NOVARTIS

Clinical Development

LCZ696/Entresto®

CLCZ696B2319 / NCT02678312

Multicenter, open-label, study to evaluate safety, tolerability, pharmacokinetics and, pharmacodynamics of LCZ696 followed by a 52-week randomized, double-blind, parallel group, active-controlled study to evaluate the efficacy and safety of LCZ696 compared with enalapril in pediatric patients from 1 month to < 18 years of age with heart failure due to systemic left ventricle systolic dysfunction

Statistical Analysis Plan (SAP)

Author: Trial Statistician, [redacted]

Document type: SAP Documentation

Document status: Final

Release date: 29-Oct-2021

Number of pages: 73

Document History – Changes compared to previous final version of SAP

Date and SAP version	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
23-Nov-2016 Part 1	Prior to DB	Initial	NA	NA
03-Mar-2017 Part 2	Prior to DB	Initial	NA	NA
27-Oct-2020 Part 1 Amendment 1	Prior to DB	Protocol amendment	Updated per protocol amendment 6	All sections
03-Nov-2020 Part 2 Amendment 1	Prior to DB	Protocol amendment	Updated per protocol amendment 6	All sections
29-Oct-2021 Combined Part 1 and Part 2 Amendment 2	Prior to DB	Combine Part 1 and Part 2 SAP	Part 1 and Part 2 combined	All sections
		Part 1 period covers long duration	Part 1 period changed: remove treatment period, dose associated treatment period; added the dose associated period to be two visits whichever occurred first	2.1.1
		Subgroup changes	NYHA/ROSS at baseline added; COVID-19 period added; Baseline eGFR removed; Race added; Prior ACEI/ARB use added.	2.2.1
		COVID-19 related analysis added	COVID-19 disposition, PD summary added	2.3.1
		Demographics added	age adjusted percentile parameters added	2.3.2
		Compliance category added	Added < 90%, 90% to < 110%, >110%.	2.4.1
		Adult dose level specified clear	Adult dose (≥ 57 kg)	2.4.1

Date and SAP version	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
		Randomization date changed to first dose date for time to first dose and permanent discontinuation of the drug	Randomization date changed to first dose date	2.4.1
		Clarification for the prior medication and concomitant medication definition	Changed to the general description.	2.4.2
		SGLT2i added	SGLT2i ATC code added	2.4.2
		Analysis cutoff date added	analysis cutoff date will be set to 58 weeks, i.e. 406 days	2.6.1 2.6.3 2.6.4 2.7.2 2.7.3
		[redacted]	[redacted]	[redacted]
		Missing baseline and post-baseline assessment handling for PGIS added for global rank	While ranking, if the baseline is missing, then the baseline will be imputed using the median assessment within each stratum. If all the post-randomization assessments is missing, then the post randomization assessment will be imputed to the worst case at week 52	2.6.3
		Supportive analysis for primary analysis added	Add supportive analysis	2.6.4
		Randomization stratification(NYHA/ROSS group) added as covariate	Cox proportional hazard model, MMRM, proportional cumulative odds model, LWYY added covariate for randomization stratification	2.7.1 2.7.2 2.7.3 2.8.1 2.8.2 2.8.3 2.8.4 4
		[redacted]	[redacted]	[redacted]
		Safety summary changes	Part 2: 4 risks expanded to AESI; Added summary by SMQ; added baseline eGFR ≤ / > 170 mL/min/1.73 m^2 seperate analysis for renal event;	2.9

Date and SAP version	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
			Clinical notable criteria for laboratory values updated; Central lab and local lab combined for events summary; Change from baseline in height Z-score added;	
		PK and PD analysis description adjustments		2.5.1 2.5.2 2.10 2.11
		[redacted]	[redacted]	[redacted]
		PD for PPS updated	PD for PPS updated	5.3
		Implementation of an Urgent Safety Measure (USM)	Analysis updated per USM	2.6 2.7 2.8

Table of contents

List of abbreviations

AE	Adverse Event
ADaM	Analysis Data Model
AGB	ADaM Governance Board
ATC	Anatomical Therapeutic Classification
AUC	Area Under the Curve
bid	bis in diem/twice a day
CRF	Case Report Form
CSR	Clinical Study Report
CV	Coefficient of variation
DMC	Data Monitoring Committee
LLOQ	Lowest Limit of Quantitation
LVEF	Left ventricular ejection fraction
MedDRA	Medical Dictionary for Drug Regulatory Affairs
NOVDTD	Novartis Drug and Therapy Dictionary
PK	Pharmacokinetics
PD	Pharmacodynamics
PDS	Programming datasets specifications
SAP	Statistical Analysis Plan
SOC	System Organ Class
UNS PK/PD	Unscheduled PK/PD
ULOQ	Upper Limit of Quantitation
USM	Urgent Safety Measure

1 Introduction

This document includes the analysis details of study CLCZ696B2319. Data will be analyzed according to CLCZ696B2319 protocol v06 Section 9 (Data analysis). Statistical methods are described in the following Section 2 (Statistical methods) and sample size considerations are provided in the following Section 3 (Sample size calculation).

The analysis results will be summarized in the clinical study report (CSR).

1.1 Study design

This study uses a seamless design which consists of two parts.

Part 1 is a multi-center, open-label, study to evaluate safety, tolerability, pharmacokinetics (PK) and pharmacodynamics (PD) of LCZ696 in pediatric patients (1 month to <18 years) with heart failure due to systemic left ventricle systolic dysfunction (left ventricular ejection fraction [LVEF] ≤ 40% or fractional shortening ≤ 20%). Eligible patients will be placed into three age groups (Age Group 1: 6 years to < 18 years, Age Group 2: 1 year to < 6 years, and Age Group 3: 1 month to < 1 year). Patient enrollment will start sequentially from the eldest age group to the youngest age group.

For each age group, available PK/PD and safety data will be reviewed to confirm or modify the doses. After completion of the PK/PD assessment visit, patients can either be maintained on open-label enalapril or their standard of care HF medical regimen until Part 2. Patients in each age group can enroll in Part 2 after the target dose for that age group is determined based on Part 1 data for the corresponding age cohort. Patients that discontinue from Part 1 (PK/PD) will be allowed to screen and enroll for participation in Part 2.

Part 2: This is a randomized, double-blind, parallel-group, active controlled, 52-week study to evaluate the efficacy, safety, and tolerability of LCZ696 compared to enalapril in pediatric HF patients (1 month to < 18 years). A screening epoch of up to 3 weeks will be used to assess eligibility. Three hundred and sixty eligible patients will be randomized to one of the two treatment arms (LCZ696 vs. enalapril) and continue treatment for 52 weeks duration. Both hospitalized patients and outpatients are eligible. Chronic HF patients that are either previously treated for HF or newly diagnosed are eligible.

An interim efficacy analysis is planned to be performed when at least 180 patients (at least 36 patients from each age group) have completed the study (i.e., reached a positively adjudicated event in Category 1 or completed the 1 year study visit), and at least 40 patients have had an event in Category 1 or 2. Further details about interim analysis are provided in Section 2.15.

If not stopped early for efficacy or futility based on the interim analysis, Part 2 will continue until at least 360 patients have completed the study and at least 80 patients have an event in Category 1 or 2. If necessary, study enrollment will continue beyond 360 patients in order to obtain at least 80 patients with an event in Category 1 or 2, based on event prediction.

A penalty-free blinded sample size adjustment may be made in the latter part of the study to ensure a power of 80% for the study.

1.2 Study objectives and endpoints

Table 1-1 summarizes the study objectives and the endpoints for the objectives.

Table 1-1 Objectives and related endpoints

OBJECTIVE	Endpoint Title, Description and Reporting Time Frame for analysis and Unit of Measure	Section
Part 1		
Primary		
To determine the pharmacokinetics (PK) and pharmacodynamics (PD) of LCZ696 in pediatric HF patients	*Title:* PK and PD of LCZ696 after single dose treatment *Unit of measure*: PK: C_{max} (ng/mL); T_{max} (h); AUC_{last}, $AUCi_{nf}$ (h▪ng/mL); Cl/F (L/h); $T_{1/2}$ (h); PD: plasma BNP, plasma NTproBNP, plasma cGMP, urine cGMP change from baseline geometric mean ratio (GMR) after single dose treatment *Description*: Part 1 PK assessment: The samples will be collected at the following intervals for the various age groups: • Patients ≥6 years of age: Blood samples are collected at pre-dose (pre-dose sample for all patients), 0.5, 1, 2, 4, 8, 10, and optional 24 hours post dosing • Patients < 6 years of age: Blood samples are collected at pre-dose (pre-dose sample only from those patients who are on valsartan), 1, 2, 4, 10, and optional 24 hours post dosing Part 1 PD assessments: The blood samples will be collected at pre-dose, 4, 8 and optional 24 hours post dosing. The urine samples will be collected at pre-dose and once between 4 to 8 hours post dosing. *Time Frame*: Single dose PK/PD visit	Section 2.5
Secondary		
To assess the safety and tolerability of LCZ696 in pediatric patients with HF	*Title:* Safety and tolerability of LCZ696 in Part 1 *Unit of Measure:* Multiple *Description:* Safety and tolerability including AEs, laboratory, vital signs *Time Frame:* Single dose PK/PD visit	Section 2.9
Part 2		
Primary		
To determine whether LCZ696 is superior to enalapril for the treatment of HF as assessed	*Title*: Global Rank endpoint through 52 weeks of treatment	Section 2.6

OBJECTIVE	Endpoint Title, Description and Reporting Time Frame for analysis and Unit of Measure	Section
using a global rank endpoint in pediatric HF patients	*Unit of Measure*: The mean global rank endpoint does not have a unit. Patients are ranked from worst to best. The ranking is based on clinical events such as death, listing for urgent heart transplant, mechanical life support requirement at end of study, worsening HF, NYHA/ROSS, patient global impression of severity (PGIS), pediatric quality of life (PedsQL) physical functioning domain. *Description*: The effects of LCZ696 and enalapril will be assessed using the Global Rank endpoint as outlined below through 52 weeks of double-blind treatment. The primary endpoint will be derived based on 5 categories ranking worst to best outcome: • Category 1: Death; UNOS status 1A listing for heart transplant or equivalent; VAD/ECMO/mechanical ventilation/intra-aortic balloon pump requirement for life support at end of study • Category 2: Worsening HF (WHF); defined by signs and symptoms of WHF that require an intensification of HF therapy • Category 3: Worsened; worse NYHA/ROSS or worse PGIS; and further ranking by PedsQL physical functioning domain • Category 4: Unchanged; unchanged NYHA/ROSS and unchanged PGIS; and further ranking by PedsQL physical functioning domain • Category 5: Improved; improved NYHA/ROSS or improved PGIS (neither can be worse); and further ranking by PedsQL physical functioning domain Within each category, patients are ranked from worst to best based on pre-defined criteria. PedsQL physical functioning will only be used for the global rank endpoint for Age Group 1. *Time Frame*: 52 Weeks	
Secondary		
To determine whether LCZ696 is superior to enalapril in delaying time to first occurrence of the composite of either Category 1 or 2 events (e.g. death, worsening HF)	*Title*: Time to first occurrence of Category 1 or Category 2 event through 52 weeks of treatment *Unit of Measure*: Days	Section 0

OBJECTIVE	Endpoint Title, Description and Reporting Time Frame for analysis and Unit of Measure	Section
	Description: Time to first occurrence of Category 1 or Category 2 event will be compared for LCZ696 and enalapril through 52 weeks of double-blind treatment *Time Frame*: 52 Weeks	
To determine whether LCZ696 is superior to enalapril for improving NYHA/ROSS functional class	*Title*: NYHA/ROSS functional class change from baseline through 52 weeks of treatment *Unit of Measure*: NYHA/ROSS classification *Description*: NYHA/ROSS functional class will be compared through 52 weeks of double-blind treatment *Time Frame*: 52 Weeks	Section 2.7.2
To determine whether LCZ696 is superior to enalapril for improving the PGIS score	*Title*: PGIS score change from baseline through 52 weeks of treatment *Unit of Measure*: PGIS scale *Description*: PGIS scale will be compared for LCZ696 and enalapril through 52 weeks of double-blind treatment *Time Frame*: 52 weeks	Section 2.7.3
To characterize the population PK of LCZ696 exposure in pediatric patients with HF	*Title*: Population PK of LCZ696 *Unit of Measure*: ng/mL/kg, ng*h/mL/kg, L/h/kg, L/kg, and 1/h *Description*: During Part 2 Efficacy, population PK will be assessed with chronic dosing. Population PK allows us to estimate clearance and total exposure. *Time frame*: week 2, 12, 52	Separate document
To assess the safety and tolerability of LCZ696 compared to enalapril in pediatric patients with HF	*Title:* Safety and tolerability through 52 weeks of treatment *Unit of Measure:* Multiple *Description:* Safety and tolerability including AEs, laboratory, ECG, vital signs data through 52 weeks of double-blind treatment *Time Frame:* 52 weeks	Section 2.9
[redacted]		

OBJECTIVE	Endpoint Title, Description and Reporting Time Frame for analysis and Unit of Measure	Section
[redacted]		

2 Statistical methods

2.1 Data analysis general information

Unless otherwise specified, data will be analyzed according to the study protocol using SAS 9.4 or higher.

An external DMC will monitor patient safety data during the course of the study. For this study, the DMC will review safety data on a regular frequency of every six months. DMC may request additional safety data review.

In addition, the safety, tolerability, pharmacokinetic, and pharmacodynamic data from open label use of LCZ696 will be evaluated by Novartis and the DMC to monitor patients safety for each age group when all the patients have completed Part 1 of the study in that age group.

Besides, one formal interim efficacy analysis will be conducted for Part 2 when at least 180 patients (at least 36 patients from each age group) have completed the study (i.e., had a positively adjudicated event in Category 1 or completed the double-blind epoch), and at least 40 patients have had an event in Category 1 or 2. Further details can be found in Section 2.15.

The analyses results for Part 1 and Part 2 will be summarized in the CSR.

In general, continuous variables will be summarized using number of observations, mean, standard deviation, median, quartiles, minimum and maximum. Categorical variables will be summarized using frequencies and percentages.

Unless otherwise specified, for Part 2, the analyses will include a summary for all the three modified age groups (6 years to < 18 years, 2 years to < 6 years, and 1 month to < 2 years) and a summary for overall; the efficacy variables will be analyzed based on the full analysis set (FAS), the safety variables will be analyzed based on the Part 2 safety set (SAF2); statistical testing for the secondary [redacted] variables will be performed at the significance level of two-sided 0.05. No multiplicity adjustment will be performed for the secondary [redacted] variables.

2.1.1 General definitions

Part 1:

Epochs

The period 1 epoch is defined as the period from the date of Visit 101 to the date of Visit 199.

The period 2 epoch is defined as the period from the date of Visit 201 to the date of Visit 299.

Dose cohorts

Dose cohort 1 is defined as the patients who were prescribed at least one dose of the study drug at dose level 0.4 mg/kg or 0.8 mg/kg within period 1.

Dose cohort 2 is defined as the patients who were prescribed at least one dose of the study drug at dose level 1.6 mg/kg or 3.1 mg/kg within period 2.

Dose cohort S is defined as patients who were prescribed one dose of study drug at dose level 3.1 mg/kg at Visit 201 and were prescribed another dose of study drug at a different dose level within period 2. This definition comes from the special cases when trial is ongoing.

Of note, it is possible that, a patient is in both dose cohort 1 and dose cohort 2.

Dose associated periods

For dose cohort 1, the dose associated period is defined as date of Visit 101 to the date of Visit 103 or Visit 199, whichever occurred first.

For dose cohort 2, the dose associated period is defined as date of Visit 201 to the date of Visit 203 or Visit 299, whichever occurred first.

Baseline assessment and period day

Unless otherwise specified, the baseline assessment for period 1 is defined as the last non-missing assessment (scheduled or unscheduled) prior to the first dose time of the study drug within period 1.

Unless otherwise specified, the baseline assessment for period 2 is defined as the last non-missing assessment (scheduled or unscheduled) prior to the first dose time of the study drug within period 2, if visit 199 and 201 happened on the same day, then the visit 199 is included.

In the Part 1 listings, the period and the period day will be displayed, the period day is defined as the date of assessment (event/visit) minus the date of first dose plus one.

Post-baseline assessment

Unless otherwise specified, the post-baseline assessments during period 1 are defined as those assessments taken during period 1 and later than the first dose time of the study drug within period 1, if visit 199 and 201 happened on the same day, then the visit 201 is included.

Unless otherwise specified, the post-baseline assessments during period 2 are defined as those assessments taken during period 2 and later than the first dose time of the study drug within period 2.

Age and Age group

For Part 1, the age at screening will be recorded on the Demography page and the age group (6 years to < 18 years, 1 year to < 6 years, 1 month to < 1 year) will be calculated based on the age at screening (Visit 1).

Unscheduled visit

Safety evaluation for unscheduled measurements will be taken into account as appropriate.

PK/PD evaluation for an unscheduled PK/PD (UNS PK/PD) visit will be mapped into Visit 101/201 as appropriate.

Unscheduled assessment will be included in the listings.

Dose associated analysis visits

In Part 1, visits in both period 1 and period 2 will be mapped into dose associated analysis visits , for example, visit 101 in period 1, visit 201 in period 2, the dose associated analysis visit is visit 101/201. The dose associated analysis visits will be used in by-visit summaries.

Part 2:

Double-blind Epoch

The double-blind epoch is defined as the period from the date of randomization to the end of the study, inclusive.

Baseline assessment and study day

Unless otherwise specified, the baseline assessment is defined as below.

