NOVARTIS

Clinical Development

RLX030A

CRLX030A2301 / NCT01870778

A multicenter, randomized, double-blind, placebo-controlled phase III study to evaluate the efficacy, safety and tolerability of Serelaxin when added to standard therapy in acute heart failure patients

RAP Module 3 – Detailed Statistical Methodology Amendment 3

Author: [redacted]

Document type: RAP Documentation

Document status: Final version 1.0

Release date: 03-Mar-2017

Number of pages: 36

Table of contents

List of tables

Document History – Changes compared to previous version of RAP module 3.

Version	Date	Changes
Amendment 1	13-Aug-2015	-add the testing strategy for WHF and secondary endpoints in case of early stopping -add sensitivity analysis for WHF using logistic regression -add substudy population definition
Amendment 2	06-Feb-2017	- Clarify the method of mapping discharge visit to scheduled visits if scheduled visits are missing - Update the list of countries that recruited subjects (Table 2-3) - Remove subgroup category of lymphocytes and education level in Table 2-4 - Use FAS in section 2.1.9 – study medication summary. - Align presentation of non-IV concomitant medications with RELAX-AHF - Clarify definition of the censoring date for death endpoints (Adjudicated CV death, adjudicated all cause death) for patients who withdrew consent in countries where vital status data is publically available, in accordance with local laws and regulations - Clarify use of discharge date as the reference date for endpoint time to the composite endpoint of CV death or rehospitalization due to HF/RF during a follow-up period of 30 days post-discharge - Clarify the statistical method used for subgroup analysis of length of stay (LOS). Clarify the imputation rule for missing discharge information to align with RELAX-AHF - Clarify the imputation of signs and symptoms through Day 5 for subjects with WHF or died at hospital to align with RELAX-AHF - Addition of two new tables per ClinicalTrials.gov and EudraCT requirement -Update Safety risk and event of interest search specification (Table 2-7) - Clarify the ECG analysis population - Update the ECG categorical analysis, add analysis of newly occurring abnormalities in ECG morphology and remove ECG shift table to enable comprehensive assessment - Update liver function categorical analysis (Table 2-8)
Amendment 3	03-Mar-2017	- Add a new protocol deviation " Site closed down for GCP reasons" leading to exclusion from the FAS and PPS -Subjects without valid written informed consent will be excluded from all analysis populations. - Remove covariate lymphocytes (%) from covariate-adjusted sensitivity analysis of CV death because of a number of missing values at baseline

List of abbreviations

ACE	Angiotensin Converting Enzyme
ACM	All-cause mortality
AE	Adverse Event
AHF	Acute Heart Failure
ALP	Alanine Phosphatase
ALT	Alanine Aminotransferase
ARB	Angiotensin-II Receptor Blocker
AST	Aspartate Aminotransferase
ATC	Anatomical Therapeutic Chemical
BNP	Brain Natriuretic Peptide
BPDE	Blood pressure decrease event
BUN	Blood Urea Nitrogen
CBPDE	Confirmed blood pressure decrease event
CCU	Coronary care unit
CEC	Clinical Endpoints Committee
CI	Confidence interval
CPK	Creatine Phosphokinase
CRF	Case Report Form
CSR	Clinical Study Report
[redacted]	[redacted]
CV	Cardiovascular
CVD	CV death
CVDR	CV death or rehospitalization due to heart failure or renal failure
CWHF	Clinical Worsening Heart Failure
dL	Deciliter
DMC	Data Monitoring Committee
DRE	Disease-related event
ECG	Electrocardiogram
EF	Ejection fraction
eGFR	Estimated Glomerular Filtration Rate
FAS	Full Analysis Set
GCP	Good Clinical Practice
HF	Heart failure
ICU	Intensive Care Unit
ITT	Intent-to-treat
IV	Intravenous
kg	Kilogram
LFT	Liver Function Toxicity
LOS	Length of hospital stay
LPLV	Last Patient Last Visit

LSM	Least-square Mean
MAP	Master Analysis Plan
MedDRA	Medical Dictionary for Regulatory Activities
MRA	Mineralocorticoid receptor antagonist
mg	Milligram
min	Minute
mL	Milliliter
mmHg	Millimeters mercury
NYHA	New York Heart Association
OF	O'Brien-Fleming
PASS	Physicians Assessment of Signs and Symptoms
PD	Protocol Deviation
PPS	Per Protocol Set
PT	Preferred Term
RAN	Randomized Set
RAP	Report and Analysis Preparation
RBC	Red Blood Cell
SAE	Serious adverse event
SAF	Safety Set
SAS	Statistical Analysis System
SBP	Systolic Blood Pressure
SCR	Screened Set
SGOT	Glutamic Oxaloacetic Transaminase
SGPT	Glutamic Pyruvic Transaminase
sMDRD	Simplified Modification of Diet in Renal Disease
SMQ	Standardized MedDRA Query
SOC	System Organ Class
TBL	Total Bilirubin
ULN	Upper Limit of Normal
WHF	Worsening heart failure
WHO-DD	World Health Organization Drug Dictionary
µg	Microgram

1 Introduction

The CRLX030A2301 study (also referred to as RELAX-AHF-2) aims to evaluate the efficacy, safety and tolerability of intravenous infusion of 30 µg/kg/day serelaxin for 48 hours, when added to standard therapy. Data from this study is intended to replicate the reduction in mortality in AHF patients observed in the RELAX-AHF study.

The purpose of this document is to outline the planned analyses to be completed to support the completion of the Clinical Study Report (CSR) for RELAX AHF-2.

2 Statistical methods planned in the protocol and determination of sample size

Data will be analyzed according to the data analysis section 10 of the study protocol which is available in Appendix 16.1.1 of the CSR.

All nominal time points will be used for analysis; actual assessment times will not be used to reclassify the time point at which a measure was taken. Time-to-event analyses will be based on actual dates (time) reported.

All references to index hospitalization refer to the initial hospitalization for AHF.

Unless specified, descriptive statistics indicates mean, median, standard deviation, minimum and maximum will be presented for continuous variables; the number and percentage of subjects in each category will be presented for categorical variables for each treatment group and all subjects (total).

2.1 Subject and treatment

2.1.1 Analysis population

The following analysis populations will be defined for statistical analysis:

- **Screened set (SCR)** – All subjects who signed the informed consent. The SCR includes only *unique screened subjects*.
- **Randomized set (RAN) -** All subjects who received a randomization number, regardless of receiving trial medication.
- **Safety set (SAF) -** All subjects who received any amount of study drug and have at least one post-baseline safety assessment. Of note, the statement that a subject had no adverse events also constitutes a safety assessment. Subjects will be analyzed according to treatment received. Subjects who received any amount of serelaxin will be included in the serelaxin treatment group.
- **Full analysis set (FAS) -** All subjects in the RAN population who were not misrandomized subjects*. Following the intent-to-treat (ITT) principle, subjects are analyzed according to the treatment they have been assigned to at the randomization. Further exclusions from the FAS were 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.

