NOVARTIS

Clinical Development

LCZ696/Entresto®

CLCZ696D2302

A 24-week, randomized, double-blind, multi-center, parallel group, active controlled study to evaluate the effect of LCZ696 on NT-proBNP, exercise capacity, symptoms and safety compared to individualized medical management of comorbidities in patients with heart failure and preserved ejection fraction

Statistical Analysis Plan (SAP)

Author: Trial Statistician, [redacted];

Document type: SAP Documentation

Document status:

Release date:

Number of pages: 57

Document History – Changes compared to previous final version of SAP

Date	Time point	Updates	Section impacted (Current)
25JUN2017	Prior to DBL		
31OCT2018	Prior to DBL	Updated per protocol amendment 2	All sections
		Add back-up strategy for model non-convergence	Section 2.1 Section 2.2.1
		Add additional sub-group of interest	Section 2.5.5
29OCT2019	Prior to DBL	Add additional sub-group analyses	Section 2.6.5
		Update safety topics	Section 2.8.1
		Update laboratory analyses	Section 2.8.3
		Update per protocol set definition	Section 5.2

Table of contents

List of abbreviations

6MWD	six-minute walking distance
6MWT	six-minute walking test
ACEi	angiotensin converting enzyme inhibitor
AE	adverse event
AF	atrial fibrillation/atrial flutter
Alb	albumin
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ARB	angiotensin receptor blocker
AST	aspartate aminotransferase
ATC	anatomical therapeutic classification
bid	twice a day
BMI	body mass index
bpm	beats per minute
B6MWD	baseline six-minute walking distance
BCSS	baseline KCCQ CSS
BLNTBNP	baseline log(NT-proBNP)
BMSS	baseline SF-36 MCS score
BOSS	baseline KCCQ OSS
BPSS	baseline SF-36 PCS score
BSBP	baseline systolic blood pressure
BUN	blood urea nitrogen
CAD	coronary artery disease
CRF	case report form
CSR	clinical study report
CSS	clinical summary score
CV	coefficient of variation
DBP	diastolic blood pressure
DMC	data monitoring committee
FAS	full analysis set
ECG	electrocardiogram
echo	echocardiogram
eGFR	estimated glomerular filtration rate
EOS	end of study
GCP	good clinical practice
HFpEF	heart failure with preserved ejection fraction
HTN	Hypertension
IA	interim analysis
IMT	individualized medical therapy
INR	international normalized ratio
IRT	interactive response technology
KCCQ	Kansas City Cardiomyopathy Questionnaire

LDL	low-density lipoprotein
LLOQ	lowest limit of quantitation
LVEF	left ventricular ejection fraction
MAR	missing at random
MCS	mental component summary
MedDRA	Medical Dictionary for Drug Regulatory Affairs
MCV	mean corpuscular volume
MCHC	mean corpuscular hemoglobin concentration
MMRM	mixed model for repeated measures
MNAR	missing not at random
NT-proBNP	N-terminal pro-brain natriuretic peptide
OSS	overall summary score
PPS	per-protocol set
PCS	physical component summary
PRO	patient-reported outcomes
PT	preferred term
QoL	quality of life
RASi	renin angiotensin system inhibitors
RBC	red blood cell count
RDW	red blood cell distribution width
SAE	serious adverse event
SAP	statistical analysis plan
SF-36	short form health survey
SOC	system organ class
TBL	total bilirubin
TEAE	treatment emergent adverse event
TESAE	treatment emergent serious adverse event
TFLs	tables, figures, listings
ULN	upper limit of normal range
ULOQ	upper limit of quantitation
WBC	white blood cell count
WHO	World Health Organization

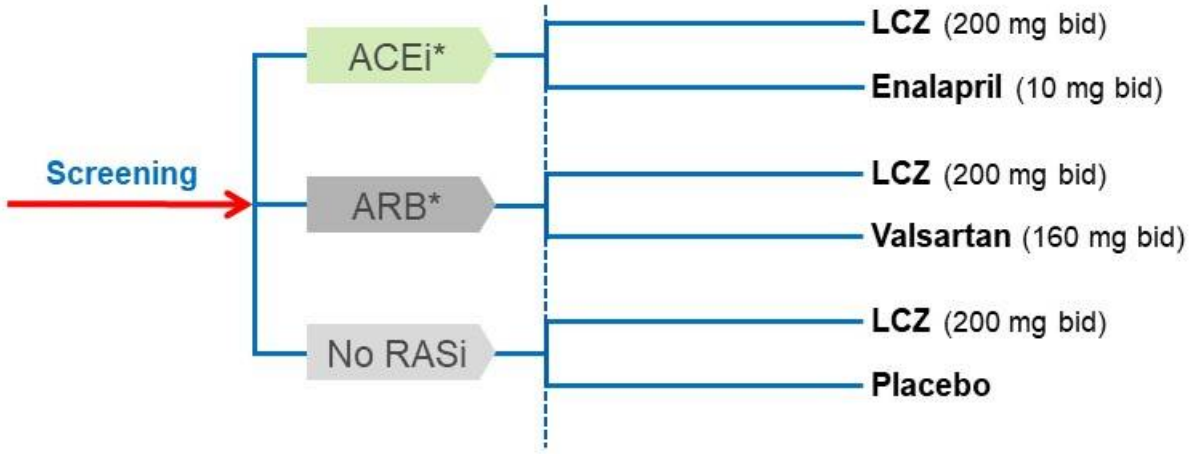
1 Introduction

The statistical analysis plan (SAP) describes the detailed methodology and implementation of the planned statistical analyses outlined in the study protocol v02 for CLCZ696D2302. The analysis results will be summarized in the clinical study report (CSR).

1.1 Study design

This is a 24-week, randomized, double-blind, multi-center, parallel group, active controlled study to evaluate LCZ696 compared to individualized medical therapy (IMT) on N-terminal pro-brain natriuretic peptide (NT-proBNP), exercise capacity, symptoms and quality of life (QoL) in patients with heart failure with preserved left ventricular ejection fraction (HFpEF) (left ventricular ejection fraction [LVEF] > 40%) (Figure 1-1).

Figure 1-1 Study design



Screening	Stratified Randomization	Double-Blind Treatment	
Up to 2 weeks	Visit 101	24 weeks	
		Titration	*Maintainance*
		1-4 weeks	20-23 weeks

Approximately 2500 patients will be randomized in this study.

Patients will be initially stratified into one of three strata according to prior renin-angiotensin system inhibitor (RASi) treatment for comorbidities: angiotensin converting enzyme inhibitor (ACEi), angiotensin receptor blocker (ARB) or no RASi. Patients will then be randomized at a 1:1 ratio to receive either LCZ696 or comparator (i.e., IMT). Patients in the ACEi stratum will be randomized to receive either LCZ696 or enalapril. Patients in the ARB stratum will be randomized to receive either LCZ696 or valsartan. Patients in the no RASi stratum will be randomized to receive either LCZ696 or placebo. Patients in the ACEi and ARB strata must have a history of hypertension (HTN). There will be no designated proportion of patients in

each stratum, the strata will populate based upon the patient's prior treatment regimen. In addition, randomization will be stratified by region.

For the purpose of blinding, the corresponding placebos are also dispended to the patients. Table 1-1 summarizes the study medications for each treatment group and each stratum (ACEi, ARB, No RASi).

Table 1-1 Study Medications

Stratum	Study Medications	
	LCZ696	IMT
ACEi	LCZ696 Active + Enalapril Placebo	LCZ696 Placebo + Enalapril Active
ARB	LCZ696 Active + Valsartan Placebo	LCZ696 Placebo + Valsartan Active
No RASi	LCZ696 Active	LCZ696 Placebo

No interim analysis (IA) for efficacy is planned.

1.2 Study objectives and endpoints

Table 1-2 presents the objectives and related endpoints

Table 1-2 Objectives and related endpoints

Objective(s)	Endpoint(s)
Primary Objective(s) • To demonstrate that LCZ696 is superior to IMT for comorbidities in reducing NT-proBNP from baseline after 12 weeks of treatment • To demonstrate that LCZ696 is superior to IMT for comorbidities in improving exercise capacity as assessed by the six-minute walk test (6MWT) at Week 24 in a subset of patients	**Endpoint(s) for primary objective(s)** • Change from baseline in NT-proBNP (in log scale) at Week 12 • Change from baseline in six-minute walk distance (6MWD) at Week 24
Secondary Objective(s) • To compare LCZ696 to IMT for comorbidities on mean change of Kansas City Cardiomyopathy Questionnaire (KCCQ) clinical summary score (CSS) at Week 24 • To compare LCZ696 to IMT for comorbidities on proportion of patients with ≥ 5-points change in KCCQ CSS at Week 24 (separate analyses for ≥ 5-points improvement and ≥ 5-points deterioration) • To compare LCZ696 to IMT for comorbidities in improving NYHA functional class at Week 24 • To compare LCZ696 to IMT for comorbidities in improving symptoms as assessed by the Short Form Health Survey (SF-36) physical component summary (PCS) score at Week 24	**Endpoint(s) for secondary objective(s)** • Mean change from baseline in KCCQ CSS at Week 24 • Proportion of patients with ≥ 5-points deterioration in KCCQ CSS at Week 24 • Proportion of patients with ≥ 5-points improvement in KCCQ CSS at Week 24 • Change from baseline in NYHA functional class at Week 24 • Change from baseline in SF-36 PCS score at Week 24

Objective(s)	Endpoint(s)
Exploratory Objective(s) • To compare LCZ696 to IMT for comorbidities in reducing NT-proBNP at Week 24 • To explore the relative effect of LCZ696 vs IMT for comorbidities on renal function as assessed by estimated glomerular filtration rate (eGFR) at Week 24 • To compare LCZ696 to IMT for comorbidities on mean change of KCCQ overall summary score (OSS) at Week 24 • To compare LCZ696 to IMT for comorbidities on proportion of patients with ≥ 5 points change in KCCQ OSS at Week 24 (separate analyses for ≥ 5-points improvement and ≥ 5-points deterioration) • To evaluate safety of LCZ696 vs IMT for comorbidities	**Endpoint(s) for exploratory objective(s)** • Change from baseline in NT-proBNP (in log scale) at Week 24 • Rate of change (slope) in eGFR from baseline • Mean change from baseline in KCCQ OSS at Week 24 • Proportion of patients with ≥ 5-points deterioration in KCCQ OSS at Week 24 • Proportion of patients with ≥ 5-points improvement in KCCQ OSS at Week 24 • Frequency of adverse events (AEs), serious adverse events (SAEs), and laboratory abnormalities

2 Statistical methods

The following section contain important information on detailed statistical methodology used for analysis and reporting purposes.

2.1 Data analysis general information

Unless otherwise specified, data will be analyzed by DATAMAP according to the protocol for CLCZ696D2302, using SAS 9.4 or higher. Further details on planned statistical analyses will be presented in the following section and in CSR Appendix 16.1.9.

An established program level data monitoring committee (DMC) independent of Novartis that reviews safety data from LCZ696 studies will be used for this study. The DMC will review patient safety data in an unblinded manner approximately every six months and determine if it is safe to continue the study. Any major recommendation from the DMC will be communicated to the Executive Committee and must be reviewed and ratified by the Executive Committee in consultation with Novartis prior to its enactment.

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

The analysis will be conducted on all data at the end of trial.

2.1.1 General definitions

Randomized Treatment Epoch

The randomized treatment 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 defiend as below.

- For efficacy variables, if the scheduled assessment at randomization (nominal visit) is performed and the assessment value is non-missing, the baseline assessment is defined as the scheduled assessment at randomization (nominal visit).
- For efficacy variables, if the scheduled assessment at randomization (nominal visit) 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 randomization date.
- For safety variables, the baseline assessment is defined as the last non-missing assessment (scheduled or unscheduled) prior to or on the first dose date of the study drug.

Specifically, the baseline systolic blood pressure (BSBP) is included in some efficacy analysis model as a covariate, in this case, the baseline assessment for the systolic blood pressure will be defined using the above criteria for efficacy variables.

In the listings, the study day will be displayed, as applicable.

- In the listings of efficacy data, the study day is defined as the date of assessment (event/visit) minus the date of randomization plus one day, i.e., Day 1 is the randomization date.
- In the listings of safety data, the study day is defined as the date of assessment (event/visit) minus the date of first dose plus one day, i.e., Day 1 is the first dose date.

Post-baseline assessment

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

Region and stratum (ACEi, ARB, No RASi)

Unless otherwise specified, the region and the stratum (ACEi, ARB, No RASi) will be loaded from the randomization data, where the region is defined based on country according to Table 2-1, which may be updated according to the final enrollment plan prior to the database lock, and the stratum is defined based on the usage of the RASi (ACEi/ARB) medication prior to the randomization.

Table 2-1 Region mapping criteria

Region	Countries
North America	Canada, United States
Latin America (including Central America)	Argentina, Brazil, Colombia, Guatemala, Mexico, Peru
Europe	Austria, Belgium, Bulgaria, Czech Republic, Denmark, Estonia, France, Germany, Hungary, Italy, Latvia, Lithuania, Netherlands, Portugal, Romania, Russia, Serbia, Slovakia (Slovak Republic), Spain, Turkey, United Kingdom

Region	Countries
Asia/Pacific and other	India, Israel, Thailand

Visit 199 mapping

For patients who completed the randomized treatment epoch, Visit 199 will be Week 24.

For patients who discontinued from the study during the randomized treatment epoch, Visit 199 will be mapped to the next scheduled visit of the last attended scheduled visit prior to visit 199 (Table 2-2).

Table 2-2 Visit 199 mapping crietria

Randomized Treatment Epoch Patient Status	Last Attended Scheduled Visit Prior to Visit 199		Mapped Visit for Visit 199	
	Visit Number	Visit Name	Visit Number	Visit Name
Discontinued	101	Randomization	102	Week 2
Discontinued	102	Week 2	103	Week 4
Discontinued	103	Week 4	104	Week 8
Discontinued	104	Week 8	105	Week 12
Discontinued	105	Week 12	106	Week 16
Discontinued	106	Week 16	107	Week 20
Discontinued	107	Week 20	108	Week 24
Completed	---	---	108	Week 24

- For each patient, the last attended scheduled visit prior to visit 199 is defined as the visit with maximal visit number among all scheduled visit (visit 101 to visit 107) in the visit panel.

Data collected at Visit 199 will be included in by-visit summaries and longitudinal modelling, if an assessment were scheduled at the mapped visit for Visit 199.

Unscheduled assessments

Unscheduled assessments will be included in over period minimum/maximum, and in over period shift table, but will not be included in by-visit summary and longitudinal modelling.