- For efficacy variables, if the scheduled assessment at visit 401 is performed and the assessment value is non-missing, the baseline assessment is defined as the scheduled assessment at visit 401.
- For efficacy variables, if the scheduled assessment at visit 401 is not performed or the assessment value is missing, the baseline assessment is defined as the last non-missing assessment (scheduled or unscheduled) prior to or on the visit 401 date and after last part 1 dose date plus five days.

- For safety variables, the baseline assessment is defined as the last non-missing assessment (scheduled or unscheduled) prior to or at the time of the first dose of double-blind study drug (i.e., study drug within the double-blind epoch) and after last part 1 dose date plus five days. When the dose and the assessment are on the same date, and the dosing time or the assessment time is missing, the assessment is considered as pre-dose.

For efficacy variables, the study day is defined as the date of assessment (event/visit) minus the date of randomization plus one, i.e., Visit 401 is Day 1.

For safety variables, the study day will be defined as the date of assessment (event/visit) minus the date of first dose of double-blind study medication plus one day, i.e., date of first dose of double-blind study medication is double-blind day 1.

Post-baseline assessment

Unless otherwise specified, for efficacy variables, the post-baseline assessments during the double-blind epoch are defined as those assessments taken during the double-blind epoch and later than the date of randomization; for safety variables, the post-baseline assessments during the double-blind epoch are defined as those assessments taken during the double-blind epoch and later than the first dose time of the study drug within the double-blind epoch.

Age, age group, modified age group and NYHA/ROSS class group at randomization

The study is randomized stratifying by age group (6 years to < 18 years, 1 year to < 6 years, 1 month to < 1 year) and NYHA/ROSS class group (Class I/II, Class III/IV) at randomization.

The age and NYHA/ROSS class at randomization will be recorded on the case report form (CRF) page. The age "Subject's Age at Randomization" will be used to calculate the age group (6 years to < 18 years, 1 year to < 6 years, 1 month to < 1 year) and modified age group (6 years to < 18 years, 2 years to < 6 years, 1 month to < 2 years).

If the information is missing on CRF page, stratification data for age group (6 years to < 18 years, 1 year to < 6 years, 1 month to < 1 year) and NYHA/ROSS class group (Class I/II, Class III/IV) at randomization will be loaded from the randomization data (Interactive Response Technology (IRT) system) to impute the age, age group, modified age group and NYHA/ROSS class. Week, Month and Year

In general, 1 week = 7 days, 1 month = 30.4375 days and 1 year = 365.25 days.

Visit 499 mapping

For patients who completed the double-blind epoch, Visit 499 will be Week 52.

For patients who discontinued from the study during the double-blind epoch, Visit 499 will be mapped according to Table 2-1.

Table 2-1 Visit 499 mapping crietria

Double-Blind Epoch Patient Status	Last Attended Scheduled Visit Prior to Visit 499		Country	Mapped Visit for Visit 499	
	Visit Number	Visit Name		Visit Number	Visit Name
Discontinued	401	Randomization	Japan	401.1	Week 1
			Other	402	Week 2
Discontinued	401.1	Week 1	Japan	402	Week 2
Discontinued	402	Week 2	All	403	Week 4
Discontinued	403	Week 4	All	404	Week 6
Discontinued	404	Week 6	All	405	Week 8
Discontinued	405	Week 8	All	406	Week 12
Discontinued	406	Week 12	All	407	Week 16
Discontinued	407	Week 16	All	408	Week 20
Discontinued	408	Week 20	All	409	Week 24
Discontinued	409	Week 24	All	410	Week 28
Discontinued	410	Week 28	All	411	Week 32
Discontinued	411	Week 32	All	412	Week 36
Discontinued	412	Week 36	All	413	Week 40
Discontinued	413	Week 40	All	414	Week 44
Discontinued	414	Week 44	All	415	Week 48
Discontinued	415	Week 48	All	416	Week 52
Completed	---	---	All	416	Week 52

- For each patient, the last attended scheduled visit prior to Visit 499 is defined as the visit with maximal visit number among all scheduled visits (Visit 401 to Visit 415) in the visit panel.

Unscheduled assessments

Data collected at Visit 499 will be included in by-visit summaries, if an assessment were scheduled at the mapped visit for Visit 499.

Unscheduled assessments will be included in over period minimum/maximum, and in over period shift table.

Unless otherwise specified, in by-visit summaries, the unscheduled assessments will be used as below.

- For efficacy variables, if a post-randomization scheduled assessment is not performed or the assessment value is missing, the last non-missing associated unscheduled assessment prior to the next scheduled visit and within 14 days of the target date could be used to impute the scheduled assessment.
- For safety variables, if a post-randomization scheduled assessment is not performed or the assessment value is missing, the last non-missing associated unscheduled assessment prior to the next scheduled visit could be used to impute the scheduled assessment.

An unscheduled assessment is considered to be associated with a scheduled assessment if the unscheduled visit number and the scheduled visit number are equal up to the first digit, e.g., visit 40x.00y is associated with visit 40x, and visit 401.10x is associated with visit 401.1.

Region

Region will be derived from country. The mapping criteria are specified below.

Table 2-2 Region mapping criteria

Region	Countries
America	Argentina, Canada, USA
Europe	Austria, Bulgaria, Croatia, Czech Republic, Finland, France, Germany, Hungary, Italy, Poland, Portugal, Romania, Russia, Spain, Sweden, Switzerland, Turkey
Asia/Pacific and other	China, India, Israel, Japan, Jordan, Lebanon, Rep. of Korea, Saudi Arabia, Singapore, South Africa, Taiwan, Thailand

2.2 Analysis sets

The following analysis sets will be defined for statistical analysis for Part 1 and Part 2.

- **Part 1 screened set (SCR1)** – All patients who signed the informed consent for Part 1. The SCR1 includes only *unique screened patients*, i.e., in the case of *re-screened patients* only the chronologically last screening data is counted.
- **Part 1 eligible set (ELG1)** – All patients who completed the Part 1 screening phase.
- **Part 1 PK set (PK1)** – All ELG1 patients with at least one dose of study drug during Part 1 of the study, at least one available valid (i.e., not flagged for exclusion) PK concentration measurement during Part 1 of the study, and with no protocol deviations with relevant impact on PK data during Part 1 of the study. Patients will be analyzed according to treatment received.
- **Part 1 PD set (PD1) –** All ELG1 patients with at least one dose of study drug during Part 1 of the study, at least one available PD measurements during Part 1 of the study and with no protocol deviations with relevant impact on PD data. Patients will be analyzed according to treatment received.
- **Part 1 safety set (SAF1)** – All ELG1 patients who received at least one dose of study drug during the Part 1 of the study. Patients will be analyzed according to treatment received.
- **Part 2 screened set (SCR2)** – All patients who signed the informed consent for Part 2. The SCR2 includes only *unique screened patients*, i.e., in the case of *re-screened patients* only the chronologically last screening data is counted.
- **Randomized set (RAN)** – All patients who received a randomization number, regardless of receiving study drug.
- **Full analysis set (FAS)** – All randomized patients with the exception of those patients who have not been qualified for randomization and have not received study drug, but have been inadvertently randomized into the study. Following the intent-to-treat principle, for the efficacy analysis, patients will be analyzed according to the treatment assigned at randomization. Efficacy variables will be analyzed based on the FAS as the primary set.
- **Per protocol set (PPS)** – A subset of the FAS which consists of the patients who do not have major deviations from the protocol procedures in the double-blind study stage. Major protocol deviations will be pre-specified prior to unblinding treatment codes for analyses. This supplemental efficacy set will be used to support the primary analysis results.

- **Part 2 safety set (SAF2)** – All randomized patients who received at least one dose of study drug. For the safety analysis, patients will be analyzed according to the treatment actually received. The treatment actually received will be considered identical to the randomized treatment if the patient has received at least one dose of the randomized treatment. The SAF2 will be used for the safety analyses for Part 2.

Rules leading to exclusion from analysis sets are given in Appendix 5.3.

2.2.1 Subgroup of interest

Table 2-3 specifies the subgroups to be used in subgroup analyses for Part 2.

Table 2-4 summarizes the subgroup analyses for Part 2.

Table 2-3 Subgroup specifications

Subgroups	Source		
	Data Source	Variable	Visit/Time Point
Modified age group at randomization • 6 years to < 18 years; • 2 years to < 6 years; • 1 month to < 2 years;	CRF Page: Subject's Age at Randomization	Age	Randomization
NYHA/ROSS class groupat baseline • NYHA/ROSS Class I/II; • NYHA/ROSS Class III/IV;	CRF Page: Signs and Symptoms of Heart Failure including New York Heart Association Classification; Modified ROSS HF Classification	NYHA/ROSS class group	Randomization
NYHA/ROSS class group at randomization • NYHA/ROSS Class I/II; • NYHA/ROSS Class III/IV;	IRT: Randomization stratification	NYHA/ROSS class group	Randomization
Gender • Female; • Male;	CRF Page: Demography	Sex	Screening or Pre-Randomization
Region • America; • Europe; • Asia/Pacific and other;	CRF Page: Demography	Country	Screening or Pre-Randomization
COVID-19 period • Pre-pandemic: EOS prior to 1-Mar-2020 • Pre and during pandemic:	CRF page: Disposition; YR domain	date	Randomization date End of study date

Subgroups	Source		
	Data Source	Variable	Visit/Time Point
Randomization date prior to 1-Mar-2020 and EOS after 1-Mar-2020 • During pandemic: Randomization date on or after 1-Mar-2020			
Race • Caucasian: White. • Black: Black or African American. • Asian: Asian. • Other: American Indian or Alaska Native; Unknown; Other.	CRF Page: Demography	Race	Screening or Pre-Randomization
Prior ACEI/ARB use • ACEI only • ARB only • ACEI and ARB both • No • Unknown	CRF Page: Prior Cardiovascular Medications-ACEI Prior Cardiovascular Medications-ARB	medication	Screening or Pre-Randomization

Table 2-4 Subgroup analyses

Subgroup Variable	Primary/Secondary Efficacy	Selected Safety
Modified age group at randomization	X	X
NYHA/ROSS class group at randomization	X	
NYHA/ROSS class group at baseline	X	
Gender	X	X
Region	X	X
COVID-19 impacted period	X	
Race	X	X
Prior ACEI/ARB use		X

2.3 Patient disposition, demographics and other baseline characteristics

2.3.1 Patient disposition

Part 1:

The number and percentage of patients who completed the Part 1 screening phase will be provided by age group. In addition, the primary reason for not completing the Part 1 screening phase will be summarized by age group using the numbers and percentages of patients not

qualifying for such reasons. For patient who was screened more than once, the information from the last screen will be used in the summary. These summaries will be performed for the SCR1.

The patient status for the dose associated period will be summarized by age group and dose cohort, using the numbers and percentages of patients who took at least one dose during the period, who completed the period, and who discontinued from study during the period; the primary reason for discontinuation will be summarized by age group, using the numbers and percentages of patients discontinued for such reasons. These summaries will be performed based on the ELG1.

The inclusion status for the analysis sets will be summarized by age group and dose cohort using the numbers and percentages of patients included in different analysis sets. These summaries will be performed based on the ELG1.

The protocol deviations will be summarized by age group and dose cohort using the numbers and percentages of patients with the protocol deviation. These summaries will be performed based on the ELG1.

In addition, the number and percentage of patients satisfying each criteria leading to exclusion from analysis sets will be provided by age group and dose cohort. These summaries will be performed based on the ELG1.

Part 2:

The number and percentage of patients who completed the pre-randomization epoch will be provided. In addition, the primary reason for not completing the pre-randomization epoch will be summarized by age group and modified age group using the numbers and percentages of patients not qualifying for such reasons. For patient who was screened more than once, the information from the last screen will be used in the summary. These summaries will be performed for the SCR2.

The patient status for the double-blind epoch will be summarized by age group/modified age group and treatment group, using the numbers of patients who were randomized, the numbers and percentages of patients who completed the double-blind epoch, who discontinued from the study during the double-blind epoch; the primary reason for discontinuation from the study will be summarized by age group/modified age group and treatment group, using the numbers and percentages of patients discontinued for such reasons. These summaries will be performed based on the RAN.

The numbers and percentages of patients who took at least one dose of the study drug during the double-blind epoch, who completed the study drug during the double-blind epoch and who permanently discontinued from the study drug during the double-blind epoch will be summarized by age group/modified age group and treatment group; the primary reason for permanent discontinuation from the study drug will be summarized by age group/modified age group and treatment group, using the numbers and percentages of patients discontinued for such reasons. These summaries will be performed based on the RAN.

The disposition will also be summarized by COVID-19 periods.

The inclusion status for the analysis sets (section 2.2) will be summarized by age group/modified age group and treatment group using the numbers and percentages of patients included in different analysis sets. These summaries will be performed based on the RAN.

The protocol deviations will be summarized by age group/modified age group and treatment group using the numbers and percentages of patients with the protocol deviation. This summary will also be summarized by COVID-19 periods. These summaries will be performed based on the RAN.

The COVID-19 related PD will also be summarized separately.

In addition, the number and percentage of patients satisfying each criteria leading to exclusion from analysis sets will be provided by age group/modified age group and treatment group. These summaries will be performed based on the RAN.

Furthermore, the numbers and percentages of patients in each country and in each region will be provided by age group/modified age group and treatment group. These summaries will be performed based on the RAN.

2.3.2 Background and demographic characteristics

Continuous variables will be summarized using number of observations, mean, standard deviation, median, minimum, the first quartile (Q1), the third quartile (Q3), and maximum. For in-text tables, minimum and maximum will not be presented. Categorical variables will be summarized using frequencies and percentages.

Part 1:

Summary statistics will be provided by age group and dose cohort for demographics and baseline characteristics including age, age group (not for within age group summary), sex, race, ethnicity, (age adjusted percentile) weight, height, body mass index (BMI), head circumference, category of prior HF medication (ARB, ACEI), HF etiology, prior HF hospitalization, vital signs, and biochemical characteristics, NYHA/Ross classification. BMI will be calculated as weight (kg) / height^2 (m^2) from the collected height and weight at screening (Visit 1).

The SAF1 will be used for the above Part 1 analyses. Baseline assessment will be the one for each period (and then each age group).

Part 2:

Summary statistics will be provided by age group/modified age group and treatment group for demographics and baseline characteristics including age, age group/modified age group (not for within age group/modified age group summary), sex, race, ethnicity, (age adjusted percentile) weight, height, body mass index (BMI), head circumference, category of prior HF medication (ARB, ACEI), HF etiology, prior HF hospitalization, vital signs, and biochemical characteristics, NYHA/ROSS classification,. BMI will be calculated as weight (kg) / height^2 (m^2) from the collected height and weight at Visit 301.

The FAS will be used for the above Part 2 analyses.

2.3.3 Medical history

Any condition entered on the relevant medical history/current medical conditions CRF will be coded using the MedDRA dictionary. Medical history includes heart failure history and other medical history in this study, which are collected at Visit 1 or Visit 301. The number and percentage of patients with each medical condition will be provided by age group/modified age group, treatment group, system organ class (SOC), and preferred term (PT).

For above analyses, the SAF1 will be used for Part 1, and FAS will be used for Part 2.

2.4 Treatments (study treatment, rescue medication, concomitant therapies, compliance)

2.4.1 Study treatment / compliance

Part 1:

The total actual dose of the study drug during a given period is defined as the sum of all doses of the study drug actually administered (recorded on the DAR-PK page, in unit mg) from the start date of the period to the end date of the period.

The planned dose of the study drug at a given visit is defined as the product of the dose prescribed of the study drug (recorded on the DAR-PK page, in unit mg/kg) at the visit and the weight (recorded on the Vital Signs page, in unit kg) at the visit.

The total planned dose of the study drug during a given period is defined as the sum of all planned doses of the study drug from the start date of the period to the end date of the period.

The compliance of the study drug during a given period is defined as the total actual dose of the study drug during the period divided by the total planned dose of the study drug during the period.

The total actual dose, the total planned dose and the compliance during the dose associated period and at the last visit will be summarized by age group and dose cohort using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum.

In addition, the compliance during the dose associated period will be summarized by age group and dose cohort using the numbers and percentages of patients with compliance < 80%, 80% to < 120%, >120% and < 90%, 90% to < 110%, >110%.

The SAF1 will be used for the above analyses.

The study drug administration and prescription will be listed based on SAF1.

Part 2:

The study drug administration will be recorded on the CRF page: “Dosage Administration Record – Double-Blind Study Medication” using start date-time, end date-time and dispensing level of the study drug. Each pair of start date and end date will be considered as a dosing interval.

For each patient, the dosing intervals will be sorted according to the start date and the end date from the earliest to the latest. It is expected that, for each patient, there should be no gaps and

overlaps among dosing intervals, with the exception that the end date of a dosing interval can be the same as the start date of the next dosing interval, when the patient may take the previous dispensing level in the morning and start a different dispensing level in the evening.

In the presence of gaps or overlaps, a data conventional imputation algorithm will be applied to the start date and end date. The detailed algorithm will be provided in the study programming datasets specifications (PDS).

For patients who permanently discontinue study treatment, there should be a dosing interval with its start date equal to the date of the permanent discontinuation of the study drug, its end date equal to the study completion/discontinuation date, and its dispensing level equal to "No treatment".

For patients who complete the treatment and study, the dose level for the end of study visit will be the dose level for the end of treatment in the dose level by visit report.

The dose level is defined based on the dispensing level according to Table 2-5 and Table 2-6 if the dispensing level is a pediatric dose level or an adult dose level. If the dispensing level is no treatment, the dose level is defined as 0.

