually answered} - 1}{4} \right) \]

### 9.1.5 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  
\[ \text{Self-Efficacy Score} = 100 \times \left( \frac{\text{mean of Questions 10 and 11 actually answered} - 1}{4} \right) \]

### 9.1.6 Quality of Life

Code responses to 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

\[
\text{Quality of Life Score} = 100\times\left(\frac{\text{mean of Questions 12, 13 and 14 actually answered} - 1}{4}\right)
\]

**9.1.7 Social Limitation**

Code responses to each of 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 are not missing, then compute

\[
\text{Social Limitation Score} = 100\times\left(\frac{\text{mean of Questions 15a-d actually answered} - 1}{4}\right)
\]

**9.1.8 Total Symptom Score**

\[= \text{mean of the following available summary scores:}\]

- Symptom Frequency Score
- Symptom Burden Score

**9.1.9 Overall Summary Score**

\[= \text{mean of the following available summary scores:}\]

- Physical Limitation Score
- Total Symptom Score
- Quality of Life Score
Social Limitation Score

9.1.10  Clinical Summary Score

= mean of the following available summary scores:

  - Physical Limitation Score
  - Total Symptom Score

9.2  EQ-5D-3L Scoring

The EQ-5D-3L consists of a 5 dimension descriptive system and the visual analogue scale (VAS). Each of the 5 dimensions of the descriptive system has 3 levels, which are coded as follows:

  - Level 1: indicating no problem = 1
  - Level 2: indicating some problems = 2
  - Level 3: indicating extreme problems = 3

A missing value is coded with 9.

9.2.1  UK Time Trade-Off (TTO) value set

The UK TTO value set is derived as follows: First, a constant term of 1 is included. This represents a state of full health (i.e. level 1 in all 5 dimensions of the descriptive system). Depending on a subject's deviation from the full health state, terms are added for the individual dimensions (see Table 9-1).

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Level 2</th>
<th>Level 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobility (MO)</td>
<td>-0.069</td>
<td>-0.314</td>
</tr>
<tr>
<td>Self care (SC)</td>
<td>-0.104</td>
<td>-0.214</td>
</tr>
<tr>
<td>Usual activities (UA)</td>
<td>-0.036</td>
<td>-0.094</td>
</tr>
<tr>
<td>Pain/discomfort (PD)</td>
<td>-0.123</td>
<td>-0.386</td>
</tr>
<tr>
<td>Anxiety/depression (AD)</td>
<td>-0.071</td>
<td>-0.236</td>
</tr>
</tbody>
</table>

In addition the following variables are assessed:

\[ N2 = I(\text{at least one dimension with level 2 or level 3}) \quad \text{and} \quad N3 = I(\text{at least one dimension with level 3}) \]

I() is the indicator function. Each of them is then multiplied with an individual term (see Table 9-2).
The mathematical representation of the model for health state X is:

\[ X = 1 - 0.081N_2 - 0.069MO_2 - 0.314MO_3 - 0.104SC_2 - 0.214SC_3 - 0.036UA_2 - 0.094UA_3 - 0.123PD_2 - 0.386PD_3 - 0.071AD_2 - 0.236AD_3 - 0.269N_3, \]

where MO_2, MO_3, SC_2, SC_3, UA_2, UA_3, PD_2, PD_3, AD_2, and AD_3 are indicator functions of the corresponding dimension being level 2 or 3, respectively. N_2 is an indicator function of any dimension being level 2 or 3, and N_3 is the indicator function of any dimension being level 3.

As an example, take a subject with the following levels in the 5 dimensions: 2,1,2,3,2. In this example, the indicators MO_2, UA_2, PD_3 and AD_2 take the value 1, while the other indicator functions for the respective dimensions take the value 0. There are several dimensions with levels 2 or 3, thus N_2=1. Also, there is one dimension with level 3, thus N_3=1. This leads to the overall value

\[
\text{State } 2,1,2,3,2 = 1 - 0.081*1 - 0.069*1 - 0.314*0 - 0.104*0 - 0.214*0 - 0.036*1 - 0.094*0 - 0.123*0 - 0.386*1 - 0.071*1 - 0.236*0 - 0.269*1 \\
= 1 - 0.081 - 0.069 - 0.036 - 0.386 - 0.071 - 0.269 \\
= 0.088
\]

9.2.2 US TTO value set

The US TTO value set is derived as follows: First, a constant term of 1 is included. This represents a state of full health (i.e. level 1 in all 5 dimensions of the descriptive system). Depending on a subject's deviation from the full health state, terms are added for the individual dimensions (see Table 9-3).