Model non-convergence

For all regression models, if the model fitting procedure does not converge, covariates could be removed from the model sequentially, as applicable, with the following order: visit-by-baseline interaction, stratum-by-BSBP interaction, stratum-by-baseline interaction, BSBP, baseline, region and stratum (ACEi, ARB, No RASi).

2.2 Analysis sets

The following analysis sets will be used for the statistical analyses.

- **Screened set (SCR)** – All patients who signed the informed consent for CLCZ696D2302. The SCR 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 in the SCR 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, and the exception of those patients who have been inadvertently randomized within a wrong stratum (ACEi, ARB, No RASi) and have not received study drug. Further exclusions from the FAS may only be justified in exceptional circumstances (e.g., *serious* GCP violations). Following the intent-to-treat principle, patients will be analyzed according to the treatment to which they were 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 will consist of the patients who do not have major deviations from the protocol procedures in the randomized treatment epoch. 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.
- **Safety set (SAF)** – All randomized patients who received at least one dose of study drug. Patients will be analyzed according to the treatment actually received. The SAF will be used for the analyses of safety variables. 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.

Patients without valid written informed consent will be excluded from all analysis sets. Further exclusions from the FAS are only justified in exceptional circumstances (e.g., site closed down for GCP reasons). The determination of which patients were excluded from the FAS is made in a blinded manner before the database lock.

Rules leading to exclusion from the analysis sets are given in Appendix 5.2.

2.2.1 Subgroup of interest

Table 2-3 specifies the subgroups to be used in sub-group analyses.

Table 2-3 Subgroup specifications

Subgroups	Source		
	Data Source	Variable	Visit/Time Point
Region • North America; • Latin America (including Central America) • Europe; • Asia/Pacific and other	Randomization	Region	Randomization
Stratum (ACEi, ARB, No RASi) • ACEi; • ARB; • No RASi;	Randomization	Stratum	Randomization

Subgroups	Source		
	Data Source	Variable	Visit/Time Point
Age group (< 65 years, ≥ 65 years) • < 65 years; • ≥ 65 years;	CRF page: Demography	Age	Screening
Age group (< 75 years, ≥ 75 years) • < 75 years; • ≥ 75 years;	CRF page: Demography	Age	Screening
Sex • Female; • Male;	CRF page: Demography	Sex	Screening
Baseline NYHA class group • Class I/II • Class III/IV	CRF page: Signs and Symptoms of Heart Failure including New York Heart Association Classification	NYHA Classification	Randomization (Baseline)
Diabetes (history or HbA1c ≥ 6.5%)* • Yes • No	CRF page: Heart failure and Diabetes History Central laboratory result	Diabetes mellitus history HbA1c	Screening
LVEF group (<= 60%, > 60%) • <= 60 % • > 60 %	CRF page: Heart failure and Diabetes History	LVEF	Screening
Prior heart failure hospitalization • Yes • No	CRF page: Heart failure and Diabetes History	Prior heart failure hospitalization	Screening
NOTE: Baseline NYHA class is defined using the baseline criteria for efficacy variables (Section 2.1.1) NOTE *: Diabetes (history or HbA1c ≥ 6.5%) is defined as No if the patient has no history of diabetes mellitus and HbA1c at screening < 6.5%; Diabetes (history or HbA1c ≥ 6.5%) is defined as Yes if the patient has history of diabetes mellitus or HbA1c at screening ≥ 6.5%.			

Table 2-4 summarizes the subgroup analyses in this study.

Table 2-4 Subgroup analyses

Subgroup Variable	Primary/Secondary Efficacy Analysis	Selected Safety Analysis
Region	X	X
Stratum (ACEi, ARB, No RASi)	X	X
Age group (< 65 years, ≥ 65 years)	X	X

Subgroup Variable	Primary/Secondary Efficacy Analysis	Selected Safety Analysis
Age group (< 75 years, ≥ 75 years)	X	X
Sex	X	X
Baseline NYHA class group	X	X
Diabetes (history or HbA1c ≥ 6.5%)	X	X
LVEF group (<= 60%, > 60%)	X	X
Prior heart failure hospitalization	X	

2.3 Patient disposition, demographics and other baseline characteristics

2.3.1 Patient disposition

Disposition for the screening epoch

The number and percentage of patients who completed the screening epoch will be provided. Besides, the primary reason for not completing the screening epoch will be summarized, using the number and percentage of patients not qualifying for such reasons. For any patient who was screened more than once, the data from the last screening will be used in the summary.

The SCR will be used for the above analyses.

Disposition for the randomized treatment epoch

The patient status for the randomized treatment epoch will be summarized by treatment group for overall and for each stratum (ACEi, ARB, No RASi), using the number of patients who were randomized, the numbers and percentages of patients who completed the epoch and who discontinued from the study during the epoch; the primary reason for discontinuation from the study during the randomized treatment epoch will be summarized by treatment group for overall and for each stratum (ACEi, ARB, No RASi), using the number and percentage of patients discontinued for such reasons.

The RAN will be used for the above analyses.

Disposition for the study treatment

The numbers and percentages of patients who took at least one dose of the study drug, who completed the study treatment and who permanently discontinued from the study treatment will be provided by treatment group, for overall and for each stratum (ACEi, ARB, No RASi); the primary reason for permanent discontinuation from the study drug will be summarized by treatment group, for overall and for each stratum (ACEi, ARB, No RASi), using the number and percentage of patients discontinued for such reasons.

The RAN will be used for the above analyses.

Protocol deviations

The protocol deviations will be summarized by treatment group and protocol deviation category, for overall and for each stratum (ACEi, ARB, No RASi), using the numbers and percentages of patients with at least one protocol deviation within the protocol deviation category.

The RAN will be used for the above analyses.

Analysis set dispositions

The numbers and percentages of patients within each analysis set will be provided by treatment group, for overall and for each stratum (ACEi, ARB, No RASi).

In addition, the number and percentage of patients satisfying each criteria leading to exclusion from analysis sets will be provided by treatment group, for overall and for each stratum (ACEi, ARB, No RASi).

2.3.2 Background and demographic characteristics

For overall and for each stratum (ACEi, ARB, No RASi), summary statistics will be provided by treatment group for demographics and baseline characteristics, including age, age group (< 65 years, ≥ 65 years), age group (< 75 years, ≥ 75 years), sex, race, ethnicity, category of prior cardiovascular medication, prior heart failure hospitalization, height, weight, body mass index (BMI), body surface area (BSA), waist circumference, hip circumference, sitting pulse, sitting systolic blood pressure (SBP), sitting diastolic blood pressure (DBP), LVEF, atrial fibrillation, atrial flutter, atrial fibrillation/atrial flutter (AF), NT-proBNP, 6MWD, HbA1c, HbA1c group (< 6.5%, ≥ 6.5%), history of diabetes mellitus, diabetes (history or HbA1c ≥ 6.5%), history of hypertension, history of coronary heart disease, NYHA class, KCCQ CSS, region, and stratum (only for overall), where BMI will be calculated as weight (kg) / height^2 (m^2) from height and weight at screening. If an above variable is scheduled at screening, the scheduled assessment at screening (nominal visit) will be summarized for this variable. If an above variable is scheduled at randomization, the baseline assessment, defined using baseline criteria for efficacy variables (Section 2.1.1), will be summarized for this variable. If an above variable is scheduled at both screening and randomization, both the assessment at screening (nominal visit) and the baseline assessment, will be summarized for this variable.

Continuous variables will be summarized using number of observations, mean, standard deviation, median, minimum, the first quartile (Q1), the third quartile (Q3), and maximum.

Categorical variables will be summarized using frequencies and percentages.

The FAS will be used for the above 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 screening. The number and percentage of patients with each medical condition will be provided by treatment group, system organ class (SOC), and preferred term (PT), for overall and for each stratum (ACEi, ARB, No RASi).

The FAS will be used for the above analyses.

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

2.4.1 Study treatment / compliance

The study drug administration will be recorded on the CRF page: "Dosage Administration Record – Double-blind Medication" using start date, end date 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 its 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, data conventional imputation algorithm will be applied to the start date and end date. Detailed algorithm will be provided in the study programming datasets specifications (PDS).

For patients who permanently discontinue from the 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 each dosing interval, the dispensing level of the study drug will be recorded together with the start date and end date. The dose level and the daily dose (mg/day) are defined based on the dispensing level according to Table 2-5.

Table 2-5 Dose levels and daily doses of the study drug

Dispensing Level	Dose Level	Dose			Daily Dose		
		LCZ696	Valsartan	Enalapril	LCZ696	Valsartan	Enalapril
No treatment	0	No Dose	No Dose	No Dose	0 mg/day	0 mg/day	0 mg/day
Dose level 1	1	50 mg bid.	40 mg bid.	2.5 mg bid.	100 mg/day	80 mg/day	5 mg/day
Dose level 2	2	100 mg bid.	80 mg bid.	5.0 mg bid.	200 mg/day	160 mg/day	10 mg/day
Dose level 3	3	200 mg bid.	160 mg bid.	10.0 mg bid.	400 mg/day	320 mg/day	20 mg/day

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 daily dose (mg/day) at the visit are defined as the dispensing level, dose level and daily dose (mg/day) 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 will be summarized by treatment group and visit, for overall and for each stratum (ACEi, ARB, No RASi), 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 stratum

(ACEi, ARB, No RASi), using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum; the daily dose (mg/day) will be summarized by treatment group and visit, for each stratum (ACEi, ARB, No RASi), but not for overall, using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum. The summary will include all scheduled visits in the randomized treatment epoch.

The SAF will be used for the above analyses.

Duration of treatment

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

The duration (week) of treatment will be converted from the duration (day) of treatment using a factor of 7 day/week, and will be summarized by treatment group, for overall and for each stratum (ACEi, ARB, No RASi), using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum.

The duration (week) of treatment will be categorized into the following categories.

- < 2 weeks
- 2 weeks to < 8 weeks
- 8 weeks to < 16 weeks
- ≥ 16 weeks

The categorized duration of treatment will be summarized by treatment group, for overall and for each stratum (ACEi, ARB, No RASi), using the number and percentage of patients within each category.

Besides, the overall patient-weeks on-treatment will be provided by treatment group, for overall and for each stratum (ACEi, ARB, No RASi), where the overall patient-weeks on-treatment is defined as the sum of duration (week) of treatment among all patients within the corresponding treatment group and stratum.

The SAF 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 (week) on each dose level will be converted from the duration (day) on this dose level using a factor of 7 day/week, and will be summarized by treatment group, for overall and for each stratum (ACEi, ARB, No RASi).

The SAF will be used for the above analyses.

Duration of study drug exposure

The duration (day) of study drug exposure is defined as the sum of the duration (day) on dose level 1 to 3.

The duration (week) of study drug exposure will be converted from the duration (day) of study drug exposure using a factor of 7 day/week, and will be summarized by treatment group, for overall and for each stratum (ACEi, ARB, No RASi), using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum.

The duration (week) of study drug exposure will be categorized into the following categories.

- < 2 weeks
- 2 weeks to < 8 weeks
- 8 weeks to < 16 weeks
- ≥ 16 weeks

The categorized duration of study drug exposure will be summarized by treatment group, for overall and for each stratum (ACEi, ARB, No RASi), using the number and percentage of patients within each category.

The SAF will be used for the above analyses.

Percentage of study drug exposure

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

The percentage of study drug exposure will be summarized by treatment group, for overall and for each stratum (ACEi, ARB, No RASi), using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum.

The SAF will be used for the above analyses.

Study treatment permanent discontinuation at each visit

For a given visit, a patient is considerred as permanently discontinued from the study treatment at this visit, if the patient permanently discontinued from the study treatment prior to or on the date of this visit.

For overall and for each stratum (ACEi, ARB, No RASi), the number and percentage of patients who permanently discontinued from the study treatment will be provided for each scheduled visit in the randomized treatment epoch.

Mean daily dose and mean dose level

For each patient, the mean dose level is defined as the weighted mean of the dose level among dose level 0 to 3 using the duration (day) on each dose level as the weight, i.e.,

$$\frac{\sum_{k=0}^{3} k \times \text{Duration (day) on Dose Level } k}{\sum_{k=0}^{3} \text{Duration (day) on Dose Level } k}.$$

The mean dose level will be summarized by treatment group, for overall and for each stratum (ACEi, ARB, No RASi), using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum.

For each patient and each dosing interval, the daily dose is defined based on the dose level and the active study medication for the actual arm according to Table 1-1 and Table 2-5. If there is no active study medication for the actual arm (i.e., the IMT arm in the No RASi stratum), the daily dose is defined as 0 mg/day.

For each patient, the mean daily dose (mg/day) is defined as the weighted mean of the daily dose (mg/day) among dose level 0 to 3 using the duration (day) on each dose level as the weight, i.e.

$$\frac{\sum_{k=0}^{3} \text{Daily Dose (mg/day) for Dose Level } k \times \text{Duration (day) on Dose Level } k}{\sum_{k=0}^{3} \text{Duration (day) on Dose Level } k}.$$

The mean daily dose (mg/day) will be summarized by treatment group, for each stratum (ACEi, ARB, No RASi), but not for overall, using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum.

The SAF will be used for the above analyses.

Duration of study exposure

The duration (day) of study exposure is defined as the date of last visit minus the date of randomization plus one day.

The duration (week) of study exposure will be derived from the duration (day) of study exposure using a factor of 7 day/week, and will be summarized by treatment group, for overall and for each stratum (ACEi, ARB, No RASi), using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum.

The duration (week) of study exposure will be categorized into the following categories.

- < 2 weeks
- 2 weeks to < 8 weeks
- 8 weeks to < 16 weeks
- ≥ 16 weeks

The categorized duration of study exposure will be summarized by treatment group, for overall and for each stratum (ACEi, ARB, No RASi), using the number and percentage of patients within each category.

The SAF will be used for the above analyses.

Last recorded dose level

The last recorded dose level is defined as the dose level for the last dosing interval.

The last recorded dose level will be summarized by treatment group, for overall and for each stratum (ACEi, ARB, No RASi), using the number and percentage of patients on each dose level.

The SAF 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 will be summarized by treatment group, for overall and for each stratum (ACEi, ARB, No RASi), using the number and percentage of patients with at least one down-titration dosing interval.

The SAF will be used for the above analyses.

Time to first down-titration of the study drug

The time (day) to first down-titration is defined as the start date of first down-titration dosing interval minus the date of first dose plus one day. If no down-titration occurred for the patient, the time (day) to first down-titration will be censored on the date of death (if the patient died), or the date of last visit (if the patient did not die) for the patient.

For overall and for each stratum (ACEi, ARB, No RASi), the time (day) to first down-titration, will be summarized by treatment group, using Kaplan-Meier curves.

The SAF will be used for the above analyses.