For a dosing interval with a pediatric dose level, the daily dose (mg/day) during the interval is defined as the product of the weight based daily dose (mg/day/kg) during the interval and the weight during the interval, where the weight based daily dose (mg/day/kg) during the interval is defined in Table 2-5 and the weight during the interval is defined as the weight from the last non-missing assessment (scheduled or unscheduled) prior to or equal to the start dose time of the study drug of the interval.

Table 2-5 Pediatric dose levels and corresponding doses of the study drug

Dispensing Level	Dose Level	Weight Based Dose		Weight Based Daily Dose	
		LCZ696	Enalapril	LCZ696	Enalapril
Age group 1 and 2					
Dose level 1	1	0.8 mg/kg bid.	0.05 mg/kg bid.	1.6 mg/day/kg	0.1 mg/day/kg
Dose level 2	2	1.6 mg/kg bid.	0.10 mg/kg bid.	3.2 mg/day/kg	0.2 mg/day/kg
Dose level 3	3	2.3 mg/kg bid.	0.15 mg/kg bid.	4.6 mg/day/kg	0.3 mg/day/kg
Dose level 4	4	3.1 mg/kg bid.	0.20 mg/kg bid.	6.2 mg/day/kg	0.4 mg/day/kg
Age group 3					
Dose level 1x	1	0.8 mg/kg bid.	0.05 mg/kg bid.	1.6 mg/day/kg	0.1 mg/day/kg
Dose level 2x	1.5	1.2 mg/kg bid.	0.075 mg/kg bid.	2.4 mg/day/kg	0.15 mg/day/kg
Dose level 3x	2	1.6 mg/kg bid.	0.1 mg/kg bid.	3.2 mg/day/kg	0.2 mg/day/kg
Dose level 4x	3	2.3 mg/kg bid.	0.15 mg/kg bid.	4.6 mg/day/kg	0.3 mg/day/kg
Dose level 5x	4	3.1 mg/kg bid	0.20 mg/kg bid.	6.2 mg/day/kg	0.4 mg/day/kg

Table 2-6 provides the weight based daily dose (mg/day/kg) for dosing intervals with adult dose levels.

Table 2-6 Adult (≥ 57 kg) dose levels and corresponding doses of the study drug

Dispensing Level	Dose Level	Dose		Weight Based Daily Dose	
		LCZ696	Enalapril	LCZ696	Enalapril
Adult dose level 1	1	50 mg bid.	2.5 mg bid.	1.6 mg/kg/day	0.1 mg/kg/day
Adult dose level 2	2	100 mg bid.	5.0 mg bid.	3.2 mg/kg/day	0.2 mg/kg/day
Adult dose level 3	3	150 mg bid.	7.5 mg bid.	4.6 mg/kg/day	0.3 mg/kg/day
Adult dose level 4	4	200 mg bid.	10.0 mg bid.	6.2 mg/kg/day	0.4 mg/kg/day

The weight based daily dose (mg/kg/day) is defined as 0 mg/kg/day for dosing intervals with no treatment.

Dispensing level, dose level and daily dose at each visit

For a given visit, the visit associated dosing interval is defined as the dosing interval with its start date prior to or equal to the date of visit AND its end date later than or equal to the date of visit; the dispensing level, dose level and weight based daily dose (mg/day/kg) at the visit are defined as the dispensing level, dose level and weight based daily dose (mg/day/kg) for the visit associated dosing interval. In the case that, the visit date is equal to the end date of an earlier dosing interval as well as the start date of a later dosing interval, the later dosing interval will be taken as the visit associated dosing interval.

The dispensing level and the dose level will be summarized by treatment group and visit, for overall and for each modified age group, using the number and percentage of patients on each level; the dose level will be summarized by treatment group and visit, for overall and for each age group and modified age group, using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum; the weight based daily dose (mg/day/kg) will be summarized by treatment group and visit, for overall and for each age group and modified age group, using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum.

The summary will include all scheduled visits in the double-blind epoch.

The SAF2 will be used for the above analyses.

Duration of treatment exposure

The duration (day) of treatment exposure during the double-blind epoch is defined as the date of the last dose of the study drug during the double-blind epoch minus the date of the first dose of the study drug during the double-blind epoch plus one day, regardless of the temporary treatment interruption.

The duration (month) of treatment exposure during the double-blind epoch will be converted from the duration (day) of treatment exposure during the double-blind epoch to months using a factor of 30.4375 day/month. The duration (year) of treatment exposure during the double-blind epoch will be converted from the duration (day) of treatment exposure during the double-blind epoch to years using a factor of 365.25 day/year. The duration (day) of treatment exposure

during the double-blind epoch will be categorized into the following categories, where 1 week = 7 days.

- < 2 weeks
- 2 weeks to < 8 weeks
- 8 weeks to < 6 months
- >= 6 months

The categorized duration of treatment exposure during the double-blind epoch will be summarized by age group and modified age group and treatment group using the numbers and percentages of patients within each category.

The duration (month) of treatment exposure during the double-blind epoch will be summarized by age group and modified age group and treatment group using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum.

Besides, the overall patient-years on-treatment will be provided by age group and modified age group and treatment group, where the overall patient-years on-treatment is defined as the sum of duration (year) of treatment exposure among the corresponding the age group and modified age group and/or the treatment group.

The SAF2 will be used for the above analyses.

Duration on each dose level

For a given dose level, the duration (day) on this dose level is defined as the number of days on this dose level during the period from the date of the first dose of the study drug to the date of the last dose of the study drug. In the case that, the end date of an earlier dosing interval is the same as the start date of a later dosing interval, the day will be counted as 0.5 day on both the earlier dose level and the later dose level. In the case that, the date of permanent discontinuation of the study drug is the same as the date of the last dose the study drug, the day will be counted as 0.5 day on the dose level of the last dose but not counted on dose level 0.

The duration (month) on each dose level during the double-blind epoch will be derived from the duration (day) on this dose level during the double-blind epoch using a factor of 30.4375 day/month, and will be summarized by age group and modified age group and treatment group using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum.

The duration (day) on each dose level during the double-blind epoch will be categorized into the following categories, where 1 week = 7 days and 1 month = 30.4375 days.

- < 2 weeks
- 2 weeks to < 8 weeks
- 8 weeks to < 6 months
- >= 6 months

The categorized duration on each dose level during the double-blind epoch will be summarized by age group and modified age group and treatment group using the numbers and percentages of patients within each category.

The SAF2 will be used for the above analyses.

Duration of study drug exposure

The duration (day) of study drug exposure during the double-blind epoch is defined as the sum of the duration (day) on dose level 1 to 4 during the double-blind epoch.

The duration (month) of study drug exposure during the double-blind epoch will be derived from the duration (day) of study drug exposure during the double-blind epoch using a factor of 30.4375 day/month, and will be summarized by age group and modified age group and treatment group using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum.

The duration (day) of study drug exposure during the double-blind epoch will be categorized into the following categories, where 1 week = 7 days and 1 month = 30.4375 days.

- < 2 weeks
- 2 weeks to < 8 weeks
- 8 weeks to < 6 months
- >= 6 months

The categorized duration of study drug exposure during the double-blind epoch will be summarized by age group and modified age group and treatment group using the numbers and percentages of patients within each category.

The SAF2 will be used for the above analyses.

Percentage of study drug exposure

For each patient, the percentage of study drug exposure during the double-blind epoch is defined as the percentage of the duration (day) of study drug exposure during the double-blind epoch out of the duration (day) of treatment exposure during the double-blind epoch.

The percentage of study drug exposure during the double-blind epoch will be summarized by age group and modified age group and treatment group using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum.

The SAF2 will be used for the above analyses.

Time to first dose of each dose level

For a given dose level, the time (day) to first dose on this dose level during the double-blind epoch is defined as the date of the first dose on this dose level during the double-blind epoch minus the date of the first dose of the study drug plus one day. If there is no dosing interval on this dose level during the double-blind epoch, the time (day) to first dose on this dose level during the double-blind epoch will be censored on the date of last visit during the double-blind epoch for this patient.

For each dose level, the time (day) to first dose on the dose level during the double-blind epoch will be summarized using the Kaplan-Meier curve for each treatment group within each age group and modified age group.

The SAF2 will be used for the above analyses.

Time to permanent discontinuation of the study drug

The time (day) to permanent discontinuation of the study drug during the double-blind epoch is defined as the date of the permanent discontinuation of the study drug minus the date of the first dose of the study drug plus one day. If the study drug is not permanently discontinued during the double-blind epoch, the time (day) to permanent discontinuation of the study drug during the double-blind epoch will be censored on the date of last visit during the double-blind epoch for this patient.

The time (day) to permanent discontinuation of the study drug during the double-blind epoch (excluding death) will be summarized for each treatment group within each age group and modified age group using the Kaplan-Meier curves.

The SAF2 will be used for the above analyses.

Mean weight based daily dose and mean dose level

For a given dose level, the duration (day) on this dose level is defined as the number of days on this dose level during the period from the date of the first dose of the double-blind study drug to the date of the last dose of the double-blind study drug. In the case that, the end date of an earlier dosing interval is the same as the start date of a later dosing interval, the day will be counted as 0.5 day on both the earlier dose level and the later dose level. In the case that, the date of permanent discontinuation of the study drug is the same as the date of the last dose the study drug, the day will be counted as 0.5 day on the dose level of the last dose but not counted on dose level 0.

For each patient, the mean weight based daily dose (mg/kg/day) during the double-blind epoch is defined as the weighted mean of the weight based daily dose (mg/kg/day) among dose level 0 to 4 using the duration (day) on each dose level during the double-blind epoch as the weight, i.e.

$$\frac{\sum_{k=0}^{4} \text{Weight Based Daily Dose (mg/kg/day) for Dose Level } k \times \text{Duration (day) on Dose Level } k}{\sum_{k=0}^{4} \text{Duration (day) on Dose Level } k}$$

the mean dose level during the double-blind epoch is defined as the weighted mean of the dose level among dose level 0 to 4 using the duration (day) on each dose level during the double-blind epoch as the weight, i.e.,

$$\frac{\sum_{k=0}^{4} k \times \text{Duration (day) on Dose Level } k}{\sum_{k=0}^{4} \text{Duration (day) on Dose Level } k}.$$

The mean weight based daily dose (mg/day/kg) during the double-blind epoch and the mean dose level during the double-blind epoch will be summarized by treatment group, for each age group and modified age group, using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum.

The SAF2 will be used for the above analyses.

Duration of study exposure

The duration (day) of study exposure during the double-bind epoch is defined as the date of last visit (Visit 416/499) minus the date of randomization plus one day.

The duration (month) of study exposure during the double-blind epoch will be derived from the duration (day) of study exposure during the double-blind epoch using a factor of 30.4375 day/month, and will be summarized by age group and modified age group and treatment group using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum.

The duration (day) of study exposure during the double-blind epoch will be categorized into the following categories, where 1 week = 7 days and 1 month = 30.4375 days.

- < 2 weeks
- 2 weeks to < 8 weeks
- 8 weeks to < 6 months
- >= 6 months

The categorized duration of study exposure during the double-blind epoch will be summarized by age group and modified age group and treatment group using the numbers and percentages of patients within each category.

The SAF2 will be used for the above analyses.

Last recorded dose level

The last recorded dose level during the double-blind epoch is defined as the dose level for the last dosing interval during the double-blind epoch.

The last recorded dose level during the double-blind epoch will be summarized by age group and modified age group and treatment group using the numbers and percentages of patients on each dose level.

The same summary will be performed for patients who permanently discontinued from the study treatment and for patients who completed the study treatment.

The SAF2 will be used for the above analyses.

Dose down titration

A down-titration dosing interval is defined as a dosing interval whose dose level is changed to a lower dose level from its previous dosing interval.

The dose down titration during the double-blind epoch will be summarized by age group and modified age group and treatment group using the numbers and percentages of patients with at least one down-titration dosing interval during the double-blind epoch.

The SAF2 will be used for the above analyses.

Dose interruption

A dose-interruption dosing interval is defined as a dosing interval with “NO TREATMENT”.

The numbers and percentages of patients with at least one dose-interruption dosing interval during the double-blind epoch will be provided by age group and modified age group and treatment group.

For each patient and each given dose-interruption dosing interval, the duration (day) of dose-interruption is defined as the number of days temporarily off-study drug in this dosing interval. In the case that the start date of a latter dosing interval is the same as the end date of this dosing interval, the end date will be counted as 0.5 day. In the case that, the end date of an earlier dosing interval is the same as the start date of this dosing interval, the start date will be counted as 0.5 day.

The numbers and percentages of patients having at least one dose-interruption dosing interval with duration (day) of dose-interruption larger than 14 days will be provided by age group and modified age group and treatment group.

The SAF2 will be used for the above analyses.

2.4.2 Prior, concomitant and post therapies

Heart failure and cardiovascular medications (prior and concomitant) will be recorded on the CRF pages: "Prior Cardiovascular Medications-ARB", and "Prior Cardiovascular Medications-ACEI". Other general prior and concomitant medications will be recorded on the page: "Prior and Concomitant Medications". Procedures and significant non-drug therapies (prior and concomitant) will be recorded on the CRF page "Surgical and Medical Procedures".

The missing or partially missing start/end date for prior/concomitant therapies will be imputed using the Novartis ADaM Governance Board (AGB) global standard approach.

Prior and concomitant medications/non-drug therapies will be identified based on recorded or imputed start and end dates. Details will be provided in the study programming datasets specifications (PDS).

'Prior medications' are defined as drugs taken prior to first dose of double-blind study medication. Any medication which has been started during the double-blind treatment epoch including medications started prior to randomization but continued in the double-blind treatment epoch are identified as **'Concomitant medications'**.

Prior and concomitant medications will be summarized separately by modified age group, treatment group, anatomical therapeutic classification (ATC) and PT. Prior and concomitant non-drug therapies will be summarized separately by modified age group, treatment group, SOC and PT.

In addition, for each of the following classes, prior medications and concomitant medications during the dose associated period will be summarized by age group and dose cohort for Part 1, and also summarized by modified age group and treatment group for Part 2.

Table 2-7 Cardiovascular and heart failure medications

Medication	ATC code
Angiotensin converting enzyme inhibitor (ACEI)	C09A, C09B
Angiotensin receptor blocker (ARB)	C09C, C09D

Beta blockers	C07
Diuretics	C03
Aldosterone antagonists (including spironolactone)	C03DA
Calcium antagonists	C08
Nitrates	C01DA
Other vasodilators	C01DX
Cardiac glycocytes (Digoxin/digitalis glycoside)	C01A
Antiarrhythmic agents	C01B
Aspirin	B01AC with preferred term of (ACETYLSALICYLIC ACID and/or ACETYLSALICYLATE LYSINE)
Other antiplatelet agents	B01AC except preferred term of (ACETYLSALICYLIC ACID and/or ACETYLSALICYLATE LYSINE)
Oral anticoagulants	B01AA, B01AE, B01AF, B01AX
Sodium-dependent glucose transporters 2 (SGLT2i)	A10BK

The use of prohibited medication during the study, if any, will also be summarized.

The SAF1 will be used for the Part 1 analyses and the SAF2 will be used for the Part 2 analyses.

2.5 Analysis of the primary objective for Part 1

2.5.1 Analysis of the PK parameters

The PK parameters will include for all LCZ696 analytes (Sacubitril, LBQ657 [also known as sacubitrilat] and Valsartan)

- C_{max}
- T_{max}
- AUC_{last}
- AUC_{inf}
- Clearance (Cl/F)
- $T_{1/2}$

In Part 1, PK samples *(as blood volume will allow)* will be taken at Visit 101 and 201 PK/PD at the following intervals for the various age groups:

1. Patients $\geq$ 6 years of age: Blood samples are collected at pre-dose (pre-dose sample for all patients) 0.5, 1, 2, 4, 8, 10, and optional 24 hours post dosing
2. Patients < 6 years of age: Blood samples are collected at pre-dose (pre-dose sample only from those patients who are on valsartan) 1, 2, 4, 10, and at an optional 24 hours post dosing

The PK parameters will be determined using non-compartmental method(s) with Phoenix v6.4, including clearance (Cl/F), $T_{1/2}$, C_{max}, T_{max}, AUC_{last}, AUC_{inf} based on Part 1 scheduled or unscheduled PK samples.

For each of the LCZ696 analytes (Sacubitril, LBQ657 and Valsartan), if data available, the PK parameters will be summarized by PK parameter, dose cohort within each age group using number of observations, mean, standard deviation, coefficient of variation (CV), geometric mean, geometric standard deviation, geometric CV, median, minimum, and maximum, where the geometric mean is the exponential transformation of the mean of the log transformed data, the geometric standard deviation is the exponential transformation of the standard deviation of the log transformed data, the geometric CV is the square root of (the exponential transformed variance of the log transformed data minus one).

These summaries will only include scheduled assessments within the dose associated period.

For each of the LCZ696 analytes (Sacubitril, LBQ657 and Valsartan), if data available, the individual LCZ696 concentration-time profiles within the dose day to the dose day plus 5 days will be plotted at patient level, by dose cohort within each age group.

The PK1 will be used for the above analyses. Listing will be provided.

2.5.2 Analysis of the PD parameters

The following PD parameters (biomarkers) will be included.

- plasma cGMP
- plasma NTproBNP
- plasma BNP
- urine cGMP

The blood samples will be collected at pre-dose, 4, 8 and optional 24 hours post dosing. The urine samples will be collected at pre-dose and once between 4 to 8 hours post dosing.

For each PD parameter, the test values and the ratios to baseline in the PD parameter will be summarized by time point, dose cohort within each age group using number of observations, geometric mean, geometric standard deviation, median, minimum, Q1, Q3, and maximum.

These summaries will only include scheduled assessments within the dose associated period.

The following analyses will not be included in the CSR but in separate report.