* Misrandomized subjects are those who have not been qualified for randomization, have been inadvertently randomized into the study and who did not take study drug. Misrandomized subjects are defined as cases where IRT calls were made by the site either prematurely or inappropriately prior to confirmation of the subject's final randomization eligibility and double-blind medication was not administered to the subject. These subjects should subsequently be discontinued from the study.

- **Per protocol set (PPS)** – all subjects in the FAS who received any amount of study medication and have no major protocol deviations.

Subjects without valid written informed consent will be excluded from all analysis populations.

Laboratory substudy and electrocardiogram (ECG) substudy populations will be defined for subjects participating laboratory or ECG substudy with informed consent.

2.1.2 Protocol deviation (PD)

All cases of prospectively defined protocol deviations (PD) will be identified prior to clinical database lock/unblinding and entered into a dedicated data panel as part of the locked database. Certain deviations may stipulate that data collected from the subjects will be excluded from PPS. This PPS is used to assess robustness of the primary and key secondary analysis results.

All exceptional cases, problems and the final decisions on the allocation of subjects to populations will be fully defined and documented before data base lock (in particular before breaking the blind where applicable) and will be fully identified and summarized in the clinical study report.

Criteria defining protocol deviations are referenced in the "PD specs and edit check' tab of Edit Check Specification document. Protocol deviations will be classified into 5 categories as appropriate:

- Selection criteria not met
- Subject not withdrawn as per protocol
- Treatment deviation
- Prohibited concomitant medication
- Other

The major protocol deviations resulting in exclusion of subjects from an analysis population are outlined in the table below:

Table 2-1 Protocol deviations resulting in exclusion of subjects from analysis populations

		Excluding from analysis set			
PD category	**Deviation**	**RAN**	**FAS**	**SAF**	**PPS**

		Excluding from analysis set			
PD category	**Deviation**	**RAN**	**FAS**	**SAF**	**PPS**
Selection criteria not met	Missing informed consent (Inclusion criteria #1)	x	x	x	x
	Patient hospitalized for reasons of non-AHF event (inclusion criteria #3)				x
	Systolic BP <125 mmHg at the start and/or at the end of screening (inclusion criteria #4)				x
	Time from presentation to randomization >16 hours (inclusion criteria #5)				x
	Dose of IV furosemide between presentation and start of screening is < 40 mg (inclusion criteria #6)				x
	eGFR is <25 or > 75 mL/min/1.73m2 at screening (inclusion criteria #7)				x
	Dyspnea is primarily due to non-cardiac causes (exclusion criteria #1)				x
	Sepsis or active infection requiring IV anti-microbial treatment reported at screening (exclusion criteria #4)				x
	AHF is due to significant arrhythmias (exclusion criteria #6)				x
	Pre-randomization eGFR < 25 mL/min/1.73m2 or ongoing or planned dialysis or ultrafiltration at screening (exclusion criteria #7)				x
Treatment derivation	Serelaxin infusion administered in diluent other than the recommended 5% dextrose bags				x
	Wrong study drug kit administered or wrong dose administered				x
	Subject misrandomized		x	x	x
Other	Site closed down for GCP reasons		x		x

In addition, deviations not defined in the protocol resulting in exclusion of subjects from an analysis population are outlined in the table below.

Table 2-2 Non-protocol deviation criteria resulting in exclusion of subjects from analysis populations

	Excluding from analysis set			
Criterion	**RAN**	**FAS**	**SAF**	**PPS**
Subject randomized but not receiving any dose of double-blind study medication			x	x

2.1.3 Visit, baseline and post baseline definitions

Except for biomarkers (described in Section 2.2.7.3), discharge visit will be mapped to Day 1, 2, 3, 4, 5, 14 (with +/- 3 day window), 60 (with +/- 7 days window), 120 (with +/- 7 days window), 180 (with +/- 7 day window) if the scheduled visits are missing.

Screening phase: The screening phase is defined as the period prior to randomization. After obtaining written informed consent, the subject will be evaluated for eligibility to participate in to the study.

Baseline definition for efficacy/safety measurements: Baseline is defined as Hour 0/Day 1 assessment. The baseline date and time will be considered the date and time of initiation of study drug, or randomization if the subject was not treated with study drug. If missing, the value at screening visit will be used.

Post-baseline phase for efficacy/safety measurements: The post-baseline phase begins at the time when study drug administration starts (Hour 0/Day 1) and ends with the last study assessment/last follow-up or the death of subject. For subjects who did not receive study drug, the phase begins at time of randomization.

2.1.4 Subgroup definitions

If not otherwise specified, the following subgroups will be defined for the study using the data collected at screening.

The key subgroups will be used in primary and key secondary endpoints.

Table 2-3 Specification of key subgroups at screening

Subgroup	Method of Derivation
Age groups (<65, ≥65; <75, ≥75 years)	Derived
Gender	eCRF
Race (White/Caucasian, Other)	Derived
Region *	Derived
Ejection fractions (<50%, ≥50%;<40%,≥40%)	Derived
Prior history of heart failure (yes/no)	eCRF

* **North America:** USA, Canada

Latin America (including Central America): Argentina, Brazil, Chile, Colombia, Peru, Mexico

Western Europe: Austria, Belgium, Denmark, France, Germany, Greece, Ireland, Italy, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, UK, Turkey, Israel

Of note, Turkey and Israel are classified in Western Europe based on medical practice.

Eastern Europe: Bulgaria, Czech Republic, Hungary, Poland, Slovakia, Romania, Russia

Other: Australia, South Africa

Additional exploratory subgroups analysis (as defined in Table 2-4, based on Metra et al 2013of primary and key secondary time to event endpoints will be presented in Appendix 16.1.9 only.

Table 2-4 Specification of additional subgroups at screening

Subgroup	Method of Derivation
NT-proBNP ≥ 3000 pg/ml or BNP ≥ 750 pg/ml	Local lab, Derived
Systolic blood pressure (<140, ≥140 mmHg)	Derived
Heart rate (< 80, ≥ 80 bpm)	Derived
History of Ischemic heart disease (yes/no)	eCRF
History of diabetic mellitus (yes/no)	eCRF
History of atrial fibrillation or flutter (yes/no)	eCRF

Time from presentation to randomization (<2, 2-10, >10 hours)	Derived
Intravenous (IV) nitrate use at randomization (yes/no)	eCRF
eGFR (<30, ≥30; <50, ≥50; <60, ≥60 mL/min/1.73 m^2)	Local Lab, Derived

2.1.5 Subject disposition

The number of subjects randomized and included in the full analysis set (FAS) will be presented by treatment group. The number of subjects screened only (SCR) will be presented. In addition, the reasons for screen failures will be provided. The number and percentage of subjects in the FAS who completed the study, who discontinued the study and the reason for discontinuation will be presented for each treatment group and all subjects.