Dose interruption

A dose-interruption dosing interval is defined as a dosing interval whose dose level is 0.

The number and percentage of patients with at least one dose-interruption dosing interval will be provided by treatment group, for overall and for each stratum (ACEi, ARB, No RASi).

The duration of dose interruption is defined as the sum of (the end date minus the start date plus one day) among all dose-interruption dosing intervals.

The number and percentage of patients whose duration of dose interruption is larger than 14 days will be provided by treatment group, for overall and for each stratum (ACEi, ARB, No RASi).

The SAF will be used for the above analyses.

Compliance

As presented in Table 1-1, in this study, the study medication includes LCZ696 Active/Placebo (for all patients), Enalapril Active/Placebo (only for patients in the ACEi stratum) and Valsartan Active/Placebo (only for patients in the ARB stratum).

For each applicable study medication, the compliance will be recorded on the CRF page: “Study Treatment Compliance per Visit”, whereby, for each post-baseline scheduled visit (Week 2, Week 4, Week 8, Week 12, Week 16, Week 20 and Week 24), the compliance from the closest previous attended scheduled visit to the current visit should be recorded.

Specifically, for patients in the ACEi stratum, the compliance for LCZ696 Active/Placebo and Enalapril Active/Placebo will be recorded; for patients in the ARB stratum, the compliance for LCZ696 Active/Placebo and Valsartan Active/Placebo will be recorded; for patients in the No RASi stratum, the compliance for LCZ696 Active/Placebo will be recorded.

For each patient and each post-baseline scheduled visit, the on treatment duration (day) from the closest previous attended scheduled visit to the current visit is defined as the number of days on any positive dose level during the period from the date of the closest previous attended scheduled visit (inclusive for the first post-baseline scheduled visit, but exclusive for all afterwards) to the date of the current visit (inclusive).

For each patient and each applicable study medication (Table 1-1) , the overall compliance for the study medication is defined as the weighted mean of the compliance using the on treatment duration (day) as the weight, i.e.

$$\frac{\sum_k \text{On Treatment Duration (day) in Period } k \times \text{Compliance}(\%) \text{ in Period } k}{\sum_k \text{On Treatment Duration (day) in Period } k},$$

where period *k* is the period from the date of the closest attended scheduled visit prior to Visit *k* (inclusive for the first post-baseline scheduled visit, but exclusive for all afterwards) to the date of Visit *k* (inclusive), and the sum is over all attended post-baseline scheduled visit.

For each stratum (ACEi, ARB, No RASi) and each applicable study medication (Table 1-1), the overall compliance (%) will be summarized separately by treatment group, using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum.

2.4.2 Prior and concomitant therapies

Prior and concomitant medications will be recorded on the CRF 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 based on the Novartis ADaM Governance Board (AGB) global standards.

Prior and concomitant medications/non-drug therapies will be identified based on recorded or imputed start dates and end dates.

Prior medications are defined as any recorded medication with its start date (recorded or imputed) prior to the date of the first dose of the study drug.

Concomitant medications are defined as any recorded medication with its end date (recorded or imputed) later than or equal to the date of the first dose of the study drug AND its start date (recorded or imputed) prior to or equal to the end date of the study.

Prior non-drug therapies are defined as any procedure/significant non-drug therapy with its start date (recorded or imputed) prior to the date of the first dose of the study drug.

Concomitant non-drug therapies are defined as any recorded procedure/significant non-drug therapy with its end date (recorded or imputed) later than or equal to the date of the first dose of the study drug AND its start date (recorded or imputed) prior to or equal to the end date of the study.

For overall and for each stratum (ACEi, ARB, No RASi), prior and concomitant medications will be summarized separately by treatment group, anatomical therapeutic classification (ATC) and PT; prior and concomitant non-drug therapies will be summarized separately by treatment group, SOC, PT.

In addition, for each of the following classes, prior medications and concomitant medications will be summarized by treatment group, for overall and each stratum (ACEi, ARB, No RASi).

- Open label ACEi
- Open label ARB
- Open label ACEi or ARB
- Open label dual RAS blockade
- Open label sacubitril/valsartan
- Renin inhibitors
- Beta blockers
- Diuretics
  - Aldoesterone antagonists
  - Loop diuretics
  - Thiazide diuretics
  - Other diuretics
- Calcium antagonists
- Nitrates
- Other vasodilators
- Cardiac glycosides (digoxin/digitalis glycoside)
- Antiarrhythmic agents
- Aspirin
- Other antiplatelet agents
- Other antithrombotic agents
  - Vitamin-K antagonist
  - NOAKs
- Statins
- Other lipid lowering agents
- Insulins
- GLP-1 agonists
- Oral anti diabetics
  - SGLT-2 inhibitors

The SAF will be used for the above analyses.

2.4.3 Comorbidity management

The lipid management will be recorded on the CRF page: “Management of Comorbidities – Lipids Management”, and will be summarized by treatment group and visit, for overall and for each stratum (ACEi, ARB, No RASi), using the number and percentage of patients having low-density lipoprotein (LDL) ≥ 100 mg/dL and (coronary artery disease [CAD] or diabetes), the number and percentage of patients with lipid lowering medications initiated/adjusted, and the

number and percentage of patients with lipid lowering medications not initiated/adjusted for each reason.

The diabetes management will be recorded on the CRF page: "Management of Comorbidities – Diabetes Management", and will be summarized by treatment group and visit, for overall and for each stratum (ACEi, ARB, No RASi), using the number and percentage of patients with HbA1c > 7.5%, the number and percentage of patients with diabetic medications initiated/adjusted, and the number and percentage of patients with diabetic medications not initiated/adjusted for each reason.

The atrial fibrillation and atrial flutter management, including rate control and anticoagulation or antiplatelet therapy, will be recorded on the CRF page: "Management of Comorbidities – Atrial Fibrillation and Atrial Flutter", and will be summarized by treatment group and visit, for overall and for each stratum (ACEi, ARB, No RASi), using the number and percentage of patients with heart rate ≥ 100 beats per minute (bpm), the number and percentage of patients with rate control medications initiated/adjusted, the number and percentage of patients receiving anticoagulation or antiplatelet therapy, the number and percentage of patients with rate control medications not initiated/adjusted for each reason, and the number and percentage of patients not receiving anticoagulation or antiplatelet therapy for each reason.

The blood pressure management will be recorded on the CRF page: "Management of Comorbidities – Blood Pressure Management", and will be summarized by treatment group and visit, for overall and for each stratum (ACEi, ARB, No RASi), using the number and percentage of patients with either SBP > 140 mmHg or DBP > 90 mmHg, the number and percentage of patients with blood pressure medications initiated/adjusted, and the number and percentage of patients with blood pressure medications not initiated/adjusted for each reason.

The SAF will be used for the above analyses.

2.5 Analysis of the primary objective

2.5.1 Primary endpoints

The primary efficacy variables are the change from baseline in log(NT-proBNP) at Week 12, and the change from baseline in 6MWD at Week 24.

Plasma NT-proBNP will be collected using central lab.

All post-screening NT-proBNP from central lab will remain blinded until data base lock. Only the NT-proBNP at screening will be reported to the investigator and the Novartis clinical study team.

For the NT-proBNP, if the test value is below the lowest limit of quantitation (LLOQ), the test value will be imputed by 0.5 × LLOQ; if the test value is above the upper limit of quantitation (ULOQ), the test value will be imputed by 1.5 × ULOQ.

For the NT-proBNP, unreliable test values will be set to missing.

The 6MWD will be recorded on the CRF page: "6 Minute Walk Test and Borg Scale".

Specifically, for the 6MWD, the baseline assessment is defined as the scheduled assessment at randomization (nominal visit).

For patients who have died, the 6MWD after the death will be defined as zero.

For patients who permanently discontinue from study treatment, assessment values collected after permanent discontinuation will generally be included in the analysis.

2.5.2 Statistical hypothesis, model, and method of analysis

Since the testing procedure involves both the primary and secondary endpoints, we describe it for them together. The analysis methods for the secondary endpoints are described in Section 2.6.

Testing strategy

The following primary null hypotheses will be included in the testing strategy.

- H_1: LCZ696 is no better than IMT in change from baseline in log(NT-proBNP) at Week 12 in the overall study population
- H_2: LCZ696 is no better than IMT in change from baseline in 6MWD at Week 24 in patients with baseline six-minute walk distance (B6MWD) ranging from 100 meters to 450 meters.

The following secondary null hypotheses will be included in the testing strategy.

- H_3: LCZ696 is no better than IMT in change from baseline in KCCQ CSS at Week 24 in the overall study population.
- H_4: LCZ696 is no better than IMT in NYHA change from baseline at Week 24 in the overall study population.

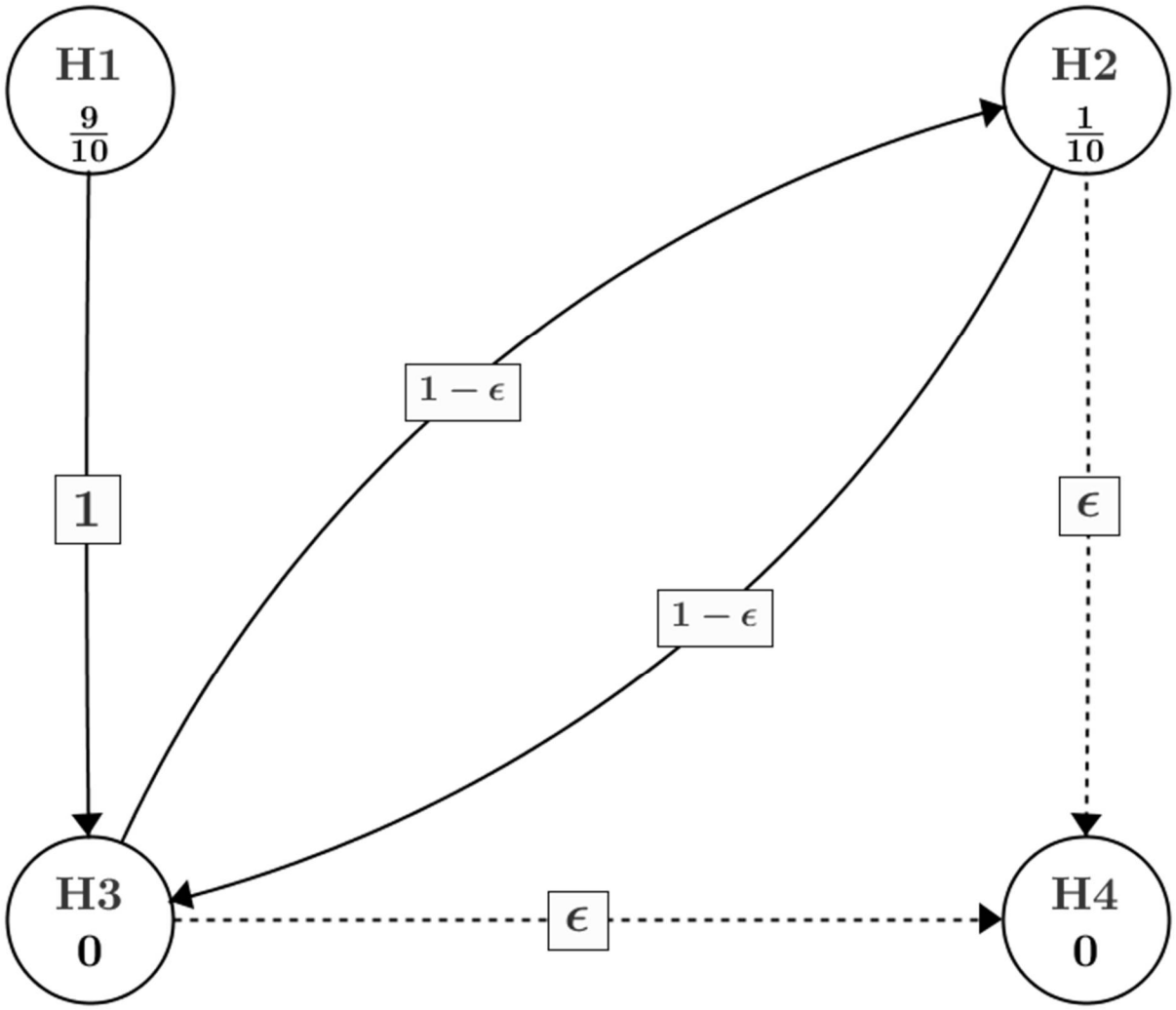
Each null hypothesis is tested against the one-sided alternative that LCZ696 is better than IMT in the corresponding variable.

In order to control the family-wise type-I error rate at the one-sided 0.025 significance level, a sequentially rejective multiple testing procedure based on the graphical presentation (Bretz et al 2009) will be employed, whereby H_1 and H_2 will be tested first at initially assigned level of one-sided $(9/10) \times \alpha = 0.0225$ and one-sided $(1/10) \times \alpha = 0.0025$, accordingly. If H_1 and/or H_2 are rejected, the alpha for the rejected null hypotheses will be propagated to H_3, such that, H_3 will be tested at the updated alpha level (one-sided 0.025 if both H_1 and H_2 are rejected; one-sided 0.0225 if H_1 is rejected but H_2 is not rejected; one-sided 0.0025 if H_2 is rejected but H_1 is not rejected); if H_3 is rejected, the alpha will be propagated to H_2 or H_4 based on the initial step rejection status.

- With H_1 rejected but H_2 not rejected at the initial step, if H_3 is rejected at one-sided 0.0225, the alpha will be propagated to H_2, such that, H_2 will be tested again at one-sided 0.025; if H_2 is rejected at one-sided 0.025, the alpha will be further propagated to H_4. Otherwise, the testing procedure will stop.
- With H_2 rejected but H_1 not rejected at the initial step, if H_3 is rejected at one-sided 0.0025, the alpha will be propagated to H_4, such that, H_4 will be tested at one-sided 0.0025. Otherwise, the testing procedure will stop.
- With both H_1 and H_2 rejected at the initial step, if H_3 is rejected at one-sided 0.025, the alpha will be propagated to H_4, such that, H_4 will be tested at full alpha (one-sided 0.025). Otherwise, the testing procedure will stop.

Figure 2-1 provides graphical illustration of the sequentially rejective testing procedure.