![Table 9-3: US TTO terms for individual dimensions](image)
In addition, the following variables are assessed:

\[ D_1 = (\text{number of dimensions with level 2 or 3}) - 1 \]
\[ I_2\text{-square} = ((\text{number of dimensions at level 2}) - 1)^2 \]
\[ I_3 = (\text{number of dimensions at level 3}) - 1 \]
\[ I_3\text{-square} = ((\text{number of dimensions at level 3}) - 1)^2 \]

Each of them is then multiplied with an individual term (see Table 9-4).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>0.140</td>
</tr>
<tr>
<td>I2-square</td>
<td>-0.011</td>
</tr>
<tr>
<td>I3</td>
<td>0.122</td>
</tr>
<tr>
<td>I3-square</td>
<td>0.015</td>
</tr>
</tbody>
</table>

The mathematical representation for a health state \( X \) is thus:

\[
X = 1 - 0.146 \text{MO}_2 - 0.558 \text{MO}_3 - 0.175 \text{SC}_2 - 0.471 \text{SC}_3 - 0.140 \text{UA}_2 - 0.374 \text{UA}_3 - 0.173 \text{PD}_2 - 0.537 \text{PD}_3 - 0.156 \text{AD}_2 - 0.450\text{AD}_3 + 0.140 D_1 - 0.011 I_2\text{-square} + 0.122 I_3 + 0.015 I_3\text{-square},
\]

where \( \text{MO}_2, \text{MO}_3, \text{SC}_2, \text{SC}_3, \text{UA}_2, \text{UA}_3, \text{PD}_2, \text{PD}_3, \text{AD}_2, \) and \( \text{AD}_3 \) are indicator functions of the corresponding dimension being level 2 or 3, respectively.

As an example, take a subject with the following levels in the 5 dimensions: 2,1,2,3,2. In this example, the indicators \( \text{MO}_2, \text{UA}_2, \text{PD}_3 \) and \( \text{AD}_2 \) take the value 1, while the other indicator functions for the respective dimensions take the value 0. As four dimensions have a level of 2 or 3, the term \( D_1=3 \). Three dimensions have the value 2, thus \( I_2\text{-square} = (3-1)^2 = 4 \). Only one dimension has level 3, thus \( I_3=0 \) and \( I_3\text{-square}=0 \), respectively. This leads to the overall value

\[
\text{State } 2,1,2,3,2 = 1 - 0.146 \cdot 1 - 0.558 \cdot 0 - 0.175 \cdot 1 - 0.471 \cdot 0 - 0.140 \cdot 1 - 0.374 \cdot 0 - 0.173 \cdot 0 - 0.537 \cdot 1 - 0.156 \cdot 1 - 0.450 \cdot 0 + 0.140 \cdot 3 - 0.011 \cdot 4 + 0.122 \cdot 0 + 0.015 \cdot 0 = 1 - 0.146 - 0.140 - 0.537 - 0.156 + 0.140 \cdot 3 - 0.011 \cdot 4 = 0.397
\]

### 9.3 Composite congestion score

The composite congestion score [14] is composed out of the following three components:

**Orthopnea**
To be coded as follows:
None = 0
Seldom = 1
Frequent = 2
Continuous = 3

**Jugular venous distension(cmH₂O)**
The coding instruction for jugular venous distension given by Ambrosy et al. were ambiguous. Therefore, the adapted coding will be as follows:
- <6 = 0
- 6-9 = 1
- 10-15 = 2
- >15 = 3

**Edema**
To be coded as follows:
- Absent/trace = 0
- Slight = 1
- Moderate = 2
- Marked = 3

The composite congestion score is the sum of codes of these three components:

Orthopnea + Jugular venous distension + Edema

In case one of the three components is missing, the composite congestion score will also be missing.