A population PK model will be developed to describe incoming data from pediatric patients based on an established model developed for the adult HF population.

Pharmacodynamic evaluations will include analysis of change in plasma cGMP, plasma BNP, plasma NTproBNP and urine cGMP. Dose and exposure-response relationship will be evaluated based on changes in plasma cGMP levels (and also plasma NTproBNP, plasma BNP, urine cGMP).

PKPD modeling of Part 1 data may be carried out, depending on the robustness of both the PK and biomarkers to be used to inform the dose justification for Part 2. The PD1 will be used for the above analyses. Listing will be provided.

2.5.3 Handling of missing values/censoring/discontinuations

All concentrations below the lowest limit of quantitation (LLOQ) or missing data will be labeled as such in the concentration data listings. Concentrations below the LLOQ will be treated as zero in summary statistics and for the calculation of PK parameters.

An UNS PK/PD assessment will be mapped into Visit 101/201 if all the following criteria are satisfied.

- The UNS PK/PD assessment is post-baseline.
- The UNS PK/PD assessment is valid (i.e., not flagged for exclusion).
- There is no valid (i.e., not flagged for exclusion) PK/PD assessment on Visit 101/201.
- The dose level prescribed on Visit 101/201 and on the UNS PK/PD visit are the same.

2.6 Analysis of the primary objective for Part 2

2.6.1 Primary endpoint

Patient strata will be defined based on the combination of the modified age group (6 years to < 18 years, 2 years to < 6 years, 1 month to < 2 years) at randomization and NYHA/ROSS class group (Class I/II, Class III/IV) at randomization. There will be six patient strata listed as following.

1. Age Group 1 (6 years to < 18 years) and NYHA/ROSS Class I/II
2. Age Group 1 (6 years to < 18 years) and NYHA/ROSS Class III/IV
3. Age Group 2a (2 years to < 6 years) and NYHA/ROSS Class I/II
4. Age Group 2a (2 years to < 6 years) and NYHA/ROSS Class III/IV
5. Age Group 3a (1 month to < 2 years) and NYHA/ROSS Class I/II
6. Age Group 3a (1 month to < 2 years) and NYHA/ROSS Class III/IV

The primary endpoint is the global rank endpoint through 52 weeks of treatment (analysis cutoff date will be set to 58 weeks, i.e. 406 days), which will be constructed through two steps within each of the six strata.

- Step 1: Patients are classified into five ordinal categories based on the logic described below in Table 2-8.
- Step 2: Within each category, patients are ranked from worst to best based first on the subcategory if applicable, and then the ranking algorithm as explained in Table 2-8.

Table 2-8 Primary endpoint algorithm using ranked analysis

Category	Sub-category	Description	Ranking algorithm
1	Death; UNOS status 1A listing for heart transplant or equivalent; VAD/ECMO/mechanical ventilation/intra-aortic ballon pump requirement for life support at end of study.		
	A	• Death; • UNOS status 1A listing for heart transplant or equivalent;	Rank within this category by time to first event. All Category 1 events are considered equal.

Category	Sub-category	Description	Ranking algorithm
		• VAD/ECMO/mechanical ventilation/intra-aortic balloon pump requirement for life support at end of study.	
2	Worsening HF (WHF); defined by signs and symptoms of WHF that requires an intensification of HF therapy.		
	B	Worsening heart failure hospitalization with intensive care unit stay.	Within Category 2, the patients will be ranked first by event subcategory, and then by number of events within each subcategory. Further ranking by time to first event in the worst subcategory.
	C	Worsening heart failure hospitalization without intensive care unit stay.	
	D	Worsening heart failure without hospitalization.	
3	Worsened; worse NYHA/ROSS or worse PGIS; and further ranking by PedsQL physical functioning domain.		
	E	NYHA/ROSS or PGIS worsened based on last available assessment compared to baseline.	Rank by combination of NYHA/ROSS and PGIS degree of change. Within a group of the same degree of NYHA/ROSS and PGIS change, further rank by PedsQL (physical functioning domain) change from baseline.
4	Unchanged; unchanged NYHA/ROSS and unchanged PGIS; and further ranking by PedsQL physical functioning domain.		
	F	NYHA/ROSS and PGIS unchanged based on last available assessment compared to baseline.	Worst baseline combination of NYHA/ROSS functional class and PGIS without change is ranked worse than a better baseline NYHA/ROSS functional class and PGIS. Within a group of the same baseline NYHA/ROSS and PGIS, further rank by PedsQL (physical functioning domain) change from baseline.
5	Improved; improved NYHA/ROSS or improved PGIS (neither can be worse); and further ranking by PedsQL physical functioning domain).		
	G	NYHA/ROSS or PGIS improved (neither worsened) based on last available assessment compared to baseline	Rank by combination of NYHA/ROSS and PGIS degree of change. Within a group of the same degree of NYHA/ROSS and PGIS change, further rank by PedsQL (physical functioning domain) change from baseline.

- PedsQL physical functioning will only be used for the global rank endpoint for Age Group 1.

Patient classification for global rank endpoint

Category 1 events include the following three events, and will be reported by the investigator and adjudicated by the adjudication committee.

- Death;
- UNOS status 1A listing for heart transplant or equivalent;
- VAD/ECMO/mechanical ventilation/intra-aortic balloon pump requirement for life support.

The investigator reported events will be recorded on the CRF pages: “Death (Endpoint)”, “Heart Transplant Listing Endpoint” and “Circulatory or Respiratory Mechanical Assistance Endpoint” together with the reported date of event; the adjudicated events will be recorded on the CRF pages: “Death (Adjudication)”, “Heart Transplant Listing Adjudication” and “Circulatory or Respiratory Mechanical Assistance Adjudication” together with the adjudicated date of event.

A patient will be classified into Category 1 if at least one positively adjudicated Category 1 event occurred during the double-blind epoch.

Within Category 1, patients will be ranked by the time to first positively adjudicated Category 1 event, which is defined as the date of first positively adjudicated Category 1 event minus the date of randomization plus one day.

Category 2 event is the worsening heart failure (WHF), which is further categorized into three sub-categories as following.

- Category 2-B: Worsening heart failure hospitalization with intensive care unit stay.
- Category 2-C: Worsening heart failure hospitalization without intensive care unit stay.
- Category 2-D: Worsening heart failure without hospitalization.

The WHF will be reported by the investigator and adjudicated by the adjudication committee. The investigator reported events will be recorded on the CRF pages: “Clinical Worsening of Heart Failure Endpoint”; the adjudicated events will be recorded on the CRF page: “Clinical Worsening of Heart Failure Endpoint (Adjudication)” together with the date of event and event sub-category.

A patient will be classified into Category 2 if at least one positively adjudicated Category 2 event occurred during the double-blind epoch AND the patient is not in Category 1.

Category 2 patients will be classified into three sub-categories (2-B, 2-C and 2-D), using the following criteria.

- A patient will be classified into Category 2-B if the patient has at least one positively adjudicated Category 2-B event during the double-blind epoch.
- A patient will be classified into Category 2-C if the patient has at least one positively adjudicated Category 2-C event during the double-blind epoch, and is not in Category 2-B.
- A patient will be classified into Category 2-D if the patient has at least one positively adjudicated Category 2-D event during the double-blind epoch, and is not in Category 2-B or 2-C.

Within Category 2, patients will be first ranked by the patient-sub-category (2-B < 2-C < 2-D). Within each patient-sub-category, patients will be ranked first by the number of positively adjudicated events in the corresponding sub-category (taking into consideration worsening hierarchy of 2-B < 2-C < 2-D), and then by the time to first positively adjudicated event in the

sub-category, which is defined as the date of the first positively adjudicated event in the sub-category minus the date of randomization plus one day.

Patients not in Category 1 and 2 will be classified into Category 3 to 5 based on NYHA/ROSS class and PGIS at Week 52.

The NYHA/ROSS class is defined based on the age at randomization.

- For patients with age at randomization from 1 month to < 6 years, the NYHA/ROSS class will be the ROSS class recorded on the CRF page: "Modified ROSS HF Classification".
- For patients with age at randomization from 6 years to < 18 years, the NYHA/ROSS class will be the NYHA class recorded on the CRF page: "Signs and symptoms of Heart Failure including NYHA Classification".

The PGIS score is also based on the age at randomization.

- For patients with age at randomization < 5 years, the PGIS will be a five-point scale (None / Mild / Moderate / Severe / Very severe) recorded on the CRF page: "Patient Global Impression of Severity – Heart Failure Symptoms – Parent/Caregiver (For patient < 5 years of age)".
- For patients with age at randomization from 5 years to < 7 years, the PGIS will be a three-point scale (Good / Neither good nor bad / Bad) recorded on the CRF page: "Patient Global Impression of Severity – Heart Failure Symptoms (Patients 5 to <7 years old)".
- For patients with age at randomization from 7 years to < 18 years, the PGIS will be a five-point scale (None / Mild / Moderate / Severe / Very severe) recorded on the CRF page: "Patient Global Impression of Severity – Heart Failure Symptoms (Patients 7 to <18 years old)".

Table 2-9 and Table 2-10 define the degree of change from baseline in NYHA/ROSS class and in PGIS at each post-baseline visit.

Table 2-9 Degree of change from baseline in NYHA/ROSS class

Degree of Change from Baseline		Post-Baseline Visit			
		Class I	Class II	Class III	Class IV
Baseline	Class I	0	-1	-2	-3
	Class II	+1	0	-1	-2
	Class III	+2	+1	0	-1
	Class IV	+3	+2	+1	0

Table 2-10 Degree of change from baseline in PGIS

Degree of Change from Baseline		Post-Baseline Visit				
		None (Good)	Mild	Moderate (NGNB)	Severe	Very severe (Bad)
Baseline	None (Good)	0	-1	-2	-3	-4
	Mild	+1	0	-1	-2	-3
	Moderate (NGNB)	+2	+1	0	-1	-2
	Severe	+3	+2	+1	0	-1
	Very severe (Bad)	+4	+3	+2	+1	0

Degree of Change from Baseline	Post-Baseline Visit				
	None (Good)	Mild	Moderate (NGNB)	Severe	Very severe (Bad)

- NGNB = neither good nor bad

Based on the degree of change defined above in Table 2-9 and Table 2-10, Table 2-11 defines an initial ranking score. For patients in Category 3 to 5, patients with lower initial ranking score are considered worse than those with higher score.

A patient not in Category 1 and 2 will be classified into Category 3, if the initial ranking score at Week 52 is lower than 30; into Category 4, if the initial ranking score at Week 52 is equal to 30, into Category 5, if the initial ranking score at Week 52 is higher than 30.

Table 2-11 Initial ranking score based on degree of change from baseline in NYHA/ROSS class and PGIS at Week 52

Initial Ranking Score		Degree of Change from baseline in PGIS at Week 52								
		-4	-3	-2	-1	0	+1	+2	+3	+4
Degree of Change from Baseline in NHYA/ROSS Class at Week 52	-3	1	2	4	7	10	14	18	22	26
	-2	3	4	6	9	13	17	21	25	28
	-1	5	7	9	12	16	20	24	27	29
	0	8	10	13	16	30	31	32	34	36
	+1	11	14	17	20	31	33	35	37	39
	+2	15	18	21	24	32	35	38	40	41
	+3	19	22	25	27	34	37	40	42	43

A subset of the physical functioning domain based on patient reported PedsQL will be used to rank patients with age at randomization from 6 years to < 18 years.

The PedsQL is an age based questionnaire. For patients with age at randomization from 6 years to < 18 years, the patient reported PedsQL will be recorded on the following CRF pages.

- "PEDsQL Pediatric Quality of Life Inventory Acute Version Teen Report (Ages 13-18)"
- "PEDsQL Pediatric Quality of Life Inventory Acute Version Child Report (Ages 8-12)"
- "PEDsQL Pediatric Quality of Life Inventory Acute Version Young Child Report (Ages 5-7)"

Table 2-12 and Table 2-13 present the items in PedsQL functioning domain used for ranking and the mapped score for the observed values at Week 52.

Table 2-12 PedsQL Ranking Items (Age at Randomization: 8 years to < 18 years)

Item ID	Item	Never	Almost Never	Some-times	Often	Almost Always
1	It is hard for me to walk more than one block	100	75	50	25	0
2	It is hard for me to run	100	75	50	25	0
3	It is hard for me to do sports activity or exercise	100	75	50	25	0
4	It is hard for me to lift something heavy	100	75	50	25	0

Item ID	Item	Never	Almost Never	Some-times	Often	Almost Always
6	It is hard for me to do chores around the house	100	75	50	25	0

Table 2-13 PedsQL Ranking Items (Age at Randomization: 6 years to 7 years)

Item ID	Item	Not at All	Sometimes	A Lot
1	It is hard for you to walk	100	50	0
2	It is hard for you to run	100	50	0
3	It is hard for you to play sports activity or exercise	100	50	0
4	It is hard for you to pick up big things	100	50	0
6	It is hard for me to do chores (like pick up your toys)	100	50	0

For each post-baseline visit, if there are three or more non-missing items, the PedsQL ranking score will be defined as the mean of the non-missing mapped values; if there are three or more missing items, the PedsQL ranking score will be considered as missing.

Within Category 3 to 5, patients with age at randomization from 6 years to < 18 years will be ranked by the initial ranking score (Table 2-11) at Week 52 and the change from baseline in the PedsQL ranking score at Week 52, while patients with age at randomization from 1 month to < 6 years will be ranked only by the initial ranking score at Week 52 but not by the PedsQL ranking score.

The response category is defined as the category which the patient is classified into.

Due to unresolvable technical reasons in study drug supply, the study treatment of all patients remaining in the study will be discontinued by the end of October 2021, two months prior to the scheduled study end. This early study drug discontinuation is not expected to compromise the scientific integrity of the study. The early study drug discontinuation affects a maximum of 31 patients, who are still on-treatment with current follow-up between 297 days to 395 days. The total on-treatment follow-up time lost is expected to be <1% (2.26 patient-years of the anticipated 375 patient-years; assuming immediate study drug discontinuation on 31-Oct-2021).

The reason for treatment discontinuation will be documented as ‘Technical problems’. In case of patient withdrawal from further study participation after study drug discontinuation due to the USM, the reason for study phase discontinuation will also be documented as ‘technical problems’.

The unforeseen intercurrent events of USM based study treatment discontinuations are not related to disease progression or to the assigned study treatment and not reflecting a real world setting, and the exposure to the study treatment and follow-up is reasonably long. Hence, an on-treatment approach is considered to be more appropriate for the handling of USM-impacted patients in the primary analysis, utilizing the relevant components of the global rank endpoint including Category 1 and Category 2 status at the unscheduled assessment at the time of treatment discontinuation.

2.6.2 Statistical hypothesis, model, and method of analysis

Within each stratum, the Mann-Whitney (MW) probability is defined as the probability of the patient from the LCZ696 group having better outcome than the patient from the Enalapril group plus half of the probability of the patient from the LCZ696 group having equal outcome to the one from the Enalapril group, when the two patients are independently sampled from the LCZ696 group and the Enalapril group. Correspondingly, the MW odds is defined as (one minus the MW probability) divided by the MW probability.

The null hypothesis is that, the MW odds in all strata are equal to one, while the alternative hypothesis is that, the MW odds is not equal to one in at least one stratum.

The hypothesis will be tested by the stratified Wilcoxon rank-sum test with modified age group (6 years to < 18 years, 2 years to < 6 years, 1 month to < 2 years) and baseline NYHA/ROSS class group (Class I/II, III/IV) as stratification factors. The testing statistic will be based on the weighted sum of the within-stratum rank-sum statistics, in which, the weight for stratum will be the reciprocal of (one plus the total sample size of the stratum), the mean and the variance will be estimated under the null hypothesis as outlined in Appendix 5.2.1. The test will be performed at an overall significance level of two-sided 0.05.

Note that the expected value of the stratified Wilcoxon rank-sum statistics is proportional to the overall MW probability (defined as weighted mean of the MW probabilities, with weights depending only on the sizes of each stratified stratum as specified in Appendix 5.2.1), with modified age group (6 years to < 18 years, 2 years to < 6 years, 1 month to < 2 years) and baseline NYHA/ROSS class group (Class I/II, III/IV) as stratification factors. The treatment effect size will be defined by the overall MW probability, which can be estimated using the stratified Wilcoxon rank-sum statistic, see Appendix 5.2.1. The corresponding confidence interval will also be provided for overall MW probability using the approach in Appendix 5.2.1, or see Kawaguchi *et.al.* 2011.

Correspondingly, the overall MW odds is defined as (one minus the overall MW probability) divided by the overall MW probability. The estimate and the corresponding 95% confidence interval for the overall MW odds will also be provided (Appendix 5.2.1).

The superiority of LCZ696 to Enalapril in the global rank endpoint will be claimed if the test is significant and the point estimate of the overall MW probability is greater than 0.5.

Besides, the response category will be summarized by modified age group, baseline NYHA/ROSS class group, and treatment group, using the numbers and percentages of patients in each category.

The FAS will be used for the above analyses.

2.6.3 Handling of missing values/censoring/discontinuations

For Category 1 and Category 2, patients will be classified into the worst possible category.

- A patient will be classified into Category 1 with event date imputed by the last known alive date (recorded on CRF page: “Survival Information”), or the date of last visit, whichever occurred later but earlier than the analysis cutoff date, if the patient discontinues from the study during the double-blind epoch without any positively

adjudicated Category 1 event and with no available clinical endpoint information to refute classification into Category 1.