Figure 2-1 Graphical illustration of the sequentially rejective testing procedure



Based on the above testing procedure, if applicable, H_1, H_2 and H_3 will be tested based on the corresponding mixed models for repeated measures (descriptions below and Section 2.6.1.1), H_4 will be tested based on the longitudinal proportional cumulative odds model (Section 2.6.2)

NT-proBNP analyses

The changes from baseline in log(NT-proBNP) will be analyzed using a mixed model for repeated measures (MMRM), in which, the response variable will be the changes from baseline in log(NT proBNP); stratum (ACEi, ARB, No RASi), region, treatment (LCZ696, IMT), visit and treatment-by-visit interaction will be included as fixed-effect factors; baseline log(NT-proBNP) (BLNTBNP), stratum-by-BLNTBNP and visit-by-BLNTBNP interactions will be included as covariates; the within-patient covariance will be modeled using an unstructured covariance matrix (a common matrix for the two treatment groups). The analysis will include all post-baseline scheduled visits up to Week 12 (Week 4, Week 12) and will be performed

based on the likelihood method with an assumption of missing at random (MAR) for missing data (Siddiqui et al 2009, National Research Council 2010).

Based on the MMRM model, the primary null hypothesis, H_1, will be tested based on the testing strategy, the estimates and the corresponding two-sided 95% confidence intervals will be provided for the adjusted geometric means for the ratio to baseline in NT-proBNP at Week 12 in both treatment groups (LCZ696 and IMT), and for the ratio of the adjusted geometric means (LCZ696/IMT). The corresponding one-sided p-values will also be provided based on the MMRM model.

The FAS will be used for the above primary analyses.

As a supportive analysis, the same approach will also be performed on the PPS.

6MWD analyses

The changes from baseline in 6MWD will be analyzed using a MMRM model, in which, the response variable will be the changes from baseline in 6MWD; stratum (ACEi, ARB, No RASi), region, treatment (LCZ696, IMT), visit and treatment-by-visit interaction will be included as fixed-effect factors; B6MWD, BSBP, stratum-by-B6MWD, stratum-by-BSBP and visit-by-B6MWD interactions will be included as covariates; the within-patient covariance will be modeled using an unstructured covariance matrix (a common matrix for the two treatment groups). The analysis will include all patients with B6MWD ranging from 100 meters to 450 meters, and all post-baseline scheduled visits up to Week 24 (Week 16 and Week 24), and will be performed based on the likelihood method with an assumption of MAR for missing data (other than due to death).

Based on the MMRM model, the primary null hypothesis, H_2, will be tested based on the testing strategy, the estimates and the two-sided 95% confidence intervals will be provided for the adjusted means for the change from baseline in 6MWD at each post-baseline scheduled visit (Week 16 and Week 24) in both treatment groups (LCZ696 and IMT), and for the adjusted mean difference (LCZ696 minus IMT) at each post-baseline scheduled visit (Week 16 and Week 24). The corresponding one-sided p-values will also be provided based on the MMRM model.

All patients in the FAS with B6MWD ranging from 100 meters to 450 meters will be used for the above primary analysis.

As a supportive analysis, the same approach will also be performed on the FAS with all patients included, regardless of B6MWD values, and on the PPS.

In addition, the improvement and the deterioration in 6MWD will be analyzed separately using longitudinal binary logistic regression models, in which, the response variables will be, the 6MWD improvement (event defined by at least 30 meters improvement in 6MWD) and the 6MWD deterioration (event defined by at least 30 meters deterioration in 6MWD), accordingly; stratum (ACEi, ARB, No RASi), region, treatment (LCZ696, IMT), visit and treatment-by-visit interaction will be included as fixed-effect factors; B6MWD, BSBP, stratum-by-B6MWD, stratum-by-BSBP and visit-by-B6MWD interactions will be included as covariates; patient-specific intercepts will be included as random effects. The analysis will include all post-baseline

scheduled visits up to Week 24 (Week 16 and Week 24) and will be performed based on the likelihood method with an assumption of MAR for missing data (other than due to death).

Based on the longitudinal binary logistic regression models, the estimates and the two-sided 95% confidence intervals will be provided for the adjusted odds ratios (LCZ696/IMT) at each post-baseline scheduled visit (Week 16 and Week 24). The corresponding one-sided p-values will also be provided based on the longitudinal binary logistic model.

Besides, the 6MWD will be summarized by treatment group and visit, for overall and for each stratum (ACEi, ARB, No RASi), using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum.

As additional supportive analyses, the Mantel-Haenszel estimates and the corresponding 95% confidence intervals will be provided by visit, for the relative risks and the risk differences, for 6MWD improvement and 6MWD deterioration, accordingly.

In addition to the 6MWD, the pre-test and post-test heart rate (HR), SBP, DBP, Borg scale for dyspnea and Borg scale for fatigue, will also be recorded on the CRF page: “6 Minute Walk Test and Borg Scale” and will be summarized by treatment group and visit, for overall and for each stratum (ACEi, ARB, No RASi), using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum.

The FAS will be used for the above analyses.

2.5.3 Handling of missing values/censoring/discontinuations

Missing data will be handled with the likelihood method with the assumption that they are MAR in the primary analysis.

2.5.4 Supportive and sensitivity analyses

In order to explore the robustness of the MAR assumption on the primary analysis, a sensitivity analysis will be carried out which assesses the situation where the data are missing not at random (MNAR); In particular, the sensitivity analysis will assume that LCZ696 patients who discontinue from the study treatment due to adverse events, death or lack of efficacy will not have adhered to therapy if they had stayed in the study.

Patients who discontinue from the study treatment due to adverse events, death or lack of efficacy are defined as those who discontinued from the study treatment during the randomized treatment epoch, and had the primary reason for discontinuation of study treatment (CRF page: “End of Study Treatment”) specified as “Adverse event”, “Death” or “Lack of efficacy”.

A controlled multiple imputation approach based on pattern mixture models will be applied (Carpenter and Kenward 2013), whereby all missing data of patients in the LCZ696 group who permanently discontinue study treatment due to: adverse events, death or lack of efficacy, will be assumed to behave like patients in the IMT group after permanent discontinuation and will be imputed based on the data from patients in the IMT group according to the stratum (ACEi, ARB, No RASi).

- For changes from baseline in log(NT-proBNP), for each post-baseline scheduled visit up to Week 12 (Week 4, Week 12), the imputation model will be a linear regression model specified using fully conditional specifications, in which, the response variable will be the

change from baseline in log(NT-proBNP) at the visit; region will be included as a fixed-effect factor; BLNTBNP, change from baseline in log(NT-proBNP) at the closest previous post-baseline scheduled visit (not applicable for Week 4, Week 4 for Week 12) will be included as covariates. For each stratum (ACEi, ARB, No RASi), the imputation model will be fitted separately based on the data from all patients in the IMT group within the stratum.

- For changes from baseline in 6MWD, for each post-baseline scheduled visit up to Week 24 (Week 16 and Week 24), the imputation model will be a linear regression model specified using fully conditional specifications, in which, the response variable will be the change from baseline in 6MWD at the visit; region will be included as a fixed-effect factor; B6MWD, BSBP, change from baseline in 6MWD at the closest previous post-baseline scheduled visit (not applicable for Week 16, Week 16 for Week 24) will be included as covariates. For each stratum (ACEi, ARB, No RASi), the imputation model will be fitted separately based on the data from patients with B6MWD ranging from 100 meters to 450 meters in the IMT group within the stratum.

All other missing data will be considered as MAR within each arm and will be imputed using a multiple imputation approach (Carpenter and Kenward 2013) based on observed data within each arm according to the stratum. Of note, patients who permanently discontinue study treatment, are not considered discontinued from the study. It is planned to continue to collect and use data at scheduled visits even after permanent discontinuation of study treatment.

- For changes from baseline in log(NT-proBNP), for each post-baseline scheduled visit up to Week 12 (Week 4, Week 12), the imputation model will be a linear regression model specified using fully conditional specifications, in which, the response variable will be the change from baseline in log(NT-proBNP) at the visit; region will be included as a fixed-effect factor; BLNTBNP, change from baseline in log(NT-proBNP) at the closest previous post-baseline scheduled visit (not applicable for Week 4, Week 4 for Week 12) will be included as covariates. For each stratum (ACEi, ARB, No RASi) and each arm, the imputation model will be fitted separately based on the data from all patients in the arm within the stratum.
- For changes from baseline in 6MWD, for each post-baseline scheduled visit up to Week 24 (Week 16 and Week 24), the imputation model will be a linear regression model specified using fully conditional specifications, in which, the response variable will be the change from baseline in 6MWD at the visit; region will be included as a fixed-effect factor; B6MWD, BSBP, change from baseline in 6MWD at the closest previous post-baseline scheduled visit (not applicable for Week 16, Week 16 for Week 24) will be included as covariates. For each stratum (ACEi, ARB, No RASi) and each arm, the imputation model will be fitted separately based on the data from patients with B6MWD ranging from 100 meters to 450 meters in the arm within the stratum.

Furthermore, the missing data patterns in the change from baseline in the log(NT-proBNP) and the change from baseline in 6MWD will be explored via by-treatment-group box plots for patients who permanently discontinue from study treatment due to adverse events, death or lack of efficacy, for patients who permanently discontinue from study treatment due to any other reasons, for patients who did not permanently discontinue from study treatment, and for patients

who take open label LCZ696 after permanent discontinuation from study treatment. Similar box plots will also be provided for on-treatment assessments.

As additional supportive analyses, the primary analyses for the primary endpoints (change from baseline in log[NT-proBNP] at Week 12, and change from baseline in 6MWD at Week_24) described in Section 2.5.2 will be performed for the following data as applicable.

- Data where the 6MWD after the death were set to missing for patients who died.
- Data where the last observation carry forward (LOCF) approach was performed for missing data.
- Data where the responses after the first dose of open-label RASi or open-label LCZ696 after permanent discontinuation of the study treatment will be set to missing and considered MAR. (This is called on-treatment analysis.)

The FAS will be used for the above analyses.

As an additional supportive analysis, the change from baseline in 6MWD will be analyzed using a partial linear MMRM model, in which, the response variable will be the changes from baseline in 6MWD; stratum (ACEi, ARB, No RASi), region, treatment (LCZ696, IMT), visit and treatment-by-visit interaction will be included as fixed-effect factors; B6MWD and BSBP will be included as non-parametric fixed-effect covariates, denoted as g(B6MWD) and h(BSBP), stratum-by-g(B6MWD) interaction, stratum-by-h(BSBP) interaction, visit-by-g(B6MWD) interaction and treatment-by-g(B6MWD) interaction will also be included as covariates. The non-parametric effects will be approximated using natural cubic splines, in which the quintiles are used as knots. Based on the model, the estimated adjusted mean difference for change from baseline in 6MWD at Week 24 will be plotted against theB6MWD.

The FAS will be used for the above analyses.

2.5.5 Subgroup analyses

For each primary endpoint (Section 2.5.1), subgroup analyses will be performed utilizing the corresponding primary analysis model (Section 2.5.2), after adding the fixed-effect factors of subgroup variable and subgroup-variable-by-treatment interaction into the original model. The following subgroup variables (Section 2.2.1, Table 2-3) will be used for the subgroup analyses for both primary endpoints.

- Region
- Stratum (ACEi, ARB, No RASi)
- Age group (< 65 years, ≥ 65 years)
- Age group (< 75 years, ≥ 75 years)
- Sex
- Baseline NYHA class group
- Diabetes (history or HbA1c ≥ 6.5%)
- LVEF group (<= 60%, > 60%)
- Prior heart failure hospitalization

For changes from baseline in log(NT-proBNP), based on the subgroup analysis models, the estimates and the two-sided 95% confidence intervals will be provided for adjusted geometric means for ratio to baseline in NT-proBNP at Week 12 in both treatment groups (LCZ696 and IMT) within each subgroup, and for the ratio of the adjusted geometric means (LCZ696/IMT) at Week 12 within each subgroup; the p-value for the subgroup-variable-by-interaction will be provided for each subgroup variable.

For changes from baseline in 6MWD, based on the subgroup analysis models, the estimates and the two-sided 95% confidence intervals will be provided for the adjusted means for the change from baseline in 6MWD at Week 24 in both treatment groups (LCZ696 and IMT) within each subgroup, and for the adjusted mean difference (LCZ696 minus IMT) at Week 24 within each subgroup; the p-value for the subgroup-variable-by-interaction will be provided for each subgroup variable.

In addition, for each primary endpoint (Section 2.5.1), multivariate subgroup analyses will be performed utilizing the corresponding primary analysis model (Section 2.5.2), after adding the fixed-effect factors of all the follwoing subgroup variables and subgroup-variable-by-treatment interactions into the original model. The p value for the global interaction test (null hypothesis: all interaction terms are equal to zero) will be presented, together with the p value for each interaction.

- Region
- Stratum (ACEi, ARB, No RASi)
- Age group (< 75 years, ≥ 75 years)
- Sex
- Baseline NYHA class group
- Diabetes (history or HbA1c ≥ 6.5%)
- LVEF group (<= 60%, > 60%)
- Prior heart failure hospitalization

The FAS will be used for the above subgroup analyses.

2.6 Analysis of secondary efficacy objective(s)

All secondary efficacy variables will be analyzed using the FAS, unless otherwise specified.

2.6.1 Kansas City Cardiomyopathy Questionnaire (KCCQ)

The KCCQ consists of 23 individual items comprising 10 summary scores.

For each individual item, the item response will be mapped into numeric response score, using the mapping criteria defined in Table 2-6.