A completer will be defined as either has died with adjudicated cause (including those adjudicated as having unknown cause), or whose actual follow up is greater or equal to 165 days (last known vital status – baseline date).

The frequency (%) of subjects with protocol deviations as well as the criteria leading to exclusion from analysis sets will be presented for RAN.

The number of subjects enrolled and randomized per region and country will be presented for FAS.

For each subject, the length of follow-up (days) will be computed by subtracting the study drug initiation date (or the randomization date for subjects not treated with study drug) from the end of study date (the last known vital status).

Descriptive statistics will be presented for length of follow-up by treatment group and overall. Categorical summaries will also be presented for the number and percent of subjects within predefined follow-up intervals as: None; >0-5 days; 6-14 days; 15-60 days; 61-120 days; 121-180 days.

2.1.6 Background and demographic characteristics

Subject demographics (age, sex, ethnicity, race, weight (continuous and <115kg, ≥ 115 kg) and height) and baseline characteristics (including alcohol history and smoking history) will be summarized for the FAS.

For baseline comparability analysis, categorical variables will be analyzed using the Chi-square test and continuous variables will be analyzed using the t-test and will be presented in Appendix 16.1.9 only.

The p-values will be provided for descriptive purpose and will not be considered to define any formal basis for determining factors to be included in statistical models. If an imbalance of treatment groups with respect to some variables does occur, supplemental analyses with addition of these variables in model may be performed to assess the potential impact on efficacy as appropriate.

The background and demographic characteristics analysis will be repeated for age, gender, race, region, PEF/REF and history of HF subgroups as defined in Table 2-2 for the FAS. The results will be presented as supplementary analysis in the statistical Appendix section.

Baseline characteristics of interest in vital signs and local laboratory values collected at screening visit will also be presented for the FAS.

2.1.7 Medical history

The number and percentage of subjects including but not limited to who had history of heart failure, New York Heart Association (NYHA) class, primary heart failure etiology (ischemic/non-ischemic), who were hospitalized for heart failure, ejection fraction (<40%, ≥40%; <50%, ≥50%), who were on heart transplant list, and who had history of diabetes mellitus will be presented for FAS. The number of HF hospitalization in the past year and ejection fraction (%) will also be presented.

The number and percentage of subjects will also be presented for each protocol solicited event as pre-printed list in eCRFs for the FAS.

Any other condition entered on the medical history not recorded on the medical history protocol solicited events or heart failure history eCRFs will be coded using the MedDRA dictionary. The number and percentage of subjects with each medical condition will be provided by treatment group for the FAS.

2.1.8 Other subject history information

The number and percentage of subject in lifestyle history, highest level of education and home health care arrangement, will be presented for the FAS.

The history drug abuse will also be presented for the FAS.

2.1.9 Study medication

For the FAS population, study drug administration details will be summarized by treatment group and overall for the time from presentation to hospital (hours), time from first IV loop diuretic for the current AHF episode to randomization (hours), time from earlier of presentation to hospital or first IV loop diuretic to randomization (hours), time from randomization to study drug administration (hours), on IV nitrates at randomization (yes/no), dose decreased due to blood pressure decrease (yes/no), study drug administered on either day (yes/no), reason study drug not administered, actual study drug received, and the number of days infused (one or two days), study drug prematurely discontinued and reasons for discontinuation.

The duration of study drug administration (in hours) and the total volume of study drug administered (estimated from infusion duration at each infusion rate) will also be summarized by treatment group and overall.

2.1.10 Concomitant medication

The use of all concomitant medications will be summarized using the SAF population.

2.1.10.1 IV loop diuretic medications

IV loop diuretic medications include the following:

- Bumetanide

- Etacrynic acid/Ethacrynic acid
- Furosemide
- Torasemide/Torsemide

The number and percentage of subjects using IV loop diuretic medications will be summarized for each of the time periods: Hour 24/Day 1 to end of study, Hour 0/Day 1, 24 hours/Day 1, 48 hours/Day 2, Day 3, Day 4, Day 5, and Discharge. Each time point reflects medication use within the preceding 24 hours.

The average total daily dose for each medication will also be presented at reported time point. In addition, cumulative use reported at 48 Hours (24 hours/Day 1 and 48 hours/Day 2), and at Day 5 (from 24 hours/Day 1 to Day 5) in furosemide equivalent (mg) will be summarized. The calculation of furosemide equivalents (mg) for Bumetanide, Ethacrynic acid/Ethacrynic acid and Torsemide/Torsemide acid are actual dose (mg) multiplied by constant 40, 0.8, or 2, respectively. For subjects who die before Day 5, the minimum of 2 times of the last recorded dose and 160 mg will be assumed for every day from the date of death up to and including Day 5.

2.1.10.2 IV medications other than loop diuretics

The number of subjects using other IV medications including Dopamine, Dobutamine, Milirinone, Enoximone, Norepinephrine, Epinephrine, Nesiritide, Levosimendan, Isosorbide Dinitrate, Nitroglycerin, Nitroprusside, Phenylephrine and other will be presented in the same way as IV loop diuretic medications.

The average dose over the time administered of the medication will be presented separately for each medication at each reported time point. No equivalent dose unit among other IV medications will be converted.

2.1.10.3 Other non-IV concomitant medications

The use of other non-IV concomitant medications is reported through Day 180. The number and percentage of subjects using other non-IV medications will be summarized for each of the time periods: 30 days to 14 prior to study drug initiation, Hour 24/Day 1 to end of study, Hour 0/Day 1, 24 hours/Day 1, 48 hours/Day 2, Day 3, Day 4, Day 5, Discharge, Day 14, Day 60, Day 120 and Day 180 for ACE inhibitor (ACEi), angiotensin inhibitor (ARB), ACEi or ARB, beta blocker, PO loop diuretics, other PO diuretics, calcium channel blockers, aldosterone antagonist and cardiac glycosides.

In addition, the number (%) of subjects who have these medications initiated during their index hospitalization and subjects who have these medications increased during the period following discharge from the index hospitalization will also be presented.

The average daily dose administered will be summarized at each time point for each medication within the loop diuretics, ACEi, ARB, beta blocker, aldosterone antagonist, ARNi and direct renin inhibitor. .

Other non-IV concomitant medications taken prior to, and after the start of study drug will be presented by the ATC and preferred terms.

2.1.11 Close-out procedure

The number of CV deaths will be predicted on an ongoing basis. [redacted]

Subjects who have been randomized to the study will be followed up until their final visit (180 days) or date of death. Attempts should be made to contact any subjects who did not complete the Day 180 visit.

The aim of the close-out procedure will be to allow for all necessary activities (study completion visits, data cleaning, raising and resolution of all queries to the clinical sites, etc.) including the adjudication of all primary and secondary events that occurred during the study and prior to final visit. The "adjudicated endpoint" refers to an endpoint and date of the endpoint that has been confirmed by an independent endpoint committee who has no access to the treatment code.