- A patient will be classified into Category 2 with event date imputed by the last known alive date (recorded on CRF page: "Survival Information"), or the date of last visit, whichever occurred later but earlier than the analysis cutoff date, if the patient discontinue from the study during the double-blind epoch without any positively adjudicated Category 1 and there is available and retrievable information for the patient that refutes classification into Category 1.
- Any positively adjudicated Category 2 event with its adjudicated sub-category missing will be classified into Category 2-B (Category 2-B: Worsening heart failure hospitalization with intensive care unit stay).
- If the event date of a positively adjudicated Category 1 event is missing, the event date will be imputed by the last known alive date (recorded on CRF page: "Survival Information"), or the date of last visit, whichever occurred later but earlier than the analysis cutoff date.
- If the event date of a positively adjudicated Category 2 event is missing, the event date will be imputed by the date of last visit but earlier than the analysis cutoff date.

For patients in Category 3 to 5, if the 52 weeks measurements of NYHA/Ross, PGIS or PedsQL score are missing then the last-observation-carried-forward (LOCF) technique will be used to impute the missing post-randomization values (for patients who do not experience any events in Category 1 or 2). While ranking, if the baseline is missing, then the baseline will be imputed using the median assessment within each stratum. If all the post-randomization assessments is missing, then the post randomization assessment will be imputed to the worst case at week 52.

2.6.4 Supportive and sensitivity analyses

PPS analysis

As a sensitivity analysis, the primary analysis approach will also be performed on the PPS.

Response category analysis

The primary analysis approach (Section 2.6.2) will be performed for the response category (Category 1 to Category 5) at Week 52 based on the ranking algorithm (Section 2.6.1) for the primary endpoint.

MI analysis

A multiple imputation approach will be applied, whereby all missing data for degree of change in NYHA/ROSS class, degree of change in PGIS, and change from baseline in PedsQL ranking score, will be considered as missing at random within each arm, within Category 3 to Category 5 and will be imputed as described below.

- Missing data of degree of change from baseline in NYHA/ROSS class and degree of change from baseline in PGIS will be imputed based on cumulative proportional odds models using fully conditional specifications. For each post-baseline scheduled visit (Week 4, Week 12, Week 24, Week 36, Week 52), the imputation model will include region, modified age group (6 years to < 18 years, 2 years to < 6 years, 1 month to < 2 years) at randomization, NYHA/ROSS class group (Class I/II, Class III/IV) at

randomization, baseline value, and value at the immediately previous scheduled visit (Not applicable for Week 4, Week 4 for Week12, Week 12 for Week 24, Week 24 for Week 36, Week 36 for Week 52) as fixed-effect factors.

- Missing data of change from baseline in PedsQL ranking score will be imputed based on linear regression models (for continuous variables) via fully conditional specifications. For each post-baseline scheduled visit (Week 12, Week 24, Week 36 and Week 52), the imputation model will include region, modified age group (6 years to < 18 years, 2 years to < 6 years, 1 month to < 2 years) at randomization, NYHA/ROSS class group (Class I/II, Class III/IV) at randomization as fixed-effect factor; baseline value, and value at the immediately previous scheduled visit (Not applicable for Week 12, Week 12 for Week 24, Week 24 for Week 36, Week 36 for Week 52) as covariates.
- The imputation will be performed within each arm for degree of change from baseline in NYHA/ROSS, degree of change from baseline in PGIS, and change from baseline in PedsQL ranking score within Category 3 to Category 5.

The stratified estimation for the MW probability and the MW odds will be estimated from each imputed datasets, and then be combined using Rubin's rules to get the final stratified estimations. (Appendix 5.2.1)

Tipping point analysis

A tipping point analysis will be performed for the primary analysis on the FAS. In the primary analysis, patients who discontinued from the study without Category 1 event will be classified (with probability 1) to the worst possible category. Alternatively, responses after discontinuation can be imputed assuming censoring at random, by combining the observed data prior to discontinuation with data of a sampled patient among those patients who are still in the study at that time. If combined data has Category 1 or 2 events, then, the discontinued patient will be classified based on the events information. If combined data has no Category 1 or 2 event, then the discontinued patient will be classified using the ranking score of the sampled patient.

The imputation/sampling will be performed within each treatment arm. The probability to utilize worst category imputation for a discontinued patient (p) will be varied from 1 to 0 in steps of 0.1 and accordingly the probability to utilize MAR imputation (1-p) will be varied from 0 to 1 in steps of 0.1. For each p, multiple imputed datasets will be produced and for each dataset, the stratified MW probability and MW odds will be estimated. The results will then be combined across datasets using Rubin's rules to get the final stratified estimations.

After such tipping point is determined, clinical judgement can be applied as to the plausibility of the assumptions underlying this tipping point. This methodology will provide a good picture of what it would take to overturn study conclusions on the basis of varying assumptions about patients discontinued from the study.

Pooled strata analysis

When there is no patient or only 1 patient in any treatment group within any of the stratum, a supportive analysis will be added using the four pooled strata, listed below:

1. Age Group 1 (6 years to < 18 years) and NYHA/ROSS Class I/II
2. Age Group 1 (6 years to < 18 years) and NYHA/ROSS Class III/IV
3. Age Group 2a and 3a (1 month to < 6 years) and NYHA/ROSS Class I/II

4. Age Group 2a and 3a (1 month to < 6 years) and NYHA/ROSS Class III/IV

USM-impacted off treatment analysis

For the USM-impacted patients, the full data, including the off-treatment part, from the impacted patients will be included as a sensitivity analyses.

Without cutoff analysis

The primary analysis approach will also be performed on the nominal week 52 assessment, i.e. without considering analysis cutoff date for NYHA/ROSS, PGIS, and PedsQL.

Supportive analysis

If possible, primary analysis will be done excluding patients who received expired drug.

The FAS will be used for above analysis.

2.6.5 Subgroup analyses

The subgroup variables in Table 2-6 will be used for subgroup analyses.

The primary analysis approach (Section 2.6.2) will be performed within each subgroup.

The FAS will be used for the subgroup analyses.

2.7 Analysis of the secondary objectives for Part 2

For the USM-impacted patients, the relevant assessments at the unscheduled visit at the time of treatment discontinuation will be utilized in the secondary analyses (on-treatment approach).

For the USM-impacted patients, the full data, including the off-treatment part, from the impacted patients will be included as a sensitivity analyses.

2.7.1 Time to first adjudicated Category 1 or 2 event

Adjudicated Category 1 or 2 events include all positively adjudicated Category 1 events and positively adjudicated Category 2 events (Section 2.6.1).

The time to first adjudicated Category 1 or 2 event is defined as the date of first occurrence of adjudicated Category 1 or 2 events minus the date of randomization plus one day.

The censoring will occur at minimum [date of death, date of withdrawal of consent, maximum (date of last visit, last known alive date)].

The time to first adjudicated Category 1 or 2 event will be analyzed using a Cox proportional hazard model, stratified by modified age group and NYHA/ROSS class group (Class I/II, Class III/IV) at randomization with treatment (LCZ696, Enalapril) included as a fixed-effect factor.

Based on the above model, the estimate and the 95% confidence interval will be provided for the adjusted hazard ratio (LCZ696 over Enalapril), the corresponding two-sided p-value will also be provided.

Besides, the time to first adjudicated Category 1 or 2 event will be summarized by modified age group and treatment group, using Kaplan-Meier curves.

For the first adjudicated Category 1 or 2 event, the annualized event rate will be presented by modified age group and treatment group.

The investigator reported Category 1 or 2 event will also be analyzed by the above method.

The FAS will be used for the above analyses.

2.7.2 NYHA/ROSS class change

The NYHA/ROSS class is an age based four-point scale (Section 2.6.1).

Table 2-14 defines the NYHA/ROSS class change at each post-baseline visit.

Table 2-14 NYHA/ROSS class change

NYHA/ROSS Class Change		Post-Baseline Visit			
		Class I	Class II	Class III	Class IV
Baseline	Class I	unchanged	worsened	worsened	worsened
	Class II	improved	unchanged	worsened	worsened
	Class III	improved	improved	unchanged	worsened
	Class IV	improved	improved	improved	unchanged

The NYHA/ROSS class change will be analyzed using a by-visit proportional cumulative odds model. For each visit (Week 4, Week 12, Week 24, Week 36 and Week 52), the proportional cumulative odds model uses NYHA/ROSS class change (order specified as "improved" < "unchanged" < "worsened") at the visit as response, and include treatment (LCZ696, Enalapril), baseline NYHA/ROSS class (Class I to Class IV), modified age group(6 years to < 18 years, 2 years to < 6 years, 1 month to < 2 years) as fixed effect. The analysis cutoff date will be considered.

Based on the above model, the estimate and the 95% confidence interval will be provided for the adjusted odds ratios (LCZ696 over Enalapril) at each post-baseline scheduled visit (Week 4, Week 12, Week 24, Week 36 and Week 52), the corresponding two-sided p-values will also be provided.

Besides, the NYHA/ROSS class and the NYHA/ROSS class change will both be summarized by modified age group, treatment group, and visit, using the numbers and percentages of patients in each category.

A sensitivity analysis will be also included without considering analysis cutoff date.

The FAS will be used for the above analyses.

2.7.3 PGIS change

The PGIS score is an age based five-point or three-point scale (Section 2.6.1).

Table 2-15 defines the PGIS change at each post-baseline visit.

Table 2-15 PGIS change

PGIS Change		Post-Baseline Visit				
		None (Good)	Mild	Moderate (NGNB)	Severe	Very severe (Bad)
Baseline	None (Good)	unchanged	worsened	worsened	worsened	worsened
	Mild	improved	unchanged	worsened	worsened	worsened
	Moderate (NGNB)	improved	improved	unchanged	worsened	worsened
	Severe	improved	improved	improved	unchanged	worsened
	Very severe (Bad)	improved	improved	improved	improved	unchanged

- NGNB = neither good nor bad

The PGIS change will be analyzed using a by-visit proportional cumulative odds model. For each visit(Week 4, Week 12, Week 24, Week 36 and Week 52), the proportional cumulative odds model uses PGIS change (order specified as "improved" < "unchanged" < "worsened") at the visit as response, and include treatment (LCZ696, Enalapril), baseline PGIS (None [Good], Mild, Moderate [Neither good nor bad], Severe, Very severe [Bad]), modified age group(6 years to < 18 years, 2 years to < 6 years, 1 month to < 2 years), NYHA/ROSS class group (Class I/II, Class III/IV) at randomization as fixed effect. The analysis cutoff date will be considered.

Based on the above model, the estimate and the 95% confidence interval will be provided for the adjusted odds ratios (LCZ696 over Enalapril) at each post-baseline scheduled visit (Week 4, Week 12, Week 24, Week 36 and Week 52), the corresponding two-sided p-values will also be provided.

Besides, the PGIS and the PGIS change will be summarized by modified age group, treatment group, and visit, using the numbers and percentages of patients in each category.

A sensitivity analysis will also be included without considering analysis cutoff date.

The FAS will be used for the above analyses.

2.7.4 Handling of missing values/censoring/discontinuations

In the analyses described in the above sections (Section 2.7.2 to 2.7.3), Category 1 event is considered the worst possible outcome and hence scheduled visits after Category 1 event will be imputed to be the worst case accordingly. For others, the missing data will not be imputed.

As a sensitivity analysis, a multiple imputation approach will be applied. Responses at scheduled visits after Category 1 event will be imputed to be the worst case. And all the other missing data for change in NYHA/ROSS class and change in PGIS will be considered as missing at random within each arm, and will be imputed as described below.

- Missing change from baseline in NYHA/ROSS class and change from baseline in PGIS will be imputed based on cumulative proportional odds models using fully conditional specifications. For each post-baseline scheduled visit (Week 4, Week 12, Week 24, Week 36, Week 52), the imputation model will include region, modified age group (6 years to < 18 years, 2 years to < 6 years, 1 month to < 2 years) at randomization, NYHA/ROSS class group (Class I/II, Class III/IV) at randomization, baseline value, and value at the immediately previous scheduled visit (Not applicable for Week 4, Week 4 for Week12,

Week 12 for Week 24, Week 24 for Week 36, Week 36 for Week 52) as fixed-effect factors.

The FAS will be used for the above sensitivity analyses.

2.7.5 Subgroup analysis for secondary endpoints

Subgroup analyses will be performed for secondary endpoints (Section 0 to 2.7.3) based on the subgroup variables in Table 2-6 without multiplicity adjustment.

Time to first adjudicated Category 1 or 2 event

For time to first adjudicated Category 1 or 2 event (Section 0), the subgroup analyses will be performed utilizing a Cox proportional hazard model, stratified by modified age group, with subgroup, treatment (LCZ696, Enalapril) and subgroup-by-treatment interaction included as fixed-effect factors.

Based on the Cox proportional hazard model, the two-sided p-value for the interaction will be provided, the estimate and the corresponding 95% confidence interval will be provided for the adjusted hazard ratio (LCZ696 over Enalapril) within each subgroup.

The estimates and the corresponding 95% confidence intervals for the adjusted hazard ratios within subgroups and for the overall adjusted hazard ratio will be presented graphically using a forest plot.

The FAS will be used for the subgroup analyses.

NYHA/ROSS class change and PGIS change

For NYHA/ROSS class change (Section 2.7.2) and PGIS change (Section 2.7.3), the subgroup analyses will be performed using a by-visit proportional cumulative odds model. For each visit (Week 4, Week 12, Week 24, Week 36 and Week 52), the proportional cumulative odds model uses corresponding NYHA/ROSS or PGIS change at the visit as response, and include treatment (LCZ696, Enalapril), corresponding baseline value (baseline NYHA/ROSS class or baseline PGIS), subgroup, subgroup by treatment, modified age group (6 years to < 18 years, 2 years to < 6 years, 1 month to < 2 years) as fixed effect.

Based on the above model, the estimate and the 95% confidence interval will be provided for the adjusted odds ratios (LCZ696 over Enalapril) at each post-baseline scheduled visit (Week 4, Week 12, Week 24, Week 36 and Week 52) within each subgroup, and the p-values for the subgroup-by-treatment interactions at all post-baseline scheduled visits will be also provided.

In addition, the estimates and the corresponding 95% confidence intervals for the adjusted odds ratios at Week 52 within subgroups and the overall adjusted odds ratio at Week 52 will be presented graphically using a forest plot.

The FAS will be used for the subgroup analyses.

[redacted]

[redacted]

[redacted]

[redacted]

2.9 Safety analyses

For Part 1, the SAF1 will be used for all safety analyses.

For Part 2, the SAF2 will be used for all safety analyses.

In addition, selected safety analysis will be done for patients who received expired drug.

2.9.1 Adverse events (AEs)

All AEs will be recorded on the CRF page "Adverse Events" and be identified using Medical Dictionary for Regulatory Activities (MedDRA). The MedDRA version used for reporting the study will be described in a footnote.

An AE with its severity increased should be considered and recorded as a new AE.

Part 1:

Treatment emergent adverse events (TEAEs) during the dose associated period are defined as any recorded AE with its start date (recorded or imputed) later than or equal to the date of the first dose of the study drug within the dose associated period AND its start date prior to or equal to the end date of the dose associated period.

TEAEs during the dose associated period will be summarized separately by age group, dose cohort, SOC, PT and maximum severity.

The following rules are applicable to the summaries.

- If a subject reported more than one adverse event with the same preferred term, the adverse event with the maximum severity will be presented.
- If a subject reported more than one adverse event within the same primary system organ class, the subject will be counted only once with the maximum severity at the system organ class level, where applicable.

Study drug related AEs are defined as any recorded AE with "Reasonable possibility that AE is related to study treatment" answered as "YES". Study drug related TEAEs during the dose associated period will be summarized by age group, dose cohort, SOC and PT.

Serious adverse events (SAEs) are defined as any recorded AE with "Does AE meet the definition of an SAE" answered as "Yes".

Treatment emergent SAEs during the dose associated period will be summarized by age group, dose cohort, SOC and PT.

TEAEs leading to death during the dose associated period will be summarized by age group, dose cohort, SOC and PT.

The SAF1 will be used for the above analyses for Part 1.

Part 2:

Treatment emergent adverse events (new or worsened) during the double-blind epoch are defined as any recorded AE with its start date (recorded or imputed) later than or equal to the date of the first dose of the study drug within the double-blind epoch.

TEAEs during the double-blind epoch will be summarized by:

1. modified age group, treatment group, SOC and PT,
2. modified age group, treatment group, SOC, PT and maximum severity.
3. modified age group, treatment group, Standard Medical Queries (SMQs) and PT.

The following rules are applicable to the summaries.

- If a patient reported more than one AE with the same PT during the double-blind epoch, the patient will be counted only once with the greatest severity at the PT level
- If a patient reported more than one AE within the same SOC during the double-blind epoch, the patient will be counted only once with the greatest severity at the SOC level, where applicable.

Study drug-related AEs are defined as any recorded AE with "Reasonable possibility that AE is related to study treatment" answered as "YES".

Study drug-related TEAEs during the double-blind epoch will be summarized by modified age group, treatment group, SOC and PT.

Serious adverse events (SAEs) are defined as any recorded AE with "Does AE meet the definition of an SAE" answered as "Yes".

Treatment emergent SAEs during the double-blind epoch will be summarized by modified age group, treatment group, SOC and PT.

AEs leading to permanent discontinuation of study treatment are defined as any recorded AE with "Action taken with study treatment" answered as "DRUG WITHDRAWN".

TEAEs leading to permanent discontinuation of study treatment during the double-blind epoch will be summarized by modified age group, treatment group, SOC and PT.

AEs leading to dose adjustment or temporarily interruption of study treatment are defined as any recorded AE with "Action taken with study treatment" answered as "DOSE REDUCED", "DOSE INCREASED" or "DRUG INTERRUPTED".