Table 2-6 KCCQ item response scoring

Item ID	Item Response	Response Score
1(a)-1(f)	EXTREMELY LIMITED	0.00
	QUITE A BIT LIMITED	0.25

Item ID	Item Response	Response Score
	MODERATELY LIMITED	0.50
	SLIGHTLY LIMITED	0.75
	NOT AT ALL LIMITED	1.00
	LIMITED FOR OTHER REASONS OR DID NOT DO THE ACTIVITY	Missing
2	MUCH WORSE	0.00
	SLIGHTLY WORSE	0.25
	NOT CHANGED	0.50
	SLIGHTLY BETTER	0.75
	MUCH BETTER	1.00
	I'VE HAD NO SYMPTOMS OVER THE LAST 2 WEEKS	0.50
3	EVERY MORNING	0.00
	3 OR MORE TIMES A WEEK, BUT NOT EVERY DAY	0.25
	1-2 TIMES A WEEK	0.50
	LESS THAN ONCE A WEEK	0.75
	NEVER OVER THE PAST 2 WEEKS	1.00
4, 6 and 8	EXTREMELY BOTHERSOME	0.00
	QUITE A BIT BOTHERSOME	0.25
	MODERATELY BOTHERSOME	0.50
	SLIGHTLY BOTHERSOME	0.75
	NOT AT ALL BOTHERSOME	1.00
	I'VE HAD NO SWELLING	1.00
	I'VE HAD NO FATIGUE	1.00
	I'VE HAD NO SHORTNESS OF BREATH	1.00
5 and 7	ALL OF THE TIME	0.00
	SEVERAL TIMES PER DAY	1/6
	AT LEAST ONCE A DAY	1/3
	3 OR MORE TIMES A WEEK, BUT NOT EVERY DAY	0.50
	1-2 TIMES A WEEK	2/3
	LESS THAN ONCE A WEEK	5/6
	NEVER OVER THE PAST 2 WEEKS	1.00
9	EVERY NIGHT	0.00
	3 OR MORE TIMES A WEEK, BUT NOT EVERY DAY	0.25
	1-2 TIMES A WEEK	0.50
	LESS THAN ONCE A WEEK	0.75
	NEVER OVER THE PAST 2 WEEKS	1.00
10	NOT AT ALL SURE	0.00
	NOT VERY SURE	0.25
	SOMEWHAT SURE	0.50
	MOSTLY SURE	0.75
	COMPLETELY SURE	1.00
11	DO NOT UNDERSTAND AT ALL	0.00

Item ID	Item Response	Response Score
	DO NOT UNDERSTAND VERY WELL	0.25
	SOMEWHAT UNDERSTAND	0.50
	MOSTLY UNDERSTAND	0.75
	COMPLETELY UNDERSTAND	1.00
12	IT HAS EXTREMELY LIMITED MY ENJOYMENT OF LIFE	0.00
	IT HAS LIMITED MY ENJOYMENT OF LIFE QUITE A BIT	0.25
	IT HAS MODERATELY LIMITED MY ENJOYMENT OF LIFE	0.50
	IT HAS SLIGHTLY LIMITED MY ENJOYMENT OF LIFE	0.75
	IT HAS NOT LIMITED MY ENJOYMENT OF LIFE AT ALL	1.00
13	NOT AT ALL SATISFIED	0.00
	MOSTLY DISSATISFIED	0.25
	SOMEWHAT SATISFIED	0.50
	MOSTLY SATISFIED	0.75
	COMPLETELY SATISFIED	1.00
14	I FELT THAT WAY ALL OF THE TIME	0.00
	I FELT THAT WAY MOST OF THE TIME	0.25
	I OCCASIONALLY FELT THAT WAY	0.50
	I RARELY FELT THAT WAY	0.75
	I NEVER FELT THAT WAY	1.00
15(a)-15(d)	SEVERELY LIMITED	0.00
	LIMITED QUITE A BIT	0.25
	MODERATELY LIMITED	0.50
	SLIGHTLY LIMITED	0.75
	DID NOT LIMIT AT ALL	1.00
	DOES NOT APPLY OR DID NOT DO FOR OTHER REASONS	Missing

KCCQ physical limitation score

The KCCQ physical limitation score will be defined as

100 × the mean of the non-missing response scores among items {1(a)-1(f)},

if there are at least three non-missing response scores among items {1(a)-1(f)}; the score will be left as missing if there are not enough non-missing response scores among these items.

KCCQ symptom stability score

The KCCQ symptom stability score will be defined as

100 × the response score for item 2,

if the response score for item 2 is non-missing; the score will be left as missing if the response score is missing.

KCCQ symptom frequency score

The KCCQ symptom frequency score will be defined as

100 × the mean of the non-missing response scores among items {3, 5 , 7, 9},

if there are at least two non-missing response scores among items {3, 5 , 7, 9}; the score will be left as missing if there are not enough non-missing response scores among these items.

KCCQ symptom burden score

The KCCQ symptom burden score will be defined as

100 × the mean of the non-missing response scores among items {4, 6, 8},

if there are at least one non-missing response scores among items {4, 6, 8}; the score will be left as missing if the response scores are missing for all items among items {4, 6, 8}

KCCQ self-efficacy score

The KCCQ self-efficacy score will be defined as

100 × the mean of the non-missing response scores among items {10, 11},

if there are at least one non-missing response scores among items {10, 11}; the score will be left as missing if the response scores are missing for both item 10 and item 11.

KCCQ quality of life score

The KCCQ quality of life score will be defined as

100 × the mean of the non-missing response scores among items {12, 13, 14},

if there are at least one non-missing response scores among items {12, 13, 14}; the score will be left as missing if the response scores are missing for all items among items {12, 13, 14}.

KCCQ social limitation score

The KCCQ social limitation score will be defined as

100 × the mean of the non-missing response scores among items {15(a)-15(d)},

if there are at least two non-missing response scores among items {15(a)-15(d)}; the score will be left as missing if there are not enough non-missing response scores among these items.

Table 2-7 KCCQ summary scores

Summary Score	Relevant Items	Minimum Number of Non-Missing Response Scores
KCCQ physical limitation score	1(a)-1(f)	3
KCCQ symptom stability score	2	1
KCCQ symptom frequency score	3, 5, 7, 9	2
KCCQ symptom burden score	4, 6, 8	1
KCCQ self-efficacy score	10, 11	1
KCCQ quality of life score	12, 13, 14	1
KCCQ social limitation score	15(a)-15(d)	2

In summary, the above seven KCCQ summary scores will be defined as 100 times the mean of the non-missing response scores for the relevant items, if there are enough items with non-missing response scores among the relevant items. The relevant items and the minimum number of non-missing response scores are given in Table 2-7.

KCCQ total symptom score

The KCCQ total symptom score will be defined as the mean of the non-missing scores among the KCCQ symptom frequency score and the KCCQ symptom burden score. The KCCQ total symptom score will be left as missing, if the KCCQ symptom frequency score and the KCCQ symptom burden score are both missing.

KCCQ clinical summary score (CSS)

The KCCQ CSS will be defined as the mean of the non-missing scores among the following.

- KCCQ physical limitation score
- KCCQ total symptom score

If the above scores are both missing, the KCCQ CSS will be left as missing.

KCCQ overall summary score (OSS)

The KCCQ OSS will be defined as the mean of the non-missing scores among the following.

- KCCQ physical limitation score
- KCCQ total symptom score
- KCCQ quality of life score
- KCCQ social limitation score

If the above scores are all missing, the KCCQ OSS will be left as missing.

2.6.1.1 KCCQ clinical summary score (CSS) change from baseline

For patients who have died, the KCCQ CSS after the death will be imputed as zero.

For patients who permanently discontinue study treatment, values collected after permanent discontinuation will generally be included in the analysis.

The changes from baseline in KCCQ CSS will be analyzed using a MMRM model, in which, the response variable will be the changes from baseline in KCCQ CSS; stratum (ACEi, ARB, No RASi), region, treatment (LCZ696, IMT), visit and treatment-by-visit interaction will be included as fixed-effect factors; baseline KCCQ CSS (BCSS), BSBP, stratum-by-BCSS, stratum-by-BSBP, and visit-by-BCSS interactions will be included as covariates; the within-patient covariance will be modeled using an unstructured covariance matrix (a common matrix for the two treatment groups). If the model fitting procedure does not converge, the AR(1) covariance structure will be used. The analysis will include all post-baseline scheduled visits up to Week 24 (Week 4, Week 12, Week 16 and Week 24) and will be performed based on the likelihood method with an assumption of MAR for missing data (other than due to death).

Based on the MMRM model, the secondary null hypothesis, H_3, will be tested (if applicable) based on the testing strategy, the estimates and the two-sided 95% confidence intervals will be provided for the adjusted mean changes from baseline in KCCQ CSS at each post-baseline scheduled visit (Week 4, Week 12, Week 16 and Week 24) in both treatment groups (LCZ696 and IMT), and for the adjusted mean difference (LCZ696 minus IMT) at each post-baseline scheduled visit (Week 4, Week 12, Week 16 and Week 24). The corresponding one-sided p-values will also be provided based on the MMRM model.

In addition, the improvement and the deterioration in KCCQ CSS will be analyzed separately, to further illustrate the clinical relevance of the observed differences in terms of proportion of responders, using longitudinal binary logistic regression models, in which, the response variables will be, the KCCQ CSS improvement (event defined by at least 5-points improvement in KCCQ CSS) and the KCCQ CSS deterioration (event defined by at least 5-points deterioration in KCCQ CSS), accordingly; stratum (ACEi, ARB, No RASi), region, treatment (LCZ696, IMT), visit and treatment-by-visit interaction will be included as fixed-effect factors; BCSS, BSBP, stratum-by-BCSS, stratum-by-BSBP and visit-by-BCSS interactions will be included as covariates; patient-specific intercepts will be included as random effects. The analysis will include all post-baseline scheduled visits up to Week 24 (Week 4, Week 12, Week 16 and Week 24) and will be performed based on the likelihood method with an assumption of MAR for missing data (other than due to death). No multiplicity adjustment will be done for these supportive responder analyses (i.e., the tests for the responder analyses will be performed at nominal level of one-sided 0.025).

Based on the longitudinal binary logistic regression models, the estimates and the two-sided 95% confidence intervals will be provided for the adjusted odds ratios (LCZ696 / IMT) at each post-baseline scheduled visit (Week 4, Week 12, Week 16 and Week 24). The corresponding one-sided p-values will also be provided based on the longitudinal binary logistic model.

Besides, the KCCQ CSS and the change from baseline in KCCQ CSS will be summarized by treatment group and visit, for overall and for each stratum (ACEi, ARB, No RASi), using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum; the improvement and the deterioration in KCCQ CSS of at least 5-points will be summarized by treatment group and visit, for overall and for each stratum (ACEi, ARB, No RASi), using the numbers and percentages of patients whose KCCQ CSS improved at least 5 points, whose KCCQ CSS deteriorated at least 5 points, whose KCCQ CSS changed less than 5 points, and whose KCCQ CSS is missing.

As additional supportive analyses, the Mantel-Haenszel estimates and the corresponding 95% confidence intervals will be provided by visit, for the relative risks and the risk differences, for KCCQ CSS improvement and KCCQ CSS deterioration, accordingly.

The FAS will be used for the above analyses.

2.6.1.2 KCCQ overall summary score (OSS) change from baseline

As additional exploratory analyses, the same analyses as KCCQ CSS will also be performed for KCCQ OSS, except that the BCSS in the analysis models will be replaced by baseline KCCQ OSS (BOSS), stratum-by-BCSS and visit-by-BCSS interactions in the analysis models will be replaced by stratum-by-BOSS and visit-by-BOSS interactions.

The FAS will be used for the above analyses.

2.6.2 NYHA class change

The heart failure symptoms and the NYHA class will be recorded on the CRF page: “Signs and Symptoms of Heart Failure including New York Heart Association Classification”.

Table 2-8 defines the NYHA class change at each post-baseline visit.

Table 2-8 NYHA Class Change

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

For patients who died, the NYHA class will be set to a firth category (death), and the NYHA class change after the death will be categorized into "worsened". For patients who permanently discontinue study treatment, values collected after permanent discontinuation will generally be included in the analysis.

The NYHA class change will be analyzed using a longitudinal proportional cumulative odds model, in which, the response variable will be the NYHA class change (order defined by "improved" < "unchanged" < "worsened"); stratum (ACEi, ARB, No RASi), region, treatment (LCZ696, IMT), visit, treatment-by-visit interaction, baseline NYHA class, stratum-by-baseline and visit-by-baseline interactions will be included as fixed-effect factors; BSBP and stratum-by-BSBP interaction will be included as covariates; patient-specific intercepts will be included as random effects. The analysis will include all post-baseline scheduled visits up to Week 24 (Week 2, Week 4, Week 8, Week 12, Week 16, Week 20 and Week 24) and will be performed based on the likelihood method with an assumption of MAR for missing data (other than due to death).

Based on the longitudinal proportional cumulative odds model, the secondary null hypothesis, H_4, will be tested based on the testing strategy (if applicable), the estimate and the two-sided 95% confidence interval will be provided for the adjusted odds ratio (LCZ696 / IMT) at each post-baseline scheduled visit (Week 2, Week 4, Week 8, Week 12, Week 16, Week 20 and Week 24). The corresponding one-sided p-values will also be provided based on the model.

In addition, the improvement and the deterioration in NYHA class will be analyzed separately using longitudinal binary logistic regression models, in which, the response variables will be, accordingly, the NYHA class improvement (event defined by NYHA class change of level "improved") and the NYHA class deterioration (event defined by NYHA class change of level "worsened"); stratum (ACEi, ARB, No RASi), region, treatment (LCZ696, IMT), visit, treatment-by-visit interaction, baseline NYHA class, stratum-by-baseline and visit-by-baseline interactions will be included as fixed-effect factors; BSBP and stratum-by-BSBP interaction will be included as covariates; patient-specific intercepts will be included as random effects. The analysis will include all post-baseline scheduled visits up to Week 24 (Week 2, Week 4, Week 8, Week 12, Week 16, Week 20 and Week 24) and will be performed based on the likelihood method with an assumption of MAR for missing data (other than due to death).

Based on the longitudinal binary logistic regression models, the estimates and the two-sided 95% confidence intervals will be provided for the adjusted odds ratios (LCZ696 / IMT) at each post-baseline scheduled visit (Week 2, Week 4, Week 8, Week 12, Week 16, Week 20 and Week 24). The one-sided p-values will also be provided based on the longitudinal binary logistic model.

Besides, the NYHA class and the NYHA class change will be summarized by treatment group and visit, for overall and for each stratum (ACEi, ARB, No RASi), using the numbers and percentages of patients within each category.

The heart failure symptoms will be summarized by treatment group and visit, for overall and for each stratum (ACEi, ARB, No RASi), using the numbers and percentages of patients with each heart failure symptom.

As additional supportive analyses, the Mantel-Haenszel estimates and the corresponding 95% confidence intervals will be provided by visit, for the relative risks and the risk differences, for NYHA improvement and NYHA deterioration, accordingly.

The FAS will be used for the above analyses.

2.6.3 Short Form Health Survey (SF-36)

The SF-36 consists of 36 individual items comprising eight scales: physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional and mental health. For each of the eight scales, a raw score ranging from 0 to 100, and a normalized score (normed to the US population) will be calculated. Based on the normalized scores, two overall summary scores: physical component summary (PCS) score and mental component summary (MCS) score will be calculated. The raw scores, the normalized scores and the overall summary scores will be provided by Optum.

The raw scores, the normalized scores and the overall summary scores will be summarized by treatment group and visit, accordingly, for overall and for each stratum (ACEi, ARB, No RASi), using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum.

2.6.3.1 SF-36 physical component summary (PCS) score

For patients who have died, the SF-36 PCS score after the death will be imputed as zero. For patients who permanently discontinue from study treatment, values collected after permanent discontinuation will generally be included in the analysis.