Only adjudicated events will be counted in the primary analyses. However, sensitivity analyses will be performed on the suspected (investigator reported) primary and secondary endpoints.

2.2 Efficacy evaluation

The primary analysis and the analyses of all secondary efficacy endpoints will be based on the FAS population following the intention-to-treat principle.

2.2.1 Efficacy variables

The following table gives an overview of efficacy variables.

Table 2-5 Primary, secondary and exploratory variables

Variables		Analysis set	Analysis method
Primary efficacy variables			
1	Time to CV death through Day 180	FAS/PPS	log rank test and Kaplan-Meier estimates
2	Time to worsening heart failure (WHF) through Day 5 (considering death in the 5-day period as WHF event)	FAS/PPS	Gehan's generalized Wilcoxon test
Secondary efficacy variables (if either primary variable is significant, follow sequential testing procedure)			
1	Time to all-cause death through Day 180	FAS/PPS	log rank test and Kaplan-Meier estimates (of note: will not be formally incorporated in a multiple testing procedure)
2	Length of total hospital stay (LOS) during the index AHF hospitalization	FAS/PPS	Wilcoxon rank-sum test
3	Time to first occurrence of the composite endpoint of CV death or rehospitalization due to heart failure (HF)/renal failure (RF) through Day 180	FAS/PPS	log rank test and Kaplan-Meier estimates
Other secondary efficacy endpoints (at α=0.05, no multiplicity adjustment)			

Variables		Analysis set	Analysis method
1	Length of ICU and/or CCU stay for the index AHF hospitalization	FAS	Wilcoxon rank-sum test
2	Change from baseline in congestive signs and symptoms of HF through Day 5	FAS	Wilcoxon rank-sum test at each post-baseline visit
3	Change from baseline in selected biomarkers from baseline through Day 14 in a subset of randomized subjects	FAS	Repeated measurement model
[redacted]			

Primary and key secondary endpoints are considered in sequential testing procedure and efficacy interim analysis

For the time to **confirmed** CV death during the follow-up period of 180 days, the confirmation of primary events is based on an adjudication process through an independent CEC (Clinical Endpoint Committee). The CEC will also adjudicate all recurrent events of the unplanned hospitalizations according to the protocol.

2.2.2 Definition of time to CV death through Day 180

Events that occur prior to or at the subject's final visit date will be part of the time-to-event analyses. Events that occurred more than 180 days will be considered as censored from the analyses. The CV death will be based on adjudicated results. If the adjudicated cause of death is unknown, it will be considered as a CV death.

The time-to-event for a given subject will be calculated as the difference between the subject's date of endpoint and the randomization date:

- Time-to-event (days) = date of CV death – randomization date,

For those subjects without events, the censoring date will be the earliest in the following dates:

- **180 days**
- **Date subject withdrew consent to all follow-up (if no subsequent vital status documented)**

Note: in accordance with local laws and regulations, vital status may be publically available

- **Date of non-CV death**
- **Date of subject's last visit or contact or vital status**

The censoring time for the analyses will be calculated as:

- Time-to-event censored (days) = censoring date – randomization date

Whenever the event date or visit date needed for defining the time-to-event or censoring date is missing or incomplete for a given subject, an imputation rule for missing/incomplete date will be applied. More details about the imputation rule are presented in Section 2.2.9.

2.2.3 Time to worsening heart failure (WHF) through Day 5 (considering death in the 5-day period as WHF event)

WHF events occurring within 120 hours of randomization will be included in the analysis. WHF is defined as clinical worsening of HF during the index hospitalization or as adjudicated rehospitalization for HF. A subject who dies by 120 hours without a WHF event will be considered to have had an event at the time of death.

Event date and time will be used at calculating the time to event for better precision, and convert it back to days.

Time-to-event (days) = (first event date/time– randomization date/time),

For those subjects without events the censoring date will be the earliest in the following dates:

- **Hour 120 (5 days)**
- **Date subject withdrew consent to all follow-up**
- **Date of subject's last visit or contact**

2.2.4 Statistical hypothesis, model, and method of analysis

CV death

The statistical hypothesis is:

H_0: $\lambda_2/\lambda_1 \geq 1$, i.e., the rate of primary event of CV death is greater or equal in the serelaxin group relative to the placebo group *versus the one-sided alternative* H_A: $\lambda_2/\lambda_1 < 1$, i.e., the rate of CV death is smaller in the serelaxin group relative to the placebo group, where λ_1 and λ_2 are the hazard rates for CV death in the placebo group and serelaxin group, respectively. The ratio λ_2/λ_1 is also called the hazard ratio of serelaxin to placebo.

The hypothesis will be tested based on the FAS with a log rank test at an initial significance level of (4/5)α within a sequentially rejective multiple testing procedure described in detail in Section 2.2.6.4. Note that the significance level for the final test will be adjusted to account for the planned interim analysis.

Number and percentage of subjects who died from cardiovascular reasons based on the number of subjects in the population as denominator will be provided by treatment group. The hazard ratio (relative risk) and its associated two-sided 95% confidence interval will be estimated based on a Cox proportional hazards model with treatment assignment as a factor.

The Kaplan-Meier estimates of the survival functions for each treatment group will be plotted. The Kaplan-Meier estimates of the cumulative event rate will also be presented in tables by treatment group for each day and also by time interval.

WHF

Time to worsening heart failure (WHF) through Day 5 will be analyzed using Gehan's generalized Wilcoxon test at an initial significance level of (1/5)α within a sequentially rejective multiple testing procedure described in detail in Section 2.2.6.4.

Number and percentage of patients who experienced WHF based on the number of patients in the population as denominator will be provided by treatment group. The hazard ratio (relative risk) and its associated two-sided 95% confidence interval will be estimated based on a Cox proportional hazards model with treatment as a factor.

Kaplan-Meier curves will be presented graphically by treatment group and Kaplan-Meier estimates for selected time points with 95% confidence intervals will be tabulated.

Model assumption

For *primary efficacy* variables analyzed by Cox proportional hazard model, the important model assumption that is made when applying a Cox proportional hazards regression model is the proportionality of the hazard function(s) over time. This assumption will be tested using the same primary Cox proportional hazards regression model and in addition a time-dependent explanatory variable created through the interaction between log time and treatment. In addition, a graphical method will be used for checking the proportional hazards assumption. A plot of log-log survivor function vs. log time for each treatment group will be used. Approximate parallelism between the curves for the treatment groups will provide supportive evidence of the proportional hazards assumption. It should be noted that the log rank test remains valid even if the proportional hazards assumption is not fulfilled.

The SAS procedures PHREG and LIFETEST will be used to conduct the analyses.