TEAEs leading to dose adjustment or temporarily interruption of study treatment during the double-blind epoch will be summarized by modified age group, treatment group, SOC and PT.

The most common TEAEs are defined as any recorded TEAE corresponding to a PT with at least 1% of patients in either treatment group having at least one TEAE of this PT during the double-blind epoch.

Most common TEAEs during the double-blind epoch will be summarized by modified age group, treatment group, SOC and PT, and will be presented in descending frequency according to the incidence in the LCZ696 group starting from the most common event.

Incidence of AEs will also be listed at a patient level by treatment group and modified age group including outcome, severity and action taken with the AE.

The SAF2 will be used for the above analyses for Part 2.

2.9.1.1 Adverse events of special interest / grouping of AEs

Specific AEs of interest will be summarized separately in addition to the above analysis:

- Anaphylaxis
- Angioedema
- Change in bone growth and density
- Cognitive impairment (Narrow SMQ)
- Embryo-fetal toxicity or lethality
- Hepatotoxicity 8 Hyperkalaemia
- Hypersensitivity (Narrow SMQ)
- Hypotension
- Malignancy
- Neonatal or infantile toxicity through exposure from breast milk
- Renal impairment (Narrow SMQ)
- Statin drug drug interaction.

Besides providing the crude percentages, exposure adjusted incidence rates per 100 patient-years will also be provided by treatment group.

The search paths for the related preferred terms (PTs), high level group term (HLGT), high level term (HLT), in Standard Medical Queries (SMQs), or NMQ in MedDRA for these risks are stored (or alternatively "summarized") in the latest version of LCZ696 Case Retrieval Strategy.

In addition to above standard analyses, for part 2, analysis for time-to-first selected AEs (Hypotension and Angioedema) by treatment group will be performed using Kaplan-Meier estimate.

The following subgroup analysis will be performed based on selected subgroups for Part 2 (gender, race, ACEI/ARB and region).

- Treatment emergent AEs during the double-blind epoch will be summarized within each subgroup, by treatment group, SOC and PT.

The renal

The SAF2 will be used for the above analyses for Part 2.

2.9.2 Death

In Part 1, the death and the investigator reported primary cause of death will be recorded on the CRF page "Death (End Point)". The death and primary cause of death will be listed based on SAF1.

In Part 2, death and primary cause of death will be reported by the investigator and adjudicated by the adjudication committee.

The investigator reported death and primary cause of death will be recorded on the CRF page: "Death (Endpoint)". The adjudicated death and primary cause of death will be recorded on the CRF page: "Death (Adjudication)".

The investigator reported death and primary cause of death will be summarized by modified age group and treatment group using the number and percentage of patients who died during the double-blind epoch (investigator reported), as well as the numbers and percentages of patients whose primary cause of death (investigator reported) is in each category and each sub-category.

Similarly, the adjudicated death and primary cause of death will be summarized by modified age group and treatment group using the number and percentage of patients who died during the double-blind epoch (adjudicated), as well as the numbers and percentages of patients whose primary cause of death (adjudicated) is in each category and each sub-category.

The SAF2 will be used for the above analyses for Part 2.

2.9.3 Laboratory data

2.9.3.1 General laboratory data

Complete laboratory evaluations (hematology, blood chemistry, and urine) as outlined in Table 2-17 will be performed in Part 1 at Visit 1 (screening), Visit 199 and/or 299, and/or Visit 301. Local laboratory will be used for Visit 1, 199, 299 and 301 for complete laboratory evaluations. Central laboratory should only be used for Visit 301, as unscheduled assessment, if local lab is unavailable.

Besides, it will also be performed in Part 2 at Visit 301, Visit 401, Visit 409 and Visit 416/499.

Local laboratory should be used for Visit 301 for complete laboratory evaluations. Central laboratory should only be used for Visit 301, if local lab is unavailable using unscheduled visit. Central laboratory must be used for Visit 401, 409 and 416/499 for complete laboratory evaluations for all patients.

Table 2-17 Laboratory examinations

Hematology	Biochemistry	Urine measurements**
Hematocrit	Alanine aminotransferase (ALT)	Specific gravity
Hemoglobin	Albumin (Alb)	pH
Platelet count	Alkaline phosphatase (ALP)	Glucose
Red blood cell count (RBC)	Aspartate aminotransferase (AST)	Protein (Total)
White blood cell count (WBC)	Blood urea nitrogen (BUN)	Ketones

Hematology	Biochemistry	Urine measurements**
WBC differential	Calcium	Bilirubin
	Magnesium	Urobilinogen
	Phosphate	Hemoglobin (blood)
Red blood cell distribution width (RDW)	Chloride	Leukocyte esterase
Mean corpuscular volume (MCV)	Creatinine*	Nitrite
Mean corpuscular hemoglobin concentration (MCHC)	Glucose	WBC
	Potassium*	RBC sediments
	Sodium*	Hyaline casts
	Bicarbonate	Granular casts
	Total bilirubin (TBL)	Waxy casts
	Fractionated bilirubin (if total bilirubin >2x ULN)	WBC casts
	Total protein	RBC casts
	Uric acid	
	eGFR*	

*Abbreviated laboratory evaluations must include these parameters

**Urinalysis with dipstick includes specific gravity, pH, glucose, total protein, bilirubin, ketones, urobilinogen, nitrite, leukocytes esterase and hemoglobin (blood). Other urine measurements listed are not required. If a urine dipstick is positive, other urine measurements such as a qualitative microscopic determination of WBC, RBC sediments (and casts) will also be measured.

Abbreviated safety laboratory evaluation consists of serum sodium, potassium, creatinine, and eGFR. Abbreviated safety laboratories will be performed locally. If the local laboratory is unavailable (either for abbreviated or complete labs), the central laboratory will be used as an unscheduled visit.

Abbreviated laboratory evaluations are to be performed at Part 1 PK/PD Visits 101 (post-dose), 201 (post-dose), 103 to 120 (when patient comes to the site for a visit about every three months), 203 to 220 (when patient comes to the site for a visit about every three months), UNS PK/PD; and also performed in Part 2 at Visit 402, Visit 403, Visit 404 (optional visit), Visit 405, Visit 406, Visit 412, Visit 415 and as an optional assessment at unscheduled study treatment discontinuation (UNS TD), and unscheduled (UNS) visit.

For those patients who have Visits 299 and 301 on the same day (for second PK dose patient) or Visits 199 and 301 on the same day (for the first PK dose only patient), laboratories will be performed only once.

Local laboratory results will allow investigators to proceed with study visit procedures without the need to wait for central laboratory results. The local laboratory results must be recorded on the appropriate CRF.

Both complete laboratory evaluations and abbreviated laboratory evaluations will be included in the summaries for laboratory data. In general, local lab and central lab data will be analyzed separately. For event summary, local lab and central lab will be summarized together.

Test values based on the local laboratory assessments and the central laboratory assessments will be summarized separately by laboratory parameter, modified age group, treatment group and visit, using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum;

The change from baseline in test values based on the local laboratory assessments and the central laboratory assessments will also be summarized separately by laboratory parameter, modified age group, treatment group and visit, using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum. The baseline values for local laboratory and central laboratory should be computed separately.

For each laboratory parameter and each specific assessment, the test value will be categorized into normal, low, or high, according to the corresponding normal range.

For each laboratory parameter and each specific post-baseline visit, the shift from baseline to the visit is defined as the shift from the baseline category to the categorized post-baseline test value (low to low, low to normal, low to high, normal to low, normal to normal, normal to high, high to low, high to normal, high to high).

The shift from baseline will be summarized by laboratory parameter, modified age group, treatment group and visit, using the numbers and percentages of patients in each category. By-visit summary will include scheduled (central or local) assessments.

For each laboratory parameter, the minimum and maximum post-baseline category during the double-blind epoch are defined based on the categorized post-baseline test values according to Table 2-18. All scheduled or unscheduled, central or local assessments will be included.

Table 2-18 Minimum and maximum post-baseline category

Range of Categorized Post-Baseline Test Values	Minimum post-baseline category	Maximum post-baseline category
Low, Normal, High, Missing	Low	High
Low, Normal, High	Low	High
Low, Normal, Missing	Low	Normal
Low, Normal	Low	Normal
Low, High, Missing	Low	High
Low, High	Low	High
Low, Missing	Low	Low
Low	Low	Low
Normal, High, Missing	Normal	High
Normal, High	Normal	High
Normal, Missing	Normal	Normal
Normal	Normal	Normal
High, Missing	High	High
High	High	High
Missing	Missing	Missing

The shift from baseline to minimum and maximum post-baseline category during the double-blind epoch will be summarized separately by laboratory parameter, modified age group and treatment group, using the numbers and percentages of patients in each category.

Table 2-19 provides criteria for clinically notable laboratory abnormality.

The shift from baseline in laboratory parameters will be summarized by laboratory parameter, age group, dose cohort, dose associated analysis visit, using the numbers and percentages of patients in each category. By-visit summary will only include scheduled assessments in the dose associated periods. Unscheduled assessments may be used for imputation if needed.

In Part 1, the clinically notable laboratory abnormality will be calculated based on the local laboratory assessments.

The clinically notable laboratory abnormality during the dose associated period will be summarized by criteria, age group and dose cohort, using the numbers and percentages of patients satisfying such criteria.

Table 2-19 Clinical notable criteria for laboratory values

Parameter	Conventional Units	SI Alert Value	SI Units
Hematology			
Red Blood Cell Count	x10E6/uL	>50% increase and >ULN; >30% decrease and <LLN.	x10E12/L
Hemoglobin	g/dL	>50% increase and >ULN; (>30% decrease and <LLN) or any value <70.	g/L
Hematocrit	%	>50% increase and >ULN; >30% decrease and <LLN.	L/L
White Blood Cell Count	x10E3/uL	>50% increase and >ULN; >50% decrease and <LLN.	x10E9/L
Platelet Count	x10E3/uL	>75% increase and >ULN; >50% decrease and <LLN.	x10E9/L
Chemistry			
BUN	mg/dL	>50% increase and >ULN	mmol/L
Creatinine	mg/dL	>50% increase and >ULN	umol/L
Albumin	g/dL	<20	g/L
Glucose	mg/dL	>50% increase and >ULN; (>50% decrease and <LLN) or any value <3.3.	mmol/L
Total Bilirubin	mg/dL	>100% increase and >ULN	umol/L
AST (SGOT)	U/L	>150% increase and >ULN	U/L
ALT (SGPT)	U/L	>150% increase and >ULN	U/L
Sodium	mEq/L	(>5% increase and >ULN) or any value >150; (>5% decrease and <LLN) or any value <125	mmol/L
Potassium	mEq/L	(>20% increase and >ULN) or any value >6; (>20% decrease and <LLN) or any value <3.	mmol/L
Chloride	mEq/L	>10% increase and >ULN; >10% decrease and <LLN	mmol/L

Parameter	Conventional Units	SI Alert Value	SI Units
Calcium	mg/dL	>10% increase and >ULN; >10% decrease and <LLN.	mmol/L
Uric Acid	mg/dL	>50% increase and >ULN	mmol/L

In the above table, increase and decrease are defined as compared to the baseline value.

The clinically notable laboratory abnormality during the double-blind epoch based on the local laboratory assessments and the central laboratory assessments will be summarized by criteria, modified age group and treatment group, using the numbers and percentages of patients satisfying such criteria.

The SAF2 will be used for the above analyses. Listing will be provided.

2.9.3.2 Special laboratory data

Liver events

Table 2-20 provides criteria for liver laboratory triggers and liver events.

The liver laboratory triggers and liver events during the double-blind epoch based on the local laboratory assessments and the central laboratory assessments will be summarized by criteria, modified age group and treatment group, using the numbers and percentages of patients satisfying such criteria.

Table 2-20 Liver event and laboratory trigger definitions

	Definition/ threshold
LIVER LABORATORY TRIGGERS	• 3 x ULN < ALT / AST ≤ 5 x ULN • 1.5 x ULN < TBL ≤ 2 x ULN
LIVER EVENTS	• ALT or AST > 5 × ULN • ALP > 2 × ULN (in the absence of known bone pathology) • TBL > 2 × ULN (in the absence of known Gilbert syndrome) • ALT or AST > 3 × ULN and INR > 1.5 • Potential Hy's Law cases (defined as ALT or AST > 3 × ULN and TBL > 2 × ULN [mainly conjugated fraction] without notable increase in ALP to > 2 × ULN) • Any clinical event of jaundice (or equivalent term) • ALT or AST > 3 × ULN accompanied by (general) malaise, fatigue, abdominal pain, nausea, or vomiting, or rash with eosinophilia • Any adverse event potentially indicative of a liver toxicity*

*These events cover the following: hepatic failure, fibrosis and cirrhosis, and other liver damage-related conditions; the non-infectious hepatitis; the benign, malignant and unspecified liver neoplasms
TBL: total bilirubin; ULN: upper limit of normal

The SAF2 will be used for the above analyses.

Renal events

The renal events during the double-blind epoch based on the local laboratory assessments and the central laboratory assessments will be summarized by criteria, modified age group and treatment group, using the numbers and percentages of patients satisfying such criteria.

The renal imparment related events during the double-blind epoch will be summarized for age group 1 and 2 for baseline eGFR≤170 mL/min/1.73 m^2 and baseline eGFR >170 mL/min/1.73 m^2 separately. And additional output for age group 3 will be presented. Normal ranges for eGFR will be provided in the reference (Heilbron 1991).

The SAF2 will be used for the above analyses.

2.9.4 Vital signs

In Part 1, the test values and the changes from baseline in vital signs (pulse rate, sitting/supine systolic blood pressure, sitting/supine diastolic blood pressure, height, head circumference and body weight) will be summarized by vital sign parameter, age group, dose cohort, dose associated analysis visit, and time point (if applicable), using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum. By-visit summary will only include scheduled visits in the dose associated periods. Unscheduled assessments may be used for imputation if needed.

In Part 2, the test values and the changes from baseline in vital signs (sitting/supine pulse rate, sitting/supine systolic blood pressure, sitting/supine diastolic blood pressure, height, head circumference and body weight) will be summarized by vital sign parameter, modified age group, treatment group, visit, and time point (if applicable), using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum. By-visit summary will only include scheduled assessments. Unscheduled assessments may be used for imputation if needed.

Table 2-21 provides criteria for clinically notable vital sign.

In Part 1, the clinically notable vital signs during the dose associated period will be summarized by criteria, age group and dose cohort, using the numbers and percentages of patients with at least one post-baseline assessment during the corresponding period satisfying such criteria, and the numbers and percentages of patients with baseline assessment not satisfying such criteria and at least one post-baseline assessment during the corresponding period satisfying such criteria.

In Part 2, the clinically notable vital signs during the double-blind epoch will be summarized by criteria, modified age group and treatment group, using the numbers and percentages of patients with at least one post-baseline assessment during the double-blind epoch satisfying such criteria, and the numbers and percentages of patients with baseline assessment not satisfying such criteria and at least one post-baseline assessment during the double-blind epoch satisfying such criteria.

Table 2-21 Criteria for clinically notable vital signs

Age	HR [min^{-1}]	SBP [mmHg]	DBP [mmHg]
1-3 months	<90, >160	<60, >95	<40, >50
3-6 months	<80, >130	<60, >100	<40, >60

Age	HR [min^{-1}]	SBP [mmHg]	DBP [mmHg]
6-12 months	<70, >130	<60, >110	<45, >70
1-3 years	<60, >120	<76, >115	<45, >75
3-6 years	<55, >120	<82, >120	<50, >80
6-12 years	<50, >105	<90, >130	<50, >80
>12 years	<45, >95	<90, >145	<55, >90

The change from baseline in height and height Z-score at end of study will be provided by treatment group and modified age group, using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum.

The SAF1 will be used for the Part 1 analyses, and the SAF2 will be used for the Part 2 analyses.

2.9.5 Electrocardiogram (ECG)

For each ECG parameter (QTcF, heart rate, PR duration, QT duration, QRS duration), the test value and the change from baseline will be summarized using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum, by ECG parameter, age group, dose cohort and dose associated analysis visit for Part 1, and also by ECG parameter, modified age group, treatment group and visit for Part 2.

For each ECG assessment, a clinically significant abnormality is defined as an answer of "Yes" to the question "any new clinically significant abnormalities" on the CRF page "12-Lead ECG Evaluation – Local Analysis".

For each specific post-baseline visit, a new onset clinically significant abnormality is defined as a clinically significant abnormality at the visit while no clinically significant abnormality at baseline.

The number and percentage of patients with new onset clinically significant abnormality will be provided by age group, dose cohort and dose associated analysis visit for Part 1, and also by modified age group, treatment group and visit for Part 2.

The number and percentage of patients with at least one post-baseline new onset clinically significant abnormality during the dose associated period will be provided for Part 1, and also will be provided for Part 2 during the double-blind epoch.

The SAF1 will be used for the Part 1 analyses, and the SAF2 will be used for the Part 2 analyses.

2.10 Pharmacokinetic (PK) endpoints

See section 2.5.1.

The population PK Model will be updated with sparse PK data from a subset of Part 2 Group 2 patients for the purpose of further confirming the target dose for Age Group 2. This will be performed prior to database lock and requires unblinding, and as such will be done by an unblinded team who is not involved in the conduct of the study.

2.11 Pharmacodynamics (PD) and PK/PD analyses

See section 2.5.2 for Part 1.

In Part 2, pre-dose PK blood sample will be collected at Visit 402, Visit 406, and Visit 416/499 for population PK analysis. If a PK sample is not obtained at Visit 402 or 406, the PK sample can be collected at a subsequent scheduled visit.

[redacted]

The population PK/PD analysis will be in a separate report.