The changes from baseline in SF-36 PCS score will be analyzed using a MMRM model, in which, the response variable will be the changes from baseline in SF-36 PCS score; stratum (ACEi, ARB, No RASi), region, treatment (LCZ696, IMT), visit and treatment-by-visit interaction will be included as fixed-effect factors; baseline SF-36 PCS score (BPSS), BSBP, stratum-by-BPSS, stratum-by-BSBP and visit-by-BPSS interactions will be included as covariates; the within-patient covariance will be modeled using an unstructured covariance matrix (a common matrix for the two treatment groups). The analysis will include all post-baseline scheduled visits up to Week 24 (Week 16 and Week 24) and will be performed based on the likelihood method with an assumption of MAR for missing data (other than due to death).

Based on the MMRM model, the estimates and the two-sided 95% confidence intervals will be provided for the adjusted mean SF-36 PCS scores and the adjusted mean changes from baseline in the SF-36 PCS score at each post-baseline scheduled visit (Week 16 and Week 24) in both treatment groups (LCZ696 and IMT), and for the adjusted mean difference (LCZ696 minus

IMT) at each post-baseline scheduled visit (Week 16 and Week 24). The one-sided p-values will also be provided based on the MMRM model.

In addition, the improvement and the deterioration in SF-36 PCS score will be analyzed separately using longitudinal binary logistic regression models, in which, the response variables will be, the SF-36 PCS score improvement (event defined by at least 3.4 points improvement in SF-36 PCS score) and the SF-36 PCS deterioration (event defined by at least 3.4 points deterioration in SF-36 PCS score), accordingly; stratum (ACEi, ARB, No RASi), region, treatment (LCZ696, IMT), visit and treatment-by-visit interaction will be included as fixed-effect factors; BPSS, BSBP, stratum-by-BPSS, stratum-by-BSBP and visit-by-BPSS interactions will be included as covariates; patient-specific intercepts will be included as random effects. The analysis will include all post-baseline scheduled visits up to Week 24 (Week 16 and Week 24) and will be performed based on the likelihood method with an assumption of MAR for missing data (other than due to death).

Based on the longitudinal binary logistic regression models, the estimates and the two-sided 95% confidence intervals will be provided for the adjusted odds ratios (LCZ696 / IMT) at each post-baseline scheduled visit (Week 16 and Week 24). The corresponding one-sided p-values will also be provided based on the longitudinal binary logistic model.

As additional supportive analyses, the Mantel-Haenszel estimates and the corresponding 95% confidence intervals will be provided by visit, for the relative risks and the risk differences, for SF-36 PCS score improvement and SF-36 PCS score deterioration, accordingly.

The FAS will be used for the above analyses.

2.6.3.2 SF-36 mental component summary score

The same analyses as SF-36 PCS score will also be performed for SF-36 MCS score, except that the BPSS in the models will be replaced by baseline SF-36 MCS score (BMSS), stratum-by-BPSS and visit-by-BPSS interactions in the analysis models will be replaced by stratum-by-BMSS and visit-by-BMSS interactions, and that the cut-off of 3.4 points for the improvement and the deterioration will be replaced by 4.6 points.

The FAS will be used for the above analyses.

2.6.4 Handling of missing values/censoring/discontinuations

In the analyses described in the above sections (Section 2.6.1 to 2.6.3), death is considered the worst possible outcome and hence scheduled visits after death will be imputed to be the worst score accordingly. Other missing data is assumed to be MAR.

As in Section 2.5.3, patients who discontinue from the study treatment due to adverse events, death or lack of efficacy are defined as patients who discontinued from the study treatment during the randomized treatment epoch, and had the primary reason for discontinuation of study treatment (CRF page: “End of Study Treatment”) specified as “Adverse event”, “Death” or “Lack of efficacy”.

As sensitivity analyses, the controlled multiple imputation approach based on pattern mixture models will be applied (Carpenter and Kenward 2013), whereby all missing data of patients in the LCZ696 group who permanently discontinue study treatment due to: adverse events, death

or lack of efficacy, will be assumed to behave like patients in the IMT group after permanent discontinuation and will be imputed based on the data from patients in the IMT group according to the stratum (ACEi, ARB, No RASi).

- For changes from baseline in KCCQ CSS, for each post-baseline scheduled visit up to Week 24 (Week 4, Week 12, Week 16 and Week 24), the imputation model will be a linear regression model specified using fully conditional specifications, in which, the response variable will be the change from baseline in KCCQ CSS at the visit; region will be included as a fixed-effect factor; BCSS, BSBP, change from baseline in KCCQ CSS at the closest previous post-baseline scheduled visit (not applicable for Week 4, Week 4 for Week 12, Week12 for Week 16, Week 16 for Week 24) will be included as covariates. For each stratum (ACEi, ARB, No RASi), the imputation model will be fitted separately based on the data from patients in the IMT group within the stratum.
- For changes from baseline in KCCQ OSS, for each post-baseline scheduled visit up to Week 24 (Week 4, Week 12, Week 16 and Week 24), the imputation model will be a linear regression model specified using fully conditional specifications, in which, the response variable will be the change from baseline in KCCQ OSS at the visit; region will be included as a fixed-effect factor; BOSS, BSBP, change from baseline in KCCQ OSS at the closest previous post-baseline scheduled visit (not applicable for Week 4, Week 4 for Week 12, Week12 for Week 16, Week 16 for Week 24) will be included as covariates. For each stratum (ACEi, ARB, No RASi), the imputation model will be fitted separately based on the data from patients in the IMT group within the stratum.
- For NYHA class change, for each post-baseline scheduled visit up to Week 24 (Week 2, Week 4, Week 8, Week 12, Week 16, Week 20 and Week 24), the imputation model will be a proportional cumulative odds model specified using fully conditional specifications, in which, the response variable will be the NYHA class change (order defined by “improved” < “unchanged” < “worsened”) at the visit; region, baseline NYHA class, NYHA class change at the closest previous post-baseline scheduled visit (not applicable for Week 2, Week 2 for Week 4, Week 4 for Week 8, Week 8 for Week 12, Week 12 for Week 16, Week 16 for Week 20 and Week 20 for Week 24) will be included as fixed-effect factors; BSBP will be included as a covariate. For each stratum (ACEi, ARB, No RASi), the imputation model will be fitted separately based on the data from patients in the IMT group within the stratum.
- For changes from baseline in SF-36 PCS score, for each post-baseline scheduled visit up to Week 24 (Week 16 and Week 24), the imputation model will be a linear regression model specified using fully conditional specifications, in which, the response variable will be the change from baseline in SF-36 PCS score at the visit; region will be included as a fixed-effect factor; BPSS, BSBP, change from baseline in SF-36 PCS score at the closest previous post-baseline scheduled visit (not applicable for Week 16, Week 16 for Week 24) will be included as covariates. For each stratum (ACEi, ARB, No RASi), the imputation model will be fitted separately based on the data from patients in the IMT group within the stratum.
- For changes from baseline in SF-36 MCS score, for each post-baseline scheduled visit up to Week 24 (Week 16 and Week 24), the imputation model will be a linear regression model specified using fully conditional specifications, in which, the response variable will

be the change from baseline in SF-36 MCS score at the visit; region will be included as a fixed-effect factor; BMSS, BSBP, change from baseline in SF-36 MCS score at the closest previous post-baseline scheduled visit (not applicable for Week 16, Week 16 for Week 24) will be included as covariates. For each stratum (ACEi, ARB, No RASi), the imputation model will be fitted separately based on the data from patients in the IMT group within the stratum.

All other missing data will be considered as MAR within each arm and will be imputed using a multiple imputation approach (Carpenter and Kenward 2013) based on observed data within each arm according to the stratum. Of note, patients who permanently discontinue study treatment, are not considered discontinued from the study. It is planned to continue to collect and use data at scheduled visits even after permanent discontinuation of study treatment.

- For changes from baseline in KCCQ CSS, for each post-baseline scheduled visit up to Week 24 (Week 4, Week 12, Week 16 and Week 24), the imputation model will be a linear regression model specified using fully conditional specifications, in which, the response variable will be the change from baseline in KCCQ CSS at the visit; region will be included as a fixed-effect factor; BCSS, BSBP, change from baseline in KCCQ CSS at the closest previous post-baseline scheduled visit (not applicable for Week 4, Week 4 for Week 12, Week12 for Week 16, Week 16 for Week 24) will be included as covariates. For each stratum (ACEi, ARB, No RASi) and each arm, the imputation model will be fitted separately based on the data from patients in the arm within the stratum.
- For changes from baseline in KCCQ OSS, for each post-baseline scheduled visit up to Week 24 (Week 4, Week 12, Week 16 and Week 24), the imputation model will be a linear regression model specified using fully conditional specifications, in which, the response variable will be the change from baseline in KCCQ OSS at the visit; region will be included as a fixed-effect factor; BOSS, BSBP, change from baseline in KCCQ OSS at the closest previous post-baseline scheduled visit (not applicable for Week 4, Week 4 for Week 12, Week12 for Week 16, Week 16 for Week 24) will be included as covariates. For each stratum (ACEi, ARB, No RASi) and each arm, the imputation model will be fitted separately based on the data from patients in the arm within the stratum.
- For NYHA class change, for each post-baseline scheduled visit up to Week 24 (Week 2, Week 4, Week 8, Week 12, Week 16, Week 20 and Week 24), the imputation model will be a proportional cumulative odds model specified using fully conditional specifications, in which, the response variable will be the NYHA class change (order defined by “improved” < “unchanged” < “worsened”) at the visit; region, baseline NYHA class, NYHA class change at the closest previous post-baseline scheduled visit (not applicable for Week 2, Week 2 for Week 4, Week 4 for Week 8, Week 8 for Week 12, Week 12 for Week 16, Week 16 for Week 20 and Week 20 for Week 24) will be included as fixed-effect factors; BSBP will be included as a covariate. For each stratum (ACEi, ARB, No RASi) and each arm, the imputation model will be fitted separately based on the data from patients in the arm within the stratum.
- For changes from baseline in SF-36 PCS score, for each post-baseline scheduled visit up to Week 24 (Week 16 and Week 24), the imputation model will be a linear regression model specified using fully conditional specifications, in which, the response variable will be the change from baseline in SF-36 PCS score at the visit; region will be included as a

fixed-effect factor; BPSS, BSBP, change from baseline in SF-36 PCS score at the closest previous post-baseline scheduled visit (not applicable for Week 16, Week 16 for Week 24) will be included as covariates. For each stratum (ACEi, ARB, No RASi) and each arm, the imputation model will be fitted separately based on the data from patients in the arm within the stratum.

- For changes from baseline in SF-36 MCS score, for each post-baseline scheduled visit up to Week 24 (Week 16 and Week 24), the imputation model will be a linear regression model specified using fully conditional specifications, in which, the response variable will be the change from baseline in SF-36 MCS score at the visit; region will be included as a fixed-effect factor; BMSS, BSBP, change from baseline in SF-36 MCS score at the closest previous post-baseline scheduled visit (not applicable for Week 16, Week 16 for Week 24) will be included as covariates. For each stratum (ACEi, ARB, No RASi) and each arm, the imputation model will be fitted separately based on the data from patients in the arm within the stratum.

Furthermore, the missing data patterns in the secondary endpoint variables will be explored via by-treatment-group box plots or bar plots as appropriate for patients who permanently discontinue from study treatment due to adverse events or lack of efficacy, for patients who permanently discontinue study treatment due to any other reasons, for patients who did not permanently discontinue study treatment and for patients who take open label LCZ696 after permanent discontinuation from study treatment. Similar box plots or bar plots will also be provided for on-treatment assessments.

As additional supportive analyses, the main analyses for the secondary endpoints described in the above sections (Section 2.6.10 to 2.6.3), will be performed for the following data.

- Data where the responses after the death were set to missing for patients who died.
- Data where the LOCF approach was performed for missing data .
- Data where the responses after the first dose of open-label RASi or open-label LCZ696 after permanent discontinuation of study treatment will be set to missing and considered MAR. (This is called on-treatment analysis.)

2.6.5 Subgroup analyses

For each secondary endpoint, subgroup analyses will be performed utilizing the corresponding secondary analysis model (Section 2.6.1 to Section 2.6.3), after adding the fixed-effect factors of subgroup variable and subgroup-variable-by-treatment interaction into the original model. The following subgroup variables (Section 2.2.1, Table 2-3) will be used for the subgroup analyses for all secondary endpoints.

- Region
- Stratum (ACEi, ARB, No RASi)
- Age group (< 65 years, ≥ 65 years)
- Age group (< 75 years, ≥ 75 years)
- Sex
- Baseline NYHA class group

- Diabetes (history or HbA1c ≥ 6.5%)
- LVEF group (<= 60%, > 60%)
- Prior heart failure hospitalization

For changes from baseline in KCCQ CSS, KCCQ OSS, SF-36 PCS score and SF-36 MCS score, based on the corresponding subgroup analysis model, the estimates and the two-sided 95% confidence intervals will be provided for the adjusted means for the change from baseline at Week 24 in both treatment groups (LCZ696 and IMT) within each subgroup, and for the adjusted mean difference (LCZ696 minus IMT) at Week 24 within each subgroup; the p-value for the subgroup-variable-by-interaction will be provided for each subgroup variable.

For NYHA class change, based on the corresponding subgroup analysis model, the estimates and the two-sided 95% confidence intervals will be provided for the adjusted odds ratios (LCZ696 / IMT) at Week 24 within each subgroup; the p-value for the subgroup-variable-by-interaction will be provided for each subgroup variable.

The FAS will be used for the above subgroup analyses.

2.7 Analysis of the exploratory objectives

2.7.1 NT-proBNP up to Week 24

Plasma NT-proBNP will be collected using central lab.

All post-screening NT-proBNP from central lab will remain blinded until data base lock. Only the NT-proBNP at screening will be reported to the investigator and the Novartis clinical study team.

For the NT-proBNP, if the test value is below the LLOQ, the test value will be imputed by 0.5 × LLOQ; if the test value is above the ULOQ, the test value will be imputed by 1.5 × ULOQ.

For the NT-proBNP, unreliable test values will be set to missing.

For patients who permanently discontinue study treatment, values collected after permanent discontinuation will generally be included in the analysis.

The changes from baseline in log(NT-proBNP) will be analyzed using a MMRM model, in which, the response variable will be the changes from baseline in log(NT-proBNP); stratum (ACEi, ARB, No RASi), region, treatment (LCZ696, IMT), visit and treatment-by-visit interaction will be included as fixed-effect factors; BLNTBNP, stratum-by-BLNTBNP and visit-by-BLNTBNP interactions will be included as covariates; the within-patient covariance will be modeled using an unstructured covariance matrix (a common matrix for the two treatment groups). If the model fitting procedure does not converge, the AR(1) covariance structure will be used. The analysis will include all post-baseline scheduled visits up to Week 24 (Week 4, Week 12 and Week 24) and will be performed based on the likelihood method with an assumption of MAR for missing data.