2.2.5 Supportive/sensitivity analyses for primary endpoint

- Both primary efficacy variables, CV death and WHF through Day 5, will also be analyzed on the PPS as supportive information.
- To explore the consistency of treatment effects across subgroups, the CV death and WHF will also be analysed for respective subgroups provided in Table 2.3 by adding the subgroup and the treatment by subgroup interaction terms to the Cox regression model of the primary efficacy endpoint. In all analyses a test for heterogeneity to assess for possible interactions between treatment and baseline variables will be applied. Results will be presented graphically as forest plots. The objective of the subgroup analyses is to evaluate the consistency of treatment effects across a wide variety of patient groups.
- A covariate-adjusted Cox-regression model will be fitted for the CV death with baseline covariates (Age (continuous), history of heart failure (yes/no), cerebrovascular accident (yes/no), respiratory rate (continuous), systolic BP (continuous), edema (categorical), orthopnea (categorical), sodium (continuous), creatinine (continuous)) identified from

RELAX-AHF exploratory analysis (Metra, et al 2013) to account for other factors that might affect prognosis using likelihood ratio test for treatment effect with exact handling of ties.

- Analyses will be performed using a modified definition of WHF, e.g. subjects with (1) worsening signs AND symptoms or (2) IV intervention and/or mechanical support. Another sensitivity analyses that exclude subjects whose only symptom is abdominal bloating/discomfort or fatigue/decreased exercise tolerance will also be provided.
- Subjects with WHF through Day 5 (yes/no) will also be analyzed using logistic regression with treatment and region as factors.

2.2.6 Key secondary analysis

There are three key secondary efficacy variables defined in this study.

All key secondary variables will be analyzed using the FAS and PPS populations, and will be repeated for the subgroups provided in Table 2-3 by adding the subgroup and the treatment by subgroup interaction terms to the model for time to event endpoints.

Whenever the event or visit time needed for defining the time-to-event or censoring date and time is missing or incomplete, an imputation rule for missing/incomplete time will be applied. More details about the imputation rule are presented in Section 2.2.9.

2.2.6.1 Time to all-cause death through Day 180

Time to all-cause death will be calculated as:

Time-to-event (days) = date of all-cause death – randomization date,

The analysis will be performed similarly to primary endpoint.

For those subjects without events, the censoring date will be the earliest in the following dates:

- **180 days**
- **Date subject withdrew consent to all follow-up (if no subsequent vital status documented)**

  Note: in accordance with local laws and regulations, vital status may be publically available
- **Date of subject's last visit or contact or vital status**

A covariate-adjusted Cox-regression model similar to the sensitivity analysis of primary efficacy variable will also be performed as sensitivity analysis.

2.2.6.2 Length of total hospital stay (LOS) during the index AHF hospitalization

Length of stay will be defined as the index hospitalization discharge date and time minus the randomization date and time, whereas time is used for better precision and converted to days.

Subjects still in the hospital at Day 60 will be censored at Day 60. Subjects who die during the initial hospitalization will be assigned the maximum length of stay (including those censored at Day 60) plus 1 day.

Subjects missing the discharge date, such that a determination of the length of stay cannot be calculated, will be assigned the mean length of stay over all subjects with not missing discharge date.

Treatment groups will be compared using a Wilcoxon rank sum test.

For subgroup analysis, a p-value for the test of the treatment by subgroup interaction term will be based on a two-way model with treatment, subgroup and treatment-by-subgroup interaction using rank-transformed data.

The van Elteren extension to the Wilcoxon rank sum test will also be carried out stratifying for region.

2.2.6.3 Time to first occurrence of the composite endpoint of CV death or rehospitalization due to heart failure (HF)/renal failure (RF) through Day 180

Time to event will be calculated as date of first event of adjudicated CV death or adjudicated rehospitalization due to HF/RF minus randomization date in days. If the reason for adjudicated rehospitalization is unknown, it will be considered as HF or RF.

The analysis will be performed similarly to primary endpoint.

For subjects without events the censoring date will be the earliest in the following dates:

- **180 days**
- **Date subject withdrew consent to all follow-up**
- **Date of subject's last visit or contact**
- **Date of non-CV death**

In addition, the estimates with 95% CI at 30, 60, and 90 days from the Kaplan-Meier method will be provided. Of note estimates at 60 and 90 days are considered as exploratory analyses.

The analysis for one of components of the composite endpoint, time to rehospitalization due to HF/RF through Day 180, will also be performed. The estimates with 95% CI and p-value at Day 180 from the Kaplan-Meier method will also be provided.

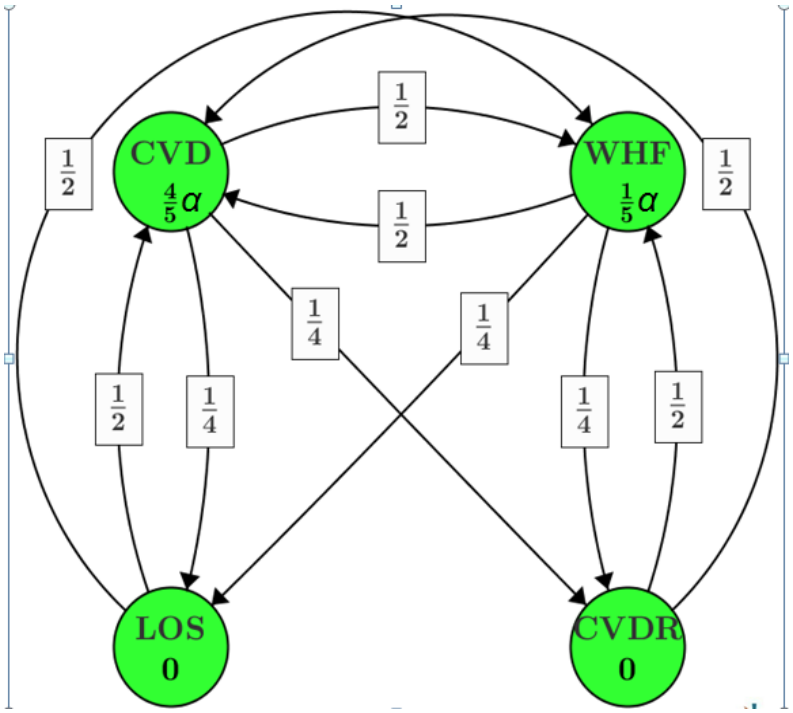
2.2.6.4 Multiple testing procedure

The key secondary efficacy variables of all-cause mortality through Day 180 will not be formally incorporated in a multiple testing procedure. It will be tested with a full α (two-sided) by itself if one of the primary endpoints, CV death or WHF is statistically significant.

The primary endpoints of CV death and WHF will be tested first. If at least one of the primary endpoints is significant, the other two key secondary variables will be tested using a sequentially rejective multiple testing procedure in testing LOS and composite of CV death or rehospitalization due to HF/RF (CVDR) in controlling overall α (Bretz, et al 2009). Figure 2-1 illustrates this multiple testing procedure for two primary and two key secondary variables

with strong control of overall α level of 5%. Note: although CV death will be tested initially at 0.02 (one-sided) significance level and WHF at 0.005 (one-sided) significance level, the final significance level for CV death and WHF will be updated according to the graphical testing strategy incorporating two secondary variables, LOS and CVDH, as depicted in below figure.