2.12 Patient-reported outcomes

See Section 2.7 and 2.8.

2.13 Biomarkers

See section 2.5 for Part 1.

[redacted]

2.15 Interim analysis

Safety monitoring

An external DMC will monitor patient safety data during the course of the study. For this study, the DMC will review safety data on a regular frequency of every six months. DMC may request additional safety data review. Such safety analyses do not inflate the type I error for the primary efficacy hypothesis testing and thus require no multiplicity adjustments.

Besides, the safety, tolerability, pharmacokinetic, and pharmacodynamic data from open label use of LCZ696 will be evaluated by Novartis and the DMC to advance patients safely for each age group when all the patients have completed Part 1 of the study in that age group. Details will be included in a separate interim statistical analysis plan.

Interim efficacy analysis

It is planned to have one formal interim efficacy analysis, when at least 180 patients (at least 36 patients from each modified age group) have completed the study (i.e., had a positively adjudicated Category 1 event or completed the double-blind epoch), and at least 40 patients have had a positively adjudicated event in Category 1 or Category 2. Some adjustment to the time of interim analysis may be made to coincide with the regular DMC meetings.

For the planned interim analysis, the primary test will be performed at a nominal significance level of one-sided 0.00153 (calculated based on the O'Brien-Fleming boundary with a sample size fraction of 0.5). The above interim analysis or the analysis set will comprise all patients who have completed the study (i.e., had a positively adjudicated Category 1 event or completed the double-blind epoch) or discontinued from the study by the cutoff date.

For the final analysis, appropriate adjustment will be made according to the actual sample size fraction to control the overall Type I error rate at two-sided 0.05.

Table 2-22 provides several examples for nominal significance levels based on actual sample size fractions.

Table 2-22 Nominal significance levels and actual sample size fraction

Actual Sample Size Fraction at the Interim Analysis	Nominal Significance Level		
	Interim Analysis (One-sided)	Final Analysis (One-sided)	Final Analysis (Two-sided)
0.20	0.00153	0.02390	0.04780
0.30	0.00153	0.02410	0.04819
0.40	0.00153	0.02430	0.04860
0.50	**0.00153**	**0.02450**	**0.04900**
0.60	0.00153	0.02469	0.04938
0.70	0.00153	0.02485	0.04970
0.80	0.00153	0.02496	0.04992

The trial may only be concluded early for efficacy, if the primary test is significant at level of one-sided 0.00153 (pre-specified boundary at the interim analysis) and the overall risk/benefit profile is considered positive.

Besides, a futility analysis will also be incorporated, whereby the study may be stopped if the conditional power for having a significant positive result at the final analysis, based on a standardized effect size which will produce a power of at least 80% at the final analysis for the sample size of 360 (or re-estimated sample size), is less than 10%. Here the standardized effect size is chosen to be 0.147 corresponding to a MW probability of 0.585 for continuous distributions, which mimics the case that we have broken all ties. See Appendix 5.2.2 for derivation details.

The interim analysis will be performed by an external independent statistician at an external data analysis center, who will not be involved in the study conduct. The results will be reviewed by the independent DMC. Investigators and others who are involved in the conduct of the trial and in the analysis of the final trial results, or who have contact with study centers, will remain blinded to the treatment codes and interim analysis results until all monitoring decisions have been made and the database has been unblinded for final analysis.

If the study is stopped early at the interim analysis due to overwhelming efficacy, the patients used in the interim analysis will be used for the CSR. In this case, at Day 30 (Week 4), Day 90 (Week 12) and Day 180 (Week 24), the global rank endpoints (as well as their components) will be defined similarly as Section 2.6.1 and will be analyzed using the same approach in Section 2.6.2. This sensitivity analysis will include all patients randomized by the cutoff date

with a positively adjudicated Category 1 event occurred or the corresponding visit [Week 4, Week 12, Week 24] completed.

The FAS will be used for this supportive analyses.

3 Sample size calculation

3.1 Sample size for Part 1

From a prior clinical study in adult HF patients, the observed CV% for body-weight normalized LBQ657 exposure was 37.1%, and the observed CV% for body-weight normalized clearance and volume of distribution of valsartan was 39.4% and 49.6%, respectively (Study CLCZ696A2117). Therefore, 49.6% of CV% was assumed for sample size estimations. A sample size of 12 observations (except Group 3 where N=6 observations) is expected to have 80% probability that the 95% CI of the point estimate for the geometric mean estimates of clearance and volume of distribution will be between 60% and 140% range (95% CI: 71 – 141%).

3.2 Sample size for the primary analysis for Part 2

In the sample size evaluation for Part 2, the stratification factors will be ignored.

The planned sample size of 360 is to provide adequate power for the primary test.

The power calculation is based on the data from the Carvedilol pediatric heart failure (CPHF) study (Shaddy et al. 2007) within the subgroup of patients with heart failure due to systemic left ventricle systolic dysfunction.

Table 3-1 presents the assumed underlying probabilities for Category 1 to 5 in the global rank endpoint (Section 2.6.1) at Month 8, based on data from the CPHF study (Shaddy et al. 2007).

- The percentages of Category 1 patients at Month 8 in the LCZ696 group and in the Enalapril are assumed to be the observed percentages of dead patients in the Carvedilol group and in the placebo group in the CPHF study (Figure 2 in Shaddy et al. 2007).
- The percentages of Category 2 patients at Month 8 in the LCZ696 group and in the Enalapril are assumed to be the observed percentages of alive patients with heart failure hospitalization in the Carvedilol group and in the placebo group in the CPHF study (Figure 2 in Shaddy et al. 2007).
- The percentages of patients in Category 5 for the LCZ696 group and for the Enalapril are assumed to be the observed percentages of patients who were categorized as “improved” for the primary endpoint in the Carvedilol group and in the placebo group in the CPHF study (Table 2 in Shaddy et al. 2007) within the subgroup of patients with systemic left ventricle.
- The percentages of patients in Category 3 and 4 are formed by splitting the remaining percentages approximately equally for these two categories.

Table 3-1 also presents the assumed underlying probabilities for Category 1 to 5 in the global rank endpoint (Section 2.6.1) at Week 52, whereby, the percentages for Category 1 and 2 are calculated based on an assumption of the exponential distribution, extending from Month 8 to Month 12; the remaining percentages within treatment group are distributed to Categories 3 to

5 according to the percentages at Month 8 (i.e., 9 : 11 : 64 for the LCZ696 group, and 14 : 13 : 51 for the enalapril group).

Table 3-1 Assumed percentage of each category in the global rank endpoint at Month 8 (Week 32) and Month 12 (Week 52)

Category	Month 8 (Week 32)		Month 12 (Week 52)	
	LCZ696 (%)	Enalapril (%)	LCZ696 (%)	Enalapril (%)
1	6	9	9	13
2	10	13	15	19
3	9	14	8.1	12.2
4	11	13	10.0	11.3
5	64	51	57.9	44.5

Assuming the percentages in Table 3-1, for the response category (Category 1 to Category 5) based on the ranking algorithm (Section 2.6.1), the MW probability is approximately 0.570, which leads to a MW odds of approximately 0.753; for a MW analysis on the response category, the power is approximately 70% with a significance level of two-sided 0.05 and a sample size of 354 (177 patients for each treatment group). Using a global ranked endpoint, the power under the same assumptions is expected to increase. While the exact increase in power is dependent on the distribution of the ranking within each category, simulation studies have demonstrated that the power for the global rank endpoint increases 10-20% more than with the ordered categorical analysis when the power of the ordered categorical endpoint is more than 50% (Sun et al. 2012). Based on the assumptions stated, the power is estimated to be at least 80% for this study.

The sample size is further increased to 360 (180 patients for each treatment group) to account for the alpha adjustment for the interim analysis.

Assuming the same percentages and response categories in Table 3-1, the following arguments illustrate how the analysis power will be increased by the effective tie breaking based on the global rank endpoint within each of the 5 response categories.

The MW probability based on the global rank endpoint can be decomposed into

$$\text{MW probability} = (1-\varrho)\psi_0 + \varrho(1-\omega)\psi + \varrho\frac{1}{2}\omega,$$

where ψ_0 is the conditional probability of the patient from the LCZ696 group having a better response category than the patient from the Enalapril group given that they are not in the same response category, ψ is the conditional probability of the patient from the LCZ696 group being better response (with global rank endpoint) than the patient from the Enalapril group given that they are in the same response category, ω is the conditional probability of the two patients having the same response given that the two patients are in the same response category, ϱ is the probability of the patient from the LCZ696 group and the patient from the Enalapril group sharing the same response category, when the two patients are independently sampled from the LCZ696 group and the Enalapril group.

The parameter, ψ, characterizes the effect of the tie breaker, while the parameter, ψ_0, can be considered as the effect of the response category (original 5 categories in Table 3-1), $1 - \omega$ can

be considered as the proportion of between-treatment-group ties being broken within each category (assumed to be the same across all the categories).

Assuming the same percentages in Table 3-1, we have $\varrho = 0.319$, $\psi_0 \approx 0.604$, and

$$\text{MW probability} \approx (1 - 0.319) \times 0.604 + 0.319 \times (1 - \omega) \times \psi + 0.319 \times \frac{1}{2}\omega.$$

Hence, the MW probability can be viewed as a function of (ψ, ω).

Assuming that a tie-breaker has the tie-breaking effect of $\psi = \psi_0 \approx 0.604$, Table 3-2 presents the powers for a total sample size of 360 and the sample sizes for a power of 80% for this tie breaker with different proportions $(1 - \omega)$ of ties being broken, where the upper bound of the variance parameter (Appendix 5.2.3) is used in the calculation (which may lead to the presented powers slightly under-estimated and the sample sizes slightly over-estimated).

According to Table 3-2, with 50% between-treatment-group ties resolved, the power for the primary test would be at least 80%.

Table 3-2 Power and sample size based on within-category proportions of between-treatment-group ties resolved ($1 - \omega$) when $\psi = \psi_0 \approx 0.604$

Within-Category Proportion of Between-Treatment-Group Ties Resolved ($1 - \omega$)	MW Probability	Power for a Total Sample Size of 360	Sample Size for a Power of 80%
10%	0.574	68%	450*
20%	0.577	72%	436
30%	0.581	76%	401
40%	0.584	79%	370
50%	**0.587**	**82%**	**342**
60%	0.591	85%	318
70%	0.594	87%	296
80%	0.597	89%	276
90%	0.601	91%	258
100%	0.604	93%	242

Note *: Since the upper bound of the variance parameter (Appendix 5.2.3) is used in the calculation, the presented powers may be under-estimated and the sample sizes may be over-estimated. The sample size corresponding to $1 - \omega = 10\%$ is 476 with this upper bound of the variance, which has been truncated to 450 in the Table since the note below.

Note that, assuming the percentages in Table 3-1, without breaking any tie, a total sample size of 450 will provide approximately 80% power for the primary test.

3.3 Sample size re-estimation for Part 2

In the sample size evaluation and sample size re-estimation for Part 2, the stratification factors will be ignored.

As discussed in the Section 3.2, assuming the percentages in Table 3-1, the effect size of the primary analysis depends on the percentage $(1 - \omega)$ of between-treatment ties resolved within each category, given the effectiveness (ψ) of ties breaker.

The value of ω is usually unknown and requires an assessment during the study to ensure the power for the primary test close to the designed power of 80%.

A blinded sample size re-estimation will be performed to avoid a penalty from unblinding the treatment codes in the middle course of the study. It will be performed, when it is about six months prior to the predicted end of patient enrollment, or when at least 180 patients have completed the study and at least 40 patients have had a positively adjudicated event in Category 1 or Category 2, whichever occurs the first.

When treatment codes are blinded, ω will not be identifiable itself. We will use the pooled probability of observing any tie for the global rank endpoint within any response category (Category 1 to Category 5) regardless of the treatment group (Appendix 5.2.3), noted as λ, to replace it, which can be considered as a reasonable "estimate" of ω. In our sample size re-estimation, we will assume the percentages in Table 3-1 and effectiveness of tie breaker $\psi = \psi_0 \approx 0.604$. With these assumptions, the MW probability is bounded below by

$$\text{MW probability} \geq \Psi = (1-\varrho)\cdot\psi_0 + \varrho\cdot(1-\lambda)\cdot\psi + \varrho\cdot\frac{1}{2}\cdot\lambda \approx 0.604 - 0.033\times\lambda.$$

The parameter, λ, will be estimated using the pooled data available at the time for the sample size re-estimation (Appendix 5.2.3), denoted as λ_M. Consequently, the MWP will be estimated using

$$\Psi_M = 0.604 - 0.033\times\lambda_M.$$

In addition, the variance parameter will be estimated by Ω_M, using the pooled data (Appendix 5.2.3). The sample size re-estimation for 80% power will be based on the following formula:

$$N_M = 4\times\Omega_M\times\left(\frac{\Psi^{-1}(0.975)+\Psi^{-1}(0.8)}{\Psi_M - 0.5}\right)^2.$$

On the other hand, we also require 80 events in Categories 1 and 2, the final re-estimated sample size will be

$$\max\{\min\{N_M, 450\}, N_{(80)}\},$$

where $N_{(80)} = 80/(P_{N,1} + P_{N,2})$ is the sample size needed for obtaining 80 adjudicated Category 1 or 2 events; $P_{N,1}$ and $P_{N,2}$ are the estimated proportions of patients in Categories 1 and Categories 2 based on the pooled data, respectively. The number of 450 is used as a sample size cap under the assumption that our study would have an event distribution outlined in Table 3-1 as the worst scenario.

Detailed formula for the sample size re-estimation are provided in Appendix 5.2.3.

4 Change to protocol specified analyses

Above efficacy analysis also considers NYHA/ROSS class group (Class I/II, Class III/IV) at randomization, i.e. the randomization stratification.

5 Appendix

5.1 Imputation rules

5.1.1 Missing or partially missing AE or concomitant medication start/end date

The missing or partially missing start/end date for AEs and prior/concomitant therapies will be imputed using the Novartis AGB global standard approach. Details will be provided in the study PDS.

5.1.2 Missing visit date

If the visit date is missing, the scheduled date (per protocol) of the visit will be used.

5.1.3 Event date missing or partially missing

If the date of an event is not known or is incomplete, the imputation rules are:

a) If the day of the event is unknown, then the 15th day of this month will be imputed for a missing day;

b) If only the month is unknown, then July will be used for imputation of the missing;

c) If only the year of the event is known, then the 1st of July will be imputed for a missing day and month;

d) The above rules are only for general case. If there is additional information available for the missing date, then the information should be used and the imputation of missing date should be treated differently. For example, if an event occurs between two visits and its date is missing, then the date in the middle of these visits may be used.

5.2 Statistical models

5.2.1 Primary analysis

Notations

Suppose that there are totally K strata.

Let $N_{k,0}$ and $N_{k,1}$ be the number of patients in the Enalapril group and in the LCZ696 group in stratum k, $N_k = N_{k,0} + N_{k,1}$ be the total number of patients in stratum k, $k = 1, \ldots, K$.

Let $F_{N,k,0}$ and $F_{N,k,1}$ be the empirical cumulative distribution function for the response variable in the Enalapril group and in the LCZ696 group in stratum k, $F_{N,k} = (N_{k,0}/N_k) \times F_{N,k,0} + (N_{k,1}/N_k) \times F_{N,k,1}$ be the pooled empirical cumulative distribution function, $k = 1, \ldots, K$.

Stratified Wilcoxon rank-sum test

For stratum *k*, the Wilcoxon rank-sum statistic is

$$\mathrm{WRS}_{\mathrm{N},k} = \frac{1}{2}\mathrm{N}_{k,1} + \frac{1}{2}\mathrm{N}_{k,1}^2 + \mathrm{N}_{k,1}\mathrm{N}_{k,0}\int \frac{1}{2}\left(\mathrm{F}_{\mathrm{N},k,0}(x) + \mathrm{F}_{\mathrm{N},k,0}(x-)\right) d\mathrm{F}_{\mathrm{N},k,1}(x)$$

whose mean and variance under the null hypothesis will be estimated by

$$\mathrm{MWRS}_{\mathrm{N},k} = \frac{1}{2}\mathrm{N}_{k,1} + \frac{1}{2}\mathrm{N}_{k,1}^2 + \frac{1}{2}\mathrm{N}_{k,1}\mathrm{N}_{k,0}, \mathrm{VWRS}_{\mathrm{N},k} = \left(\mathrm{N}_{k,1}\mathrm{N}_{k,0}\right)^2\left(\frac{\mathrm{V}_{\mathrm{N},k,0}^{\mathrm{NULL}}}{\mathrm{N}_{k,0}} + \frac{\mathrm{V}_{\mathrm{N},k,1}^{\mathrm{NULL}}}{\mathrm{N}_{k,1}}\right),$$

where

$$\mathrm{V}_{\mathrm{N},k,0}^{\mathrm{NULL}} = \int \left(\frac{1}{2}\left(\mathrm{F}_{\mathrm{N},k,1}(x) + \mathrm{F}_{\mathrm{N},k,1}(x-)\right) - \frac{1}{2}\right)^2 d\mathrm{F}_{\mathrm{N},k,0}(x),$$

$$\mathrm{V}_{\mathrm{N},k,1}^{\mathrm{NULL}} = \int \left(\frac{1}{2}\left(\mathrm{F}_{\mathrm{N},k,0}(x) + \mathrm{F}_{\mathrm{N},k,0}(x-)\right) - \frac{1}{2}\right)^2 d\mathrm{F}_{\mathrm{N},k,1}(x).$$

The stratified Wilcoxon rank-sum statistic is

$$\mathrm{WRS}_{\mathrm{N}} = \sum_{k=1}^{\mathrm{K}} \frac{\mathrm{WRS}_{\mathrm{N},k}}{\mathrm{N}_k + 1},$$

whose mean and variance under the null hypothesis will be estimated by

$$\mathrm{MWRS}_{\mathrm{N}} = \sum_{k=1}^{\mathrm{K}} \frac{\mathrm{MWRS}_{\mathrm{N},k}}{\mathrm{N}_k + 1} = \frac{1}{2}\sum_{k=1}^{\mathrm{K}} \mathrm{N}_{k,1},$$

$$\mathrm{VWRS}_{\mathrm{N}} = \sum_{k=1}^{\mathrm{K}} \frac{\mathrm{VWRS}_{\mathrm{N},k}}{(\mathrm{N}_k + 1)^2} = \sum_{k=1}^{\mathrm{K}} \left(\frac{\mathrm{N}_{k,1}\mathrm{N}_{k,0}}{\mathrm{N}_k + 1}\right)^2 \left(\frac{\mathrm{V}_{\mathrm{N},k,0}^{\mathrm{NULL}}}{\mathrm{N}_{k,0}} + \frac{\mathrm{V}_{\mathrm{N},k,1}^{\mathrm{NULL}}}{\mathrm{N}_{k,1}}\right).$$

The testing statistic for the stratified Wilcoxon rank-sum statistic is

$$\mathrm{TWRS}_{\mathrm{N}} = \mathrm{VWRS}_{\mathrm{N}}^{-1/2}(\mathrm{WRS}_{\mathrm{N}} - \mathrm{MWRS}_{\mathrm{N}}),$$

which leads to a one-sided p-value (right-side) of

$$\mathrm{PVRS}_{\mathrm{N}} = 1 - \Phi(\mathrm{TWRS}_{\mathrm{N}}),$$

and a two-sided p-value of

$$\mathrm{PVTS}_{\mathrm{N}} = 2\left(1 - \Phi(|\mathrm{TWRS}_{\mathrm{N}}|)\right),$$

where Φ is the cumulative distribution function for the standard Gaussian distribution.