Based on the MMRM model, the estimates and the two-sided 95% confidence intervals will be provided for the adjusted geometric means for the ratio to baseline in NT-proBNP at each post-baseline scheduled visit (Week 4, Week 12 and Week 24) in both treatment groups (LCZ696 and IMT), and for the ratio of the adjusted geometric means (LCZ696/IMT) at each post-

baseline scheduled visit (Week 4, Week 12 and Week 24). The corresponding one-sided p-values will also be provided based on the MMRM model.

As an additional supportive analysis, the changes from baseline in log(NT-proBNP) will be analyzed using a MMRM model, in which, the response variable will be the changes from baseline in log(NT-proBNP); stratum (ACEi, ARB, No RASi), region, visit, treatment (LCZ696, IMT) and treatment-by-visit interaction will be included as fixed-effect factors; BLNTBNP, BSBP, stratum-by-BLNTBNP, stratum-by-BSBP and visit-by-BLNTBNP interactions will be included as covariates; the within-patient covariance will be modeled using an unstructured covariance matrix (a common matrix for the two treatment groups). If the model fitting procedure does not converge, the AR(1) covariance structure will be used. The analysis will include all post-baseline scheduled visits up to Week 24 (Week 4, Week 12 and Week 24) and will be performed based on the likelihood method with an assumption of MAR for missing data.

Based on the supportive MMRM model, the estimates and the two-sided 95% confidence intervals will be provided for the adjusted geometric means for the ratio to baseline in NT-proBNP at each post-baseline scheduled visit (Week 4, Week 12 and Week 24) in both treatment groups (LCZ696 and IMT), and for the ratio of the adjusted geometric means (LCZ696/IMT) at each post-baseline scheduled visit (Week 4, Week 12 and Week 24). The corresponding one-sided p-values will also be provided based on the MMRM model.

Besides, the NT-proBNP test values and NT-proBNP ratio to baseline will be summarized by treatment group and visit, for overall and for each stratum (ACEi, ARB, No RASi), using number of observations, mean, standard deviation, geometric mean, 95% confidence interval for geometric mean, median, minimum, Q1, Q3, maximum, number of observations below the LLOQ and number of observations above the ULOQ.

The FAS will be used for the above analyses.

2.7.2 Estimated glomerular filtration rate (eGFR)

The eGFR will be evaluated by the central laboratory, but local laboratory assessments will be performed in addition to the required central laboratory.

Only the central laboratory assessments will be used in the following analyses.

For the rate change in eGFR, the eGFR slope will be estimated from a mixed model, in which, the response variable will be the eGFR; stratum (ACEi, ARB, No RASi), region, and treatment (LCZ696, IMT) will be included as fixed-effect factors; time (in month) and treatment-by-time interaction will be included as covariates; a patient level random intercept and a patient level random slope (time) will be included as random effects with a unstructured covariance matrix (a common matrix for the two treatment groups). Here, the time (in month) will be computed as (date of assessment minus date of randomization plus one day)/(30.4375 day/month). The analysis will include all assessments from central laboratory at all visits up to Week 24 (Baseline to Week 24, scheduled or unscheduled) and will be performed based on the likelihood method with an assumption of MAR for missing data.

Based on the mixed model, the estimates and the two-sided 95% confidence intervals will be provided for the adjusted mean slope in each of the two treatment groups (LCZ696 and IMT), and for the adjusted mean difference (LCZ696 minus IMT).

Besides, the eGFR test values and the change from baseline in eGFR will be summarized by treatment group and visit, separately, for overall and for each stratum (ACEi, ARB, No RASi), using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum. The summary will include all scheduled assessments.

The FAS will be used for the above analyses.

2.8 Safety analyses

The SAF will be used for all safety analyses.

2.8.1 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.

The missing or partially missing start/end date for AEs will be imputed based on the Novartis AGB global standards.

Treatment emergent adverse events (TEAEs) 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.

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

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

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".

Most common TEAEs are defined as any recorded TEAE corresponding to a PT with at least 2% of patients in either treatment group (LCZ696 or IMT) having at least one TEAE of this PT.

SAEs are defined as any recorded AE with "Does AE meet the definition of an SAE" answered as "Yes".

The following rules are applicable to the summaries.

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

The following summaries will be performed.

- TEAEs will be summarized by treatment group, SOC and PT, and summarized by treatment group, SOC, PT and maximum severity, for overall and for each stratum (ACEi, ARB, No RASi).

- Study drug related TEAEs will be summarized by treatment group, SOC and PT, and summarized by treatment group, SOC, PT and maximum severity, for overall and for each stratum (ACEi, ARB, No RASi).
- TEAEs leading to permanent discontinuation of study treatment will be summarized by treatment group, SOC and PT, for overall and for each stratum (ACEi, ARB, No RASi).
- TEAEs leading to dose adjustment or temporarily interruption of study treatment will be summarized by treatment group, SOC and PT, for overall and for each stratum (ACEi, ARB, No RASi).
- Most common TEAEs will be summarized by treatment group, SOC and PT, for overall and for each stratum (ACEi, ARB, No RASi).
- Study drug related most common TEAEs will be summarized by treatment group, SOC and PT, for overall and for each stratum (ACEi, ARB, No RASi).
- Treatment emergent SAEs (TESAEs) will be summarized by treatment group, SOC and PT for overall and for each stratum (ACEi, ARB, No RASi).
- Most common TESAEs will be summarized by treatment group, SOC and PT, for overall and for each stratum (ACEi, ARB, No RASi).
- Study drug related TESAEs will be summarized by treatment group, SOC and PT, for overall and for each stratum (ACEi, ARB, No RASi).

The SAF will be used for the above analyses.

Table 2-9 shows the safety topics of interest for LCZ696 and the specifically specified analyses for each safety topic. In addition, the following standard analyses will be applied to all safety topics.

- Numbers and percentages of patients with any TEAE within the safety topic name (or PT within safety topic name) will be provided by treatment group, safety topic name, PT and maximum severity, for overall and for each stratum (ACEi, ARB, No RASi).
- Exposure adjusted incidence rates per 100 patient-years for TEAEs within the risk name will be provided by treatment group and safety topic name, for overall and for each stratum (ACEi, ARB, No RASi).
- Listing of patients numbers per safety topic.

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 the safety topics in Table 2-9 are store in the LCZ696 Case Retrieval Strategy (2019).

Table 2-9 Safety Topics of Interest

Safety Topic of Interest	Analysis
Anaphylaxis	Standard analyses;
Angioedema	Standard analyses;
Change in bone growth	Standard analyses;
Cognitive impairment (Broad SMQ)	Standard analyses;
Cognitive impairment (Narrow SMQ)	Standard analyses;
Embryo-fetal toxicity	Standard analyses;

Safety Topic of Interest	Analysis
Hepatotoxicity	1. Standard analyses; 2. Summary of the shift from baseline to minimum and maximum post-baseline category (Section 2.8.3.1) in alkaline phosphatase (ALP), alanine aminotransferase (ALT), aspartate aminotransferase (AST) and total bilirubin (TBL) 3. Numbers and percentages of patients with each of the following events will be provided by treatment group, for overall and for each stratum (ACEi, ARB, No RASi); a. (ALT > 3 × ULN) or (AST > 3 × ULN) b. (ALT > 5 × ULN) or (AST > 5 × ULN) c. (ALT > 8 × ULN) or (AST > 8 × ULN) d. (ALT > 10 × ULN) or (AST > 10 × ULN) e. (ALT > 20 × ULN) or (AST > 20 × ULN) f. [(ALT > 3 × ULN) or (AST > 3 × ULN)] and (TBL > 1.5 × ULN) g. [(ALT > 3 × ULN) or (AST > 3 × ULN)] and (TBL > 2 × ULN) h. [(ALT > 5 × ULN) or (AST > 5 × ULN)] and (TBL > 2 × ULN) i. [(ALT > 8 × ULN) or (AST > 8 × ULN)] and (TBL > 2 × ULN) j. [(ALT > 10 × ULN) or (AST > 10 × ULN)] and (TBL > 2 × ULN) k. [(ALT > 20 × ULN) or (AST > 20 × ULN)] and (TBL > 2 × ULN) l. (ALP > 2 × ULN) m. (ALP > 3 × ULN) n. (ALP > 5 × ULN) o. (TBL > 1.5 × ULN) p. (TBL > 2 × ULN) q. (TBL > 3 × ULN) r. (ALP > 3 × ULN) and (TBL > 2 × ULN) s. (ALP > 5 × ULN) and (TBL > 2 × ULN) t. [(ALT > 3 × ULN) or (AST > 3 × ULN)] and [nausea or vomiting or fatigue or general malaise or abdominal pain or (rash and eosinophilia)] u. [(ALT > 3 × ULN) or (AST > 3 × ULN)] and (TBL > 2 × ULN) and (ALP ≤ 2 × ULN) v. ([(ALT > 3 × ULN) or (AST > 3 × ULN)] and [(TBL > 2 × ULN) and (ALP ≤ 2 × ULN)]) or reported Hy's Law case w. (TBL > 3 × ULN) and [(AST ≤ 3 × ULN) or (ALT ≤ 3 × ULN)] and (ALP ≤ 1.5 × ULN) x. (ALP > 3 × ULN) and AST, ALT, TBL are within normal range
Hyperkalemia	1. Standard analyses;

Safety Topic of Interest	Analysis
	2. Numbers and percentages of patients with treatment emergent hyperkalemia leading to study treatment interruption will be provided by treatment group, for overall and for each stratum (ACEi, ARB, No RASi); 3. Numbers and percentages of patients with treatment emergent hyperkalemia leading to study treatment permanent discontinuation will be provided by treatment group, for overall and for each stratum (ACEi, ARB, No RASi); 4. Numbers and percentages of patients with any post-baseline serum potassium ≥5.5 mEq/L, >6.0 mEq/L and >6.5 mEq/L will be provided by treatment group, for overall and for each stratum (ACEi, ARB, No RASi);
Hypersensitivity (Broad SMQ)	Standard analyses;
Hypersensitivity (Narrow SMQ)	Standard analyses;
Hypotension	1. Standard analyses; 2. Blood pressures and change from baseline in blood pressures will be summarized by treatment group and visit, for overall and for each stratum (ACEi, ARB, No RASi), using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum; 3. Numbers and percentages of patients with the following events will be provided by treatment group, for overall and for each stratum (ACEi, ARB, No RASi); a. Post-baseline SBP < 100 mmHg; b. At least 30 mmHg drop in SBP (post-baseline assessments comparing to baseline); c. Simultaneous post-baseline SBP < 100 mmHg and at least 30 mmHg drop in SBP (post-baseline assessments comparing to baseline);
Malignancy	Standard analyses;
Neonatal toxicity through breast milk	Standard analyses;
Renal impairment (Broad SMQ)	1. Standard analyses; 2. For serum creatinine and eGFR, the test values and changes from baseline will be summarized by treatment group and visit, for overall and for each stratum (ACEi, ARB, No RASi), using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum; 3. Numbers and percentages of patients with each of the following events will be provided by treatment group, for overall and for each stratum (ACEi, ARB, No RASi); a. eGFR decline by >25%, >40%, >50%, and >30 mL/min/1.73 m^2 (post-baseline assessments comparing to baseline); b. Serum creatinine increase by >50% and >0.5 mg/dL (post-baseline assessments comparing to baseline); c. Post-baseline serum creatinine >2.0 mg/dL, >2.5 mg/dL, >3.0 mg/dL;

Safety Topic of Interest	Analysis
	4. Standard analyses within the subgroups: patients with baseline eGFR < 60 mL/min/1.73 m^2 and >= 60 mL/min/1.73 m^2;
Renal impairment (Narrow SMQ)	Standard analyses;
Rhabdomyolysis and Myoglobinuria	Standard analyses;
Statin drug drug interaction	Standard analyses;

The following subgroup analyses will be based on selected subgroup variables (Section 2.2.1).

- Treatment emergent SAEs will be summarized within each subgroup, by treatment group, SOC and PT.
- Numbers and percentages of patients with any TEAE within the safety topic name (or SOC/PT within safety topic name) for identified and potential safety topics will be provided within each subgroup, by treatment group, safety topic name, SOC, PT and maximum severity.

The SAF will be used for the subgroup analyses.

2.8.2 Death

Death and primary cause of death will be reported by the investigator.

The investigator reported death and primary cause of death will be recorded on the CRF page: "Death".

The investigator reported death and primary cause of death will be summarized by treatment group, for overall and for each stratum (ACEi, ARB, No RASi), using the number and percentage of patients who died, as well as the numbers and percentages of patients whose primary cause of death is in each category and each sub-category.

The SAF will be used for the above analyses.

2.8.3 Laboratory data

2.8.3.1 General laboratory data

Complete laboratory evaluations (hematology, biochemistry, and urine) outlined in Table 2-10 will be performed at screening, randomization, Week 12 and Week 24.

Abbreviated laboratory evaluations will be performed at Week 2, Week 4, Week 8, Week 16, and Week 20. Abbreviated laboratory includes: creatinine, potassium, eGFR, and blood urea nitrogen (BUN).

A central laboratory will be used for analysis of all specimens collected.

In addition to the required central laboratory assessments, a local laboratory may be used for the assessment of creatinine, potassium and eGFR during the up titration period. The results from the local laboratory during the up titration period will be allowed to be used for decision making regarding the eligibility of the patient to continue on in the study and will be recorded on the appropriate eCRF.

In addition, local laboratory assessments may be performed on an as-needed basis to monitor tolerability to study drug at unscheduled visits during the randomized treatment period and will be recorded in the appropriate eCRF.

Table 2-10 Complete laboratory evaluations

Hematology	Biochemistry	Urine measurements
Hematocrit	Alanine aminotransferase (ALT)	Urinalysis
Hemoglobin	Albumin (Alb)	
Platelet count	Alkaline phosphatase (ALP)	
Red blood cell count (RBC)	Aspartate aminotransferase (AST)	
White blood cell count (WBC)	Blood urea nitrogen (BUN)*	
WBC differential	Calcium	
Red blood cell distribution width (RDW)	Chloride	
Mean corpuscular volume (MCV)	Creatinine*	
Mean corpuscular hemoglobin concentration (MCHC)	eGFR*	
	Glucose	
	Hemoglobin A1C	
	Lipid profile (total cholesterol, LDL, HDL, and triglycerides)	
	Phosphate	
	Potassium*	
	Serum pregnancy test	
	Sodium	
	Total bilirubin (TBL)	
	Fractionated bilirubin (if total bilirubin > 2 x upper limit of normal range [ULN])	
	Total protein	
	Uric acid	

*Laboratory assessments for the abbreviated laboratory evaluation at visits where the complete laboratory evaluation is not performed

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 the laboratory assessments, if the test value is below the LLOQ, the test value will be imputed by 0.5 × LLOQ; if the test value is above the ULOQ, the test value will be imputed by 1.5 × ULOQ.