Figure 2-1 Sequentially rejective multiple testing procedure for two key secondary efficacy variables



Where CVD is CV death, WHF is worsening heart failure, LOS is length of hospital stay, and CVDR is a composite variable of CV death or rehospitalization due to heart failure or renal failure.

2.2.7 Other secondary analysis

Other secondary efficacy variables comprise:

2.2.7.1 Length of ICU and/or CCU stay for the index AHF hospitalization

The days stayed in ICU/CCU will be obtained from eCRF page directly. If subjects die during the initial hospitalization, the maximum number of days plus 1 day in the ICU/CCU over all subjects will be assigned.

The length of ICU and/or CCU stay for the index AHF hospitalization will be analyzed similarly to the length of hospital stay, i.e., it will be compared using a Wilcoxon rank sum

test. As a sensitivity analysis, the van Elteren extension to the Wilcoxon rank sum test will also be carried out stratifying for region.

2.2.7.2 Change from baseline in congestive signs and symptoms of HF through Day 5

Detailed criteria to assess subject's signs and symptoms (dyspnea on exertion, orthopnea, edema, rales, and jugular venous pulse) of heart failure can be found in protocol Appendix 15.

Subjects who die or have a worsening heart failure event (either during the index hospital or rehospitalization due to heart failure) will have their score imputed to be the worst observed score in any subject at any time point in the respective analysis population carried forward from the day of death or WHF regardless of whether their original data are missing through Day 5. For patients who are alive with missing values, the last preceding non-missing value will be used for imputation.

A time-to-event (response defined as >=1 point improvement from baseline) analysis will be applied to provide treatment comparisons. Hazard ratios (serelaxin/placebo) of being a responder for each of the signs and symptoms variables will be estimated using a Cox regression model stratified by region with treatment as a factor.

Since the assessment only occurs on the scheduled visits, the nominal scheduled day will be used directly to determine the time to event with exact handling of ties.

For those subjects without events, the censoring date will be the earliest in the following dates:

- **6 days**
- **Date subject withdrew consent to all follow-up**
- **Date of subject's last visit or contact**

Change from baseline in congestive signs and symptoms of HF through Day 5 will be analyzed by displaying the percentage of subjects in each category by treatment and time points for each of the signs and symptoms variables. The scores will be presented as both categorical and continuous statistics (i.e., rales will be defined as 0, 1, 2, 3 for its 4 categories; jugular venous pulse will be defined as 0, 1, 2 for its 3 categories). Change from baseline will be computed as the baseline value subtracted from the post-baseline value. Treatment groups will be compared with respect to change from baseline at each time point using Wilcoxon rank sum tests.

2.2.7.3 Change from baseline in selected biomarkers from baseline through Day 14 in a subset of randomized subjects

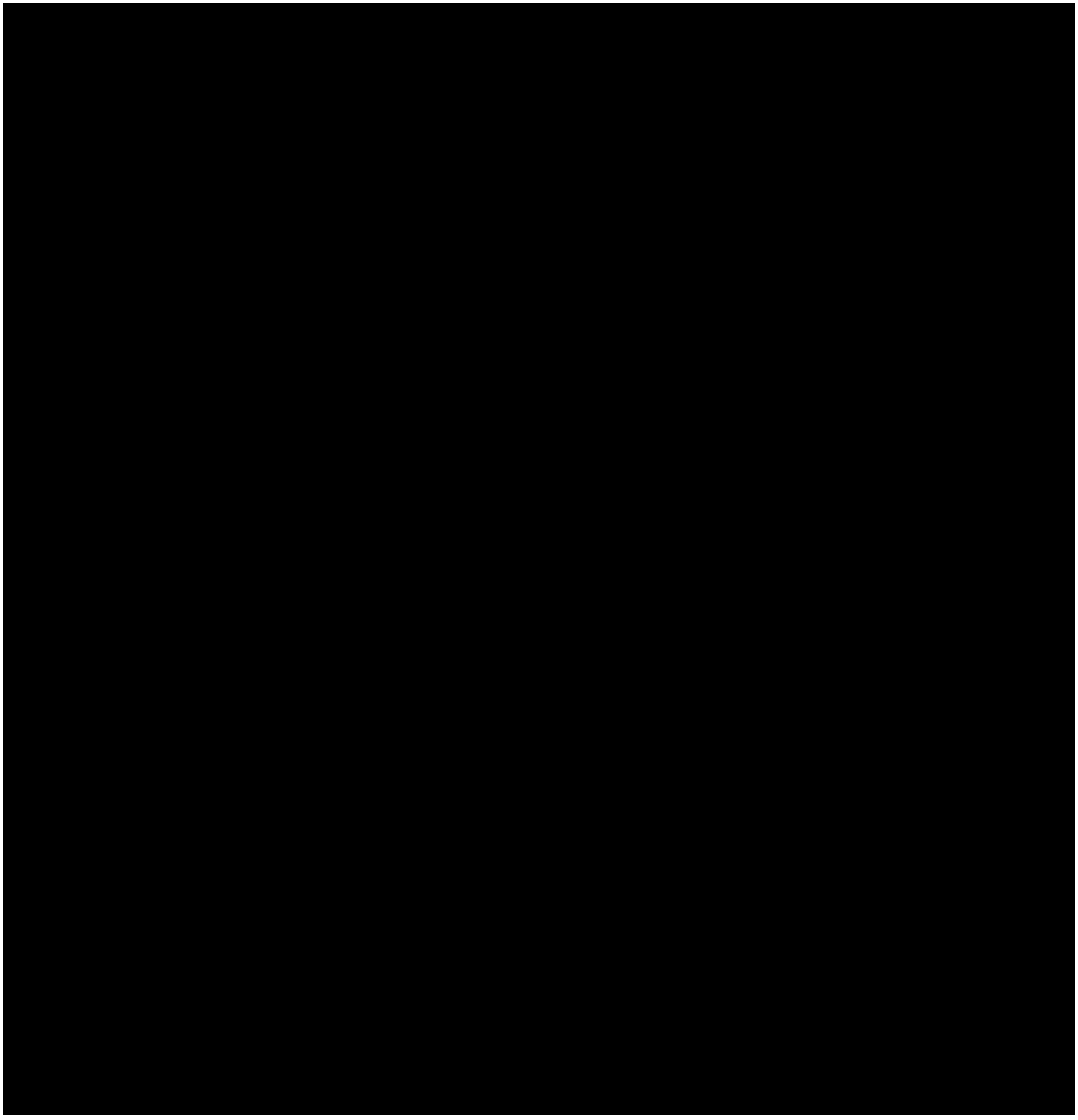
A biomarker substudy as measured by a core laboratory will be conducted to enroll at least 1600 subjects.