Stratified estimation for the MW probability and the MW odds

For stratum *k*, the MW probability within stratum *k* will be estimated by

$$\mathrm{MWP}_{\mathrm{N},k} = \frac{\mathrm{WRS}_{\mathrm{N},k} - \mathrm{N}_{k,1}\left(\mathrm{N}_{k,1} + 1\right)/2}{\mathrm{N}_{k,1}\mathrm{N}_{k,0}} = \int \frac{1}{2}\left(\mathrm{F}_{\mathrm{N},k,0}(x) + \mathrm{F}_{\mathrm{N},k,0}(x-)\right) d\mathrm{F}_{\mathrm{N},k,1}(x),$$

whose variance will be estimated by

$$\mathrm{VMWP}_{\mathrm{N},k} = \frac{\mathrm{V}_{\mathrm{N},k,0}}{\mathrm{N}_{k,0}} + \frac{\mathrm{V}_{\mathrm{N},k,1}}{\mathrm{N}_{k,1}},$$

where

$$V_{N,k,0} = \int \left(\frac{1}{2}\left(F_{N,k,1}(x) + F_{N,k,1}(x-)\right) - \left(1 - MWP_{N,k}\right)\right)^2 dF_{N,k,0}(x),$$

$$V_{N,k,1} = \int \left(\frac{1}{2}\left(F_{N,k,0}(x) + F_{N,k,0}(x-)\right) - MWP_{N,k}\right)^2 dF_{N,k,1}(x).$$

The overall MW probability is defined as

$$MWP = \sum_{k=1}^{K} \frac{N_{k,0}N_{k,1}}{N_k + 1} MWP_k \Big/ \sum_{k=1}^{K} \frac{N_{k,0}N_{k,1}}{N_k + 1},$$

where MWP_k denotes the MW probability in stratum k.

The overall MW probability will be estimated by

$$MWP_N = \sum_{k=1}^{K} \frac{N_{k,1}N_{k,0}}{N_k + 1} MWP_{N,k} \Big/ \sum_{k=1}^{K} \frac{N_{k,1}N_{k,0}}{N_k + 1},$$

whose variance will be estimated by

$$VMWP_N = \left[\sum_{k=1}^{K} \left(\frac{N_{k,0}N_{k,1}}{N_k + 1}\right)^2 VMWP_{N,k}\right] \Big/ \left[\sum_{k=1}^{K} \frac{N_{k,0}N_{k,1}}{N_k + 1}\right]^2.$$

The 1-α confidence interval for the overall MW probability will be given as

$$\left(\frac{1}{1 + \exp\{\theta_N^U\}}, \frac{1}{1 + \exp\{\theta_N^L\}}\right),$$

where

$$\theta_N^L = \log\frac{1 - MWP_N}{MWP_N} - \frac{\Phi^{-1}(1 - \alpha/2)VMWP_N^{1/2}}{MWP_N(1 - MWP_N)},$$

$$\theta_N^U = \log\frac{1 - MWP_N}{MWP_N} + \frac{\Phi^{-1}(1 - \alpha/2)VMWP_N^{1/2}}{MWP_N(1 - MWP_N)},$$

with Φ^{-1} being the quantile function for the standard Gaussian distribution.

The overall MW odds will then be estimated by

$$MWO_N = \frac{1 - MWP_N}{MWP_N},$$

the 1-α confidence interval for the overall MW odds will be given as

$$\left(\exp\{\theta_N^L\}, \exp\{\theta_N^U\}\right)$$

5.2.2 Futility analysis

Suppose there are M patients for the planned efficacy interim analysis.

Conditioning on the testing statistic for the interim analysis, the conditional power for the test at the final analysis is approximately

$$CP_{N,M} = 1 - \Phi\left(\sqrt{\frac{N}{N - M}}\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) - \sqrt{\frac{M}{N - M}}\Phi^{-1}(1 - PVRS_M) - \sqrt{N - M}\cdot\delta\right),$$

where $PVRS_M$ is the one-sided (right-side) p-value for the primary test at the interim analysis, N is the sample size at the final analysis, α is the two-sided nominal significance level for the final analysis based on the sample size fraction M/N, and δ is an assumed standardized effect size:

$$\delta = \frac{\text{MWP} - 0.5}{2\sqrt{\Omega}}$$

with

$$\Omega = \mathbf{Var}\left(\mathbf{P}(Y_i < Y_j | Y_j) + \frac{1}{2}\mathbf{P}(Y_i = Y_j | Y_j)\right) = \frac{1}{12}\left[1 - \sum_x \left(\mathbf{P}(Y_i = x)\right)^3\right],$$

in which, the sum in the last term is over the mass point.

5.2.3 Sample size calculation and blinded re-estimation

Notations

Let Y_i be the global rank endpoint for patient i, G_i be the group indicator for patient i ($G_i = 1$ if patient i is in the LCZ696 group, $G_i = 0$ if patient i is in the Enalapril group), $C(Y_i)$ be the response category (Category 1 to Category 5) for patient i.

Let patient i and patient j be two independent patients.

Let ϱ be the probability of a patient from the LCZ696 group and a patient from the Enalapril group sharing one category, i.e., $\varrho = \mathbf{P}(C(Y_i) = C(Y_j) | G_i = 1, G_j = 0)$.

Effect size decomposition

Note that, the MW probability (MWP) is

$$\begin{aligned}\text{MWP} &= \mathbf{P}(Y_i > Y_j | G_i = 1, G_j = 0) + \frac{1}{2}\mathbf{P}(Y_i = Y_j | G_i = 1, G_j = 0) \\ &= (1 - \varrho) \cdot \psi_0 + \varrho \cdot (1 - \omega) \cdot \psi + \varrho \cdot \frac{1}{2} \cdot \omega,\end{aligned}$$

where ψ is defined as

$$\psi = \mathbf{P}(Y_i > Y_j | Y_i \neq Y_j, C(Y_i) = C(Y_j), G_i = 1, G_j = 0),$$

ψ_0 is defined as

$$\psi_0 = \mathbf{P}(C(Y_i) > C(Y_j) | C(Y_i) \neq C(Y_j), G_i = 1, G_j = 0).$$

and ω is defined as

$$\omega = \mathbf{P}(Y_i = Y_j | C(Y_i) = C(Y_j), G_i = 1, G_j = 0).$$

Assuming the percentages in Table 3-1, and that $\psi = \psi_0$, we have $\varrho \approx 0.319$ and $\psi = \psi_0 \approx 0.604$, which leads to a MWP of

$$\text{MWP} \approx 0.604 \times (1 - 0.319) + 0.319 \times 0.604 \times (1 - \omega) + 0.319 \times \frac{1}{2} \times \omega.$$

Power and sample size under the contiguous alternative

For a two-sided test of $\alpha = 0.05$, under the contiguous alternative, the power for a sample size of n is approximately

$$\Phi\left(\frac{1}{2}\sqrt{\frac{n}{\Omega}}\left(\text{MWP}-\frac{1}{2}\right)-\Phi^{-1}\left(1-\frac{\alpha}{2}\right)\right),$$

and the sample size for a power of $1-\beta=0.8$, is approximately

$$4\cdot\Omega\cdot\left(\frac{\Phi^{-1}(1-\alpha/2)+\Phi^{-1}(1-\beta)}{\text{MWP}-0.5}\right)^2,$$

where Ω is the variance parameter defined as

$$\Omega=\mathbf{Var}\left(\mathbf{P}\left(\text{Y}_i<\text{Y}_j\middle|\text{Y}_j\right)+\frac{1}{2}\mathbf{P}\left(\text{Y}_i=\text{Y}_j\middle|\text{Y}_j\right)\right)=\frac{1}{12}\left[1-\sum_x\left(\mathbf{P}(\text{Y}_i=x)\right)^3\right],$$

in which, the sum in the last term is over the mass point.

Note that the variance parameter, Ω, is bounded above by

$$\Omega=\frac{1}{12}\left[1-\sum_x\left(\mathbf{P}(\text{Y}_i=x)\right)^3\right]\leq\frac{1}{12}.$$

Thus, the power for a sample size of n has a lower bound of approximately

$$\Phi\left(\sqrt{3n}(\text{MWP}-0.5)-\Phi^{-1}(1-\alpha/2)\right),$$

and the sample size for a power of $1-\beta=0.8$ has an upper bound of approximately

$$\frac{1}{3}\cdot\left(\frac{\Phi^{-1}(1-\alpha/2)+\Phi^{-1}(1-\beta)}{\text{MWP}-0.5}\right)^2.$$

Blinded sample size re-estimation

The probability, ω, is non-identifiable in the blinded analysis and is bounded above by

$$\lambda=\mathbf{P}\left(\text{Y}_i=\text{Y}_j\middle|\text{C}(\text{Y}_i)=\text{C}(\text{Y}_j)\right),$$

which is the pooled probability of observing ties within any response category, regardless of the treatment group. Hence, under the previous assumptions, the MW probability is bounded below by

$$\text{MW probability}\geq\Psi=(1-\varrho)\cdot\psi_0+\varrho\cdot(1-\lambda)\cdot\psi+\varrho\cdot\frac{1}{2}\cdot\lambda\approx0.604-0.033\times\lambda,$$

the power for a sample size of n is approximately bounded below by

$$\Phi\left(\frac{1}{2}\cdot\sqrt{\frac{n}{\Omega}}\cdot\left(\Psi-\frac{1}{2}\right)-\Phi^{-1}\left(1-\frac{\alpha}{2}\right)\right),$$

and the sample size for a power of $1-\beta=0.8$, is approximately bounded above by

$$4\cdot\Omega\cdot\left(\frac{\Phi^{-1}(1-\alpha/2)+\Phi^{-1}(1-\beta)}{\Psi-0.5}\right)^2.$$

Suppose there are M patients available for the blinded sample size re-estimation.

At the blinded interim analysis for sample size re-estimation, the within-category probability of observing ties, λ, will be estimated by

$$\lambda_{\text{M}}=\frac{\sum_{i=1}^{\text{M}}\sum_{j=1}^{\text{M}}\text{I}\{\text{Y}_i=\text{Y}_j,j\neq i\}}{\sum_{i=1}^{\text{M}}\sum_{j=1}^{\text{M}}\text{I}\{\text{C}(\text{Y}_i)=\text{C}(\text{Y}_j),j\neq i\}}$$

the lower bound for the MW probability, Ψ, will be estimated by

$$\Psi_{\mathrm{M}} = 0.604 - 0.033 \times \lambda_{\mathrm{M}},$$

and the variance parameter, Ω, will be estimated by

$$\Omega_{\mathrm{M}} = \frac{1}{\mathrm{M}-1}\sum_{i=1}^{\mathrm{M}}\left(\frac{1}{\mathrm{M}}\sum_{j=1}^{\mathrm{M}}\frac{1}{2}\left(\mathrm{I}\{\mathrm{Y}_j < \mathrm{Y}_i\} + \mathrm{I}\{\mathrm{Y}_j \le \mathrm{Y}_i\}\right) - \frac{1}{2}\right)^2,$$

the power for a sample size of $n = 360$ is approximately bounded below by

$$\Phi\left(\frac{1}{2}\cdot\sqrt{\frac{360}{\Omega_{\mathrm{M}}}}\cdot\left(\Psi_{\mathrm{M}} - \frac{1}{2}\right) - \Phi^{-1}(0.975)\right),$$

and the sample size for a power of $1 - \beta = 0.8$, is approximately bounded above by

$$4\cdot\Omega_{\mathrm{M}}\cdot\left(\frac{\Phi^{-1}(0.975) + \Phi^{-1}(0.8)}{\Psi_{\mathrm{M}} - 0.5}\right)^2.$$

5.3 Rule of exclusion criteria of analysis sets

Criteria defining protocol deviations are referenced in the "Protocol Deviations' tab of the Data Review Plan document. Protocol deviations will be classified into 4 categories as appropriate:

- Selection criteria not met
- Treatment deviation
- Prohibited concomitant medication
- Other

The protocol deviation criteria are listed in the sheet "Protocol Deviations" in the data review plan. The protocol deviation criteria leading to exclusion from the analysis sets are provided below.

Table 5-1 Protocol deviations leading to exclusion of patients from from analysis sets for Part 1

ID	Description of Protocol Deviation	Excluding from analysis set				
		SCR1	ELG1	PK1	PD1	SAF1
INCL01	Informed consent missing	x	x	x	x	x

Table 5-2 Non-protocol deviation criteria leading to exclusion of patients from analysis sets for Part 1

Criterion	Excluding from analysis set				
	SCR1	ELG1	PK1	PD1	SAF1
Screen failure		x	x	x	x
No available valid (i.e., not flagged for exclusion) PK concentration measurement			x		
No available PD measurements				x	
Patient not receiving any dose of open-label study drug			x	x	x

Table 5-3 Protocol deviation criteria leading to exclusion from the analysis sets for Part 2

Protocol Deviation		Excluding from Analysis Set				
ID	**Description**	**SCR2**	**RAN**	**FAS**	**PPS**	**SAF2**
INCL01	Informed consent missing	X	X	X	X	X
INCL04	At Screening NYHA/ROSS Class 1 then NYHA/ROSS Class 2 to 4 in MH				X	
INCL05a	Systemic left ventricular EF <=45% or fractional shortening <=22.5% (assessed by echocardiogram, MRI, MUGA or left ventricular angiogram)				X	
INCL05b	Systemic left ventricular EF <=45% or fractional shortening <=22.5% (assessed by echocardiogram, MRI, MUGA or left ventricular angiogram) within 1 month before patient begins Part 2				X	
EXCL01	Patients with single ventricle or systemic right ventricle				X	
WITH01	Blind was broken				X	
WITH02	Withdrawal of informed consent and patient continued in the trial				X	
OTH13	Blind was broken in error.				X	
TRT01	Overall drug Compliance rate < 80%				X	
TRT03	Incorrect treatment taken during double-blind epoch				X	
TRT04	Patients that were randomized in error and did not ingest study medication during the Randomized treatment.			X	X	X

Table 5-4 Patient classification for Part 2

Analysis Set	Protocol Deviation ID leading to exclusion of patients	Non-protocol deviation criteria leading to exclusion from analysis sets
SCR2	INCL01	Not applicable
RAN	INCL01	• Patient without a randomization number
FAS	INCL01, TRT04	• Patient without a randomization number • GCP issue present at site
PPS	INCL01, INCL04, EXCL01 INCL05a, INCL05b WITH01, WITH02, OTH13, TRT01, TRT03, TRT04	• Patient without a randomization number • Patient randomized but not receiving any study drug during the double-blind epoch • GCP issue present at site
SAF2	INCL01, TRT04	• Patient without a randomization number • Patient randomized but not receiving any study drug during the double-blind epoch

6 Reference

1. Kawaguchi, A, Koch, GG, Wang, X (2011) Stratified multivariate Mann–Whitney estimators for the comparison of two treatments with randomization based covariance adjustment. Statistics in Biopharmaceutical Research; 3:217-231.
2. Shaddy RE, Boucek MM, Hsu DT, et al. (2007) Carvedilol for children and adolescents with heart failure: A randomized controlled trial. JAMA; 298:1171-1179.
3. Sun H, Davison BA, Cotter G, et al. (2012) Evaluating treatment efficacy by multiple end points in phase II acute heart failure clinical trials: Analyzing data using a global method. Circulation Heart Failure; 5:742-9.
4. Lin DY, Wei LJ, Yang I, et al. (2000) Semiparametric regression for the mean and rate functions of recurrent events. J R Stat Soc B; 62(Pt 4):711-30.
5. Heilbron DC, Holliday MA, al-Dahwi A, Kogan BA. Expressing glomerular filtration rate in children. Erratum in: Pediatr Nephrol 1991 May;5(3):370.