For each laboratory parameter, the baseline assessment is defined as the last non-missing central laboratory assessment (scheduled or unscheduled) prior to or on the first dose date of the study drug.

Test values and change from baseline in test values based on the central laboratory assessment will be summarized, separately, for overall and for each stratum (ACEi, ARB, No RASi), by laboratory parameter, treatment group and visit, using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum.

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, treatment group and visit, for overall and for each stratum (ACEi, ARB, No RASi), using the numbers and percentages of patients within each shift category. By-visit summaries will be conducted for central laboratory assessments and local laboratory assessments, separately. The baseline in summaries for local laboratory assessments will be taken from the central laboratory assessments.

For each laboratory parameter and each visit, the visit based minimum and maximum category are defined based on categorized test values at this visit according to Table 2-11. All central or local assessments will be included in the calculation.

The shift from baseline to visit based minimum and maximum category will be summarized separately, by laboratory parameter, treatment group and visit, for overall and for each stratum (ACEi, ARB, No RASi), using the numbers and percentages of patients within each category.

For each individual laboratory parameter, the minimum and maximum post-baseline category are defined based on categorized post-baseline test values according to Table 2-11. All central or local, scheduled or unscheduled assessments will be included in the calculation.

Table 2-11 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

Range of Categorized Post-Baseline Test Values	Minimum post-baseline category	Maximum post-baseline category
Missing	Missing	Missing

The shift from baseline to minimum and maximum post-baseline category will be summarized separately, by laboratory parameter and treatment group, for overall and for each stratum (ACEi, ARB, No RASi), using the numbers and percentages of patients within each category.

Table 2-12 provides criteria for clinically notable laboratory abnormality.

Table 2-12 Clinical notable criteria for laboratory values

Parameter	Criteria
Hematology	
Hematocrit	> 50% increase, > 20% decrease
Hemoglobin	> 50% increase, > 20% decrease
Platelet count	> 75% increase, > 50% decrease
RBC count	> 50% increase, > 20% decrease
WBC count	> 50% increase, > 50% decrease
Blood Chemistry	
Alkaline phosphatase	> 100% increase
ALT (SGPT)	> 150% increase
AST (SGOT)	> 150% increase
BUN	> 50% increase
Calcium	> 10% increase, > 10% decrease
Chloride	> 10% increase, > 10% decrease
Creatinine	> 50% increase
Potassium	> 20% increase, > 20% decrease
Total bilirubin	> 100% increase
Uric acid	> 50% increase

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

The clinically notable laboratory abnormality based on the local laboratory assessments and the central laboratory assessments will be summarized separately, by criteria and treatment group, for overall and for each stratum (ACEi, ARB, No RASi), using the numbers and percentages of patients satisfying such criteria.

The SAF will be used for the above analyses.

2.8.3.2 Special laboratory data

Liver function related laboratory data based on the local laboratory and the central laboratory assessments will be summarized separately, using the hepatotoxicity analyses in Table 2-9.

Renal function related laboratory data based on the local laboratory and the central laboratory assessments will be summarized separately, using the renal impairment analyses in Table 2-9.

The SAF will be used for the above analyses.

2.8.4 Other safety data

2.8.4.1 Electrocardiogram (ECG)

The ECG parameters (heart rate, RR interval, PR interval, QRS duration, QT interval, QTcF interval) will be recorded on the CRF page: "12 Lead ECG Evaluation - Local Analysis".

The test value and the change from baseline in ECG parameters, will be summarized by ECG parameter, treatment group and visit, for overall and for each stratum (ACEi, ARB, No RASi), using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum.

In addition, any of the following clinically significant abnormalities, will also be recorded on the same CRF page: "12 Lead ECG Evaluation - Local Analysis".

- atrial fibrillation
- atrial flutter
- LBB block
- RBB block
- pathological Q waves
- left ventricular hypertrophy
- paced rhythm

The clinically significant abnormalities will be summarized by treatment group and visit, for overall and for each stratum (ACEi, ARB, No RASi), using the numbers and percentages of patients with each abnormality and with each newly onset abnormality, where the newly onset abnormality is defined as an abnormality not present at baseline but present at the visit

The SAF will be used for the above analyses.

2.8.4.2 Vital signs

The test values and the changes from baseline in vital signs (sitting pulse, sitting SBP, sitting DBP, weight, waist circumference and hip circumference) will be summarized by parameter, treatment group and visit, for overall and for each stratum (ACEi, ARB, No RASi), using number of observations, mean, standard deviation, median, minimum, Q1, Q3, and maximum. By-visit summary will only include scheduled assessments.

Table 2-13 provides criteria for clinically notable vital sign.

The clinically notable vital signs will be summarized by criteria and treatment group, for overall and for each stratum (ACEi, ARB, No RASi), using the numbers and percentages of patients with at least one post-baseline assessment satisfying such criteria.

Table 2-13 Criteria for clinically notable vital signs

Vital Sign (unit)	Clinically notable criteria
Weight (kg)	• decrease > 7% from baseline • increase > 7% from baseline
Sitting pulse (bpm)	• < 50 bpm and decrease from baseline of > 15 bpm

Vital Sign (unit)	Clinically notable criteria
Sitting SBP (mmHg)	• > 120 bpm and increase from baseline of >15 bpm • < 90 mmHg and decrease from baseline of > 20 mmHg • > 180 mmHg and increase from baseline of > 20 mmHg
Sitting DBP (mmHg)	• < 50 mmHg and decrease from baseline of > 15 mmHg • > 105 mmHg and increase from baseline of > 15 mmHg

The SAF will be used for the above analyses.

2.9 Pharmacokinetic endpoints

Not applicable.

2.10 PD and PK/PD analyses

Not applicable.

2.11 Patient-reported outcomes

See Section 2.6.1 and Section 2.6.3.

2.12 Biomarkers

Plasma NT-proBNP will be analyzed according to Section 2.5 and Section 2.7.1.

2.13 Other Exploratory analyses

Not applicable.

2.14 Interim analysis

No formal interim efficacy analysis is planned.

An established program level DMC independent of Novartis that reviews safety data from LCZ696 studies will be used for this study. The DMC will review SAEs in an unblinded manner on a regular basis and determine if it is safe to continue the study. Any major recommendation from the DMC will be communicated to the Executive Committee and must be reviewed and ratified by the Executive Committee in consultation with Novartis prior to its enactment.

3 Sample size calculation

The sample size (2500 randomized patients) has been chosen in order to provide adequate power to evaluate both primary and secondary endpoints.

With an alpha of one-sided 0.0225, the power for the NT-proBNP will range from 92% to more than 99%, to detect a relative reduction ranging from 11% to 24% in change from baseline to Week 12 in NT-proBNP, assuming a standard deviation of 0.81 for change from baseline in the log-transformed NT-proBNP (based on PARAMOUNT-HF data of patients with baseline KCCQ CSS < 75) and an overall drop-out rate of 10%.

With an alpha of one-sided 0.0025, the power for the 6MWD will range from 90% to 99%, to detect a mean difference ranging from 22 meters to 30 meters in change from baseline to Week 24 in 6MWD, assuming a standard deviation of 120 meters (Ingle et al 2005), an overall drop-out rate of 10%, and an overall proportion of 88% for patients with B6MWD ranging from 100 meters to 450 meters.

With an alpha of one-sided 0.0225, the power for the KCCQ CSS will range from 87% to 99%, to detect a mean difference ranging from 2 points to 3 points in change from baseline to Week 24 in KCCQ CSS, assuming a standard deviation of 15.52 points for change from baseline in the KCCQ CSS (based on PARAMOUNT-HF data of patients with baseline KCCQ CSS < 75) and an overall drop-out rate of 5%.

All alpha levels in the following sample size calculations will be one-sided 0.025.

This sample size of 2500 randomized patients aims to provide a power of approximately 90%, with the one-sided alpha level of 0.025, to demonstrate a 5% responder advantage in at least 5-point deterioration for patients who are treated with LCZ696 over IMT when assuming the proportion of patients with at least 5-point deterioration from baseline to Week 24 in the IMT group is 20% (based on PARAMOUNT-HF data of patients with baseline KCCQ CSS < 75) and an overall 5% dropouts.

With the same assumption of the dropout rate, this sample size will also provide a power of 69% to detect a 5% responder advantage in at least 5-points improvement in the KCCQ clinical summary score for patients who are treated with LCZ696 over IMT when assuming the proportion of patients with at least 5-points improvement from baseline to Week 24 in the IMT group is 55% (based on PARAMOUNT-HF data of patients with baseline KCCQ CSS < 75).

The sensitivity of power for at least 5-points deterioration in the KCCQ clinical summary score to changes in assumptions is outlined in Table 3-1.

Table 3-1 Sensitivity of power for at least 5-points deterioration in KCCQ CSS to changes in assumptions for N = 2500

Assumed probability of at least 5-points deterioration			Power for at least 5-points deterioration (one-sided alpha = 0.025)			
IMT	LCZ696	Difference (LCZ696 - IMT)	With 5% drop-out rate	With 10% drop-out rate	With 15% drop-out rate	With 20% drop-out rate
20%	15%	-5%	90%	88%	86%	84%
35%	30%	-5%	74%	72%	69%	67%

The sensitivity of power for at least 5-points improvement in the KCCQ CSS to changes is outlined in Table 3-2.

Table 3-2 Sensitivity of power for at least 5-points improvement in KCCQ CSS to changes in assumptions for N = 2500

Assumed probability of at least 5-points improvement			Power for at least 5-points improvement (one-sided alpha = 0.025)			
IMT	LCZ696	Difference (LCZ696 - IMT)	With 5% drop-out rate	With 10% drop-out rate	With 15% drop-out rate	With 20% drop-out rate
55%	60%	5%	69%	67%	65%	62%
41%	46%	5%	69%	67%	64%	62%

The assumed dropout rates (Table 3-1 and Table 3-2) are for the purpose of power sensitivity analysis. During the conduct of the study, all efforts will be made to reduce missing data.

For other secondary endpoints, this sample size will provide adequate powers for reasonable effect size assumptions: power of 90% for NYHA class change, assuming 14% of the patients in the comparator and 19% of the patients in the LCZ696 group are improved in NYHA class change, 6% of the patients in the comparator and 5% of the patients in the LCZ696 group are worsened in NYHA class change (based on Week 24 PARAMOUNT-HF data of patients with baseline KCCQ CSS < 75); and power of approximately 90% to detect a between treatment group mean difference of 1.4 points or more in change from baseline in SF-36 physical component summary score, assuming a standard deviation of 10-points (Edelmann et al 2013).

East 6.3 was used for the sample size calculations.

4 Change to protocol specified analyses

No change from protocol.

5 Appendix

5.1 Imputation rules

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

5.2 Rule of exclusion criteria of analysis sets

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 in Table 5-1, which may be updated prospectively and will be finalized before database lock.

Table 5-1 Protocol deviation criteria leading to exclusion from the analysis sets

Protocol Deviation		Excluding from Analysis Set				
ID	Description	SCR	RAN	FAS	PPS	SAF
INCL01A	No study informed consent obtained	X	X	X	X	X

	Protocol Deviation	Excluding from Analysis Set		
INCL03A	Patient's qualifying LVEF <= 40% or the qualifying echo does not support LVEF >40%		X	
INCL04	Heart failure symptoms not requiring diuretic treatment for at least 30 days prior to Visit 1		X	
INCL05	NYHA Class I at Visit 1 or no heart failure symptoms at Visit 1		X	
INCL08A	NT-proBNP at Visit 1 <= 220 pg/mL and no atrial fibrillation/atrial flutter (AF) on the Visit 1 ECG		X	
INCL08B	NT-proBNP at Visit 1 <= 600 pg/mL and atrial fibrillation/atrial flutter (AF) on the Visit 1 ECG		X	
INCL09	KCCQ clinical summary score >= 75 at Visit 1		X	
EXCL08	Walking distance primarily limited by non-cardiac comorbid conditions		X	
TRT03	Patient was randomized in error and no double-blind study medication was taken	X	X	X
TRT05	Patient was randomized within the wrong stratum and no double-blind study medication was taken	X	X	X
OTH01	Major GCP violation at site	X	X	
OTH02	Blind was broken accidently at the site		X	

Table 5-2 presents patient classification criteria including protocol deviation criteria and non-protocol-deviation criteria leading to exclusion from analysis sets.

Table 5-2 Patient classification

Analysis Set	Protocol Deviation ID leading to exclusion of patients	Non-protocol deviation criteria leading to exclusion from analysis sets
SCR	INCL01A	Not applicable
RAN	INCL01A	• Patient without a randomization number
FAS	INCL01A, TRT03, TRT05, OTH01	• Patient without a randomization number
PPS	INCL01A, INCL03A, INCL04, INCL05, INCL08A, INCL08B, INCL09, EXCL08, TRT03, TRT05, OTH01, OTH02	• Patient without a randomization number • Patient randomized but not receiving any study drug
SAF	INCL01A, TRT03, TRT05	• Patient without a randomization number • Patient randomized but not receiving any study drug

6 Reference

Bretz F, Maurer W, Brannath W, et al (2009) A graphical approach to sequentially rejective multiple test procedures. Statistics in Medicine; 28(4): 586-604.

Carpenter JR. and Kenward, MG (2013) Multiple imputation and its application, New York: John Wiley &Sons.

Ingle, L, Rhidian J, Shelton AS. Rigby, et al (2005) The reproducibility and sensitivity of the 6-min walk test in elderly patients with chronic heart failure. Eur Heart J; 26(17): 1742-1751.

Edelmann F, Wachter R, Schmidt AG, et al (2013) Effect of spironolactone on diastolic function and exercise capacity in patients with heart failure with preserved ejection fraction: the Aldo-DHF randomized controlled trial. JAMA; 309(8):781-91.

National Research Council (2010): The Prevention and Treatment of Missing Data in Clinical Trials. Panel on Handling Missing Data in Clinical Trials. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. The National Academies Press, Washington, DC.

Novartis ADaM Governance Board (AGB) global standard

Siddiqui O, Hung HM, O'Neill R (2009). MMRM vs. LOCF: a comprehensive comparison based on simulation study and 25 NDA datasets. J Biopharm Stat;19(2):227-46.

The Kansas City Cardiomyopathy Questionnaire Scoring Instructions

Ware, J. E., Kosinski, M., Bjorner, J. B., Turner-Bowker, D. M., Gandek, B., & Maruish, M. E. (2008): *User's manual for the SF-36v2 Health Survey*. Quality Metric.