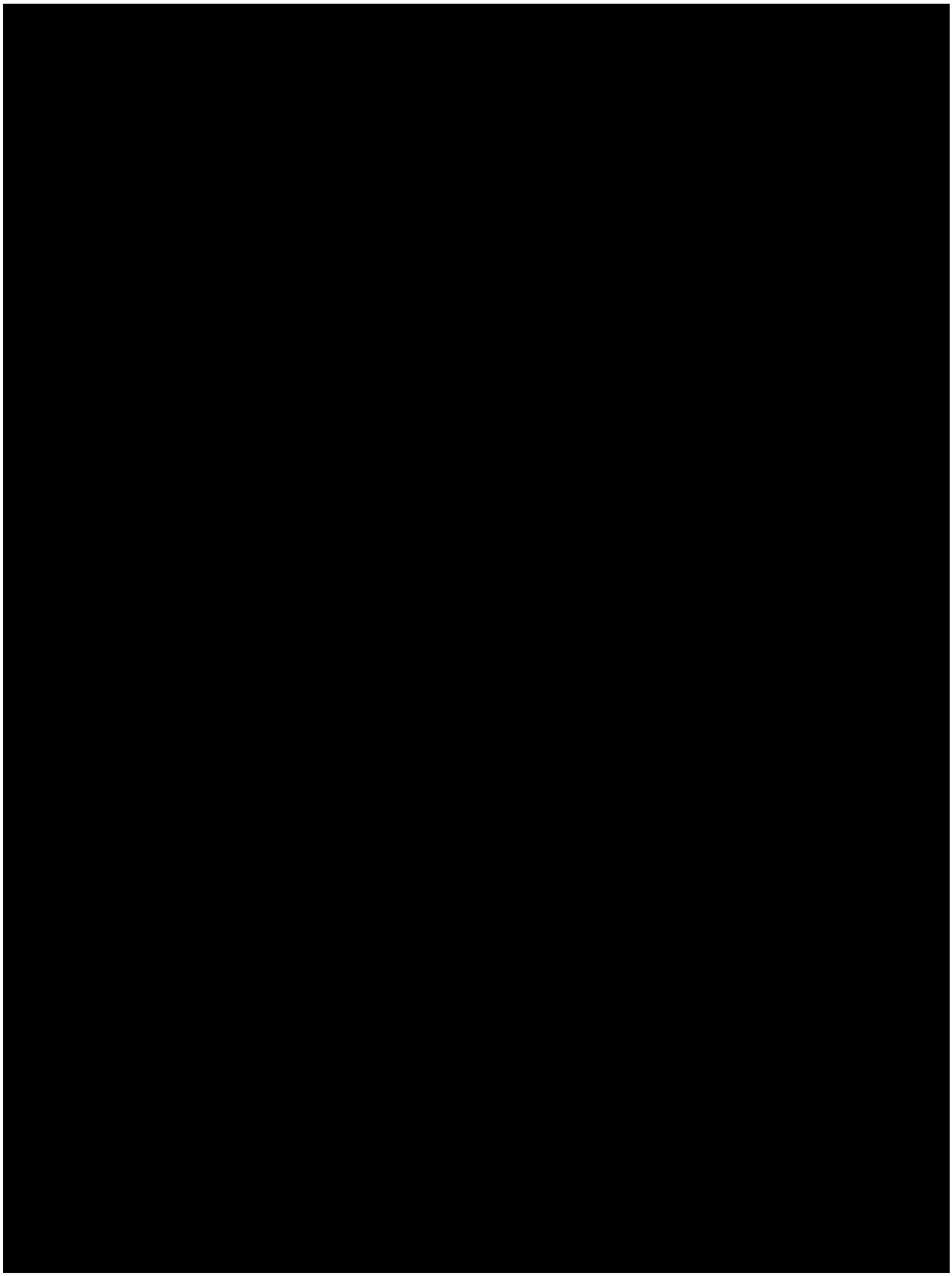
Change from baseline through Day 14 will be performed on the key biomarkers high sensitivity troponin T, cystatin C and NT-proBNP using a repeated measurement model. It is anticipated that the assumption of normally distributed residuals for biomarkers may not hold. Therefore, the model will be performed on log-transformed biomarker. Treatment group, visit (Day 2, Day 5 and Day 14) and region will be fitted as factors and log-baseline biomarker will

be fitted as a continuous covariate. Treatment group by visit and visit by log baseline will be included as interaction terms in the model.

The analysis will be presented by treatment group and visit (Day 2, Day 5 and Day 14). The detailed analysis of changes in biomarkers and analyses of associations of baseline biomarker levels with clinical outputs are described in the separate biomarker statistical analysis plan.

Of note the discharge visit for biomarkers will be mapped to Day 2, Day 5 (within +/- 1 day window) or Day 14 (within +/- 3 days window) if Day 2, Day 5 or Day 14 visits are missing.





2.2.9 Handling of missing values/censoring/discontinuations in efficacy analyses

The detailed censoring for time-to-event is mentioned in previous sections.

Missing date

It is assumed that the day and month of all event dates will be known. If the date's day of the month is missing such that a determination of whether the event occurred within the reporting window cannot be made, the day will be imputed following an attempt to get an approximate date from the family or other records:

a) if only the month of the event is known, then the 15th day of this month will be imputed for a missing day in case the month and year are after the month and year of double blind treatment start. If the month and year are equal to the double blind treatment start then the event date will be imputed by the treatment start date.
b) if only the year of the event is known, then the 1st July will be imputed for a missing day and month in case the 1st July is after double-blind treatment start (otherwise treatment start date will be used), or
c) for events occurring between two visits after randomization, without other information available, the date in the middle of these visits will be used.

An adjudicated death with unknown cause will be classified as a CV death. Similarly, investigator suspected deaths of unknown cause will be classified as CV deaths.

Missing time

For efficacy variables require date and time in calculation, if time is missing or incomplete, the following will be applied:

a) if hour and minute are both missing, then the 12:00pm (24hr clock) of the day will be imputed in case the hour and minute are after double blind treatment start (otherwise, treatment start hour and minute will be used).

b) if only minute is missing, 30th minute of the hour will be imputed in case the minute is after double blind treatment start (otherwise, treatment start minute will be used).

2.3 Safety evaluation

Safety will be assessed by comparing the study drug to the placebo group with regard to the frequency of AEs, changes in clinical laboratory test results (chemistry and hematology), confirmed blood pressure decreases and vital signs. All safety analyses will be presented for the SAF.

The treatment group will be based on the actual treatment received for each subject.

2.3.1 Adverse events

The incidence of AEs and non-serious AEs recorded through Day 5 and the incidence of SAEs recorded through Day 14 will be presented for the Safety population.

For presentation, AE verbatim text will be coded into a MedDRA (Medical Dictionary for Regulatory Activities) term and classified by MedDRA System Organ Class (SOC) and Preferred Term (PT). The incidence of all reported AEs will be summarized by SOC and PT. A subject reporting the same event more than once will be counted once when reporting the number of subjects with that particular event. Percentages will be based on the number of subjects in SAF.

SAEs will be summarized by time period of onset: from study drug initiation through Day 5 and from Day 6 through Day 14 separately. All AEs through Day 14 will also be provided. Additionally, a summary of AEs by PT and severity, using the worst reported severity grade for each event for the subject, will be provided.

All study-drug-related AEs, AEs leading to discontinuation of study drug, and study-drug-related SAEs, SAEs with an outcome of death, and SAEs leading to study drug discontinuation will be summarized by the number and percentage of subjects reporting particular events. For analysis purposes, if the relationship to study drug is unknown or missing, the AE will be considered to be drug-related.

For the legal requirements of ClinicalTrials.gov and EudraCT, two required tables summarizing treatment emergent adverse events which are not serious adverse events with an incidence greater than 0.5% and summarizing treatment emergent serious adverse events and SAE suspected to be related to study treatment will be provided by system organ class and preferred term on the Safety population.

If for a same subject, several consecutive AEs (irrespective of study treatment causality, seriousness and severity) occurred with the same SOC and PT:

- a single occurrence will be counted if there is ≤ 1 day gap between the end date of the preceding AE and the start date of the consecutive AE
- more than one occurrence will be counted if there is > 1 day gap between the end date of the preceding AE and the start date of the consecutive AE

For occurrence, the presence of at least one SAE / SAE suspected to be related to study treatment / non SAE has to be checked in a block e.g., among AE's in a ≤ 1 day gap block, if at least one SAE is occurring, then one occurrence is calculated for that SAE.

Detailed AE listings will also be provided.

In addition, AEs representing identified and potential safety risk in addition to events of interest will be grouped. The incidence (absolute and relative frequency) of each potential risk will be summarized by treatment group. Odds ratios and associated 95% confidence intervals comparing serelaxin with placebo will be presented using the Mantel-Haenszel test. In addition, subjects experiencing each of these risks will be listed. Such safety risks and events of interest are displayed in below table.

Table 2-7 Safety risk and event of interest search specification

Type of Risk	Specification
Safety risk	
Hypotension	Search terms include the following PTs: Blood pressure ambulatory decreased Blood pressure decreased Blood pressure diastolic decreased Blood pressure immeasurable Blood pressure orthostatic decreased Blood pressure systolic decreased CT hypotension complex Diastolic hypotension Dizziness postural Hypotension Mean arterial pressure decreased Neonatal hypotension Orthostatic hypotension Procedural hypotension Altered state of consciousness Blood pressure abnormal Blood pressure ambulatory abnormal Blood pressure diastolic abnormal Blood pressure fluctuation Blood pressure inadequately controlled Blood pressure orthostatic abnormal Blood pressure systolic abnormal Blood pressure systolic inspiratory decreased Consciousness fluctuating Depressed level of consciousness Dizziness Dizziness exertional Labile blood pressure Loss of consciousness Presyncope

Type of Risk	Specification
	Schellong test
	Syncope
	Tilt table test positive
Hemoglobin/ hematocrit transient decrease	SMQ 'Haemorrhages' (Broad)
	HLGT 'Anaemias nonhaemolytic and marrow depression'
	SMQ 'Haematopoietic erythropenia' (Broad)
Fetotoxicity and Teratogenicity	SMQ 'Pregnancy and neonatal topics' (Broad)
Event of interest	
Hypokalemia	Search terms include the following PTs:
	Alkalosis hypokalaemic
	Blood potassium abnormal
	Blood potassium decreased
	Hypokalaemia
	Hypokalaemic syndrome
	Hyperkaliuria
	Transtubular potassium gradient
	Urine potassium abnormal
	Urine potassium decreased
	Urine potassium increased
Menorrhagia/Metrorrhagia	HLGT 'Menstrual cycle and uterine bleeding disorders'
Potential promotion of cancers	SMQ 'Malignancies' (Broad)
Immunogenicity/Hypersensitivity	SMQ 'Hypersensitivity' (Broad)
	SMQ 'Angioedema'(Broad)
	SMQ 'Anaphylactic reaction'(Broad)
	SOC 'Immune system disorders'
	HLGT 'Immunology and allergy investigations'
	HLGT 'Administration site reactions'
	SMQ 'Severe cutaneous adverse reactions'(Broad)
Cardiac failure	SMQ 'Cardiac failure' (Broad)
QT prolongation	SMQ 'Torsade de pointes/QT prolongation' (Broad)
	SMQ 'Cardiac arrhythmias' (Broad)
Thromboembolic events	SMQ 'Embolic and thrombotic events'
Renal impairment	SMQ 'Acute renal failure' (Broad)
Respiratory failure	SMQ 'Respiratory failure' (Broad)
Hepatotoxicity	SMQ 'Drug related hepatic disorders – comprehensive search' (Broad)

AEs with an onset between informed consent and study drug initiation will be listed.

For each AE listed, the verbatim term and PT, the corresponding severity, relationship to study drug, seriousness (yes/no), and start and stop date and time of the event will be specified and displayed by treatment group and subject number.

2.3.2 Laboratory data

2.3.2.1 Hematology and chemistry

In addition to local laboratory evaluations for all randomized subject, a laboratory substudy will be conducted at a subset of selected sites to enroll at least 1600 subjects. Subjects without written informed consent for the laboratory substudy before any study-specific assessment is performed will be excluded from the population.

Descriptive summaries of actual and change from baseline will be presented for all continuous clinical laboratory values by treatment and time point.

Clinical shift tables showing the pattern of change from baseline versus each post-baseline time point will be presented based on the cross-classification of the number of subjects with clinical laboratory values below (low), within (normal), or above (high) normal ranges for each laboratory test.

All laboratory data will be listed for each subject. Laboratory values that are outside the normal ranges will be flagged high (H) or low (L).

The number and percentage of subjects with clinically notable laboratory results, including categorical analysis on laboratory parameters of interest after baseline will be presented. Clinically notable laboratory values are defined in protocol Appendix 1. The notable laboratory values related to liver function are detailed in Section 2.3.7.

The data from local and central laboratory will be presented separately.

2.3.2.2 Creatinine and eGFR changes

eGFR will be calculated according to simplified Modification of Diet in Renal Disease (sMDRD).

The proportion of subjects with > 0.3 mg/dL increase, >0.5 mg/dL increase, > 1.0 mg/dL increase and >50% increase from baseline in serum creatinine levels at each time point, and at any time point, will be presented.

The proportion of subjects with decreases from baseline GFR of >25%, >40%, >50%, and >30 mL/min/1.73 m^2 at each time point, and at any time point, will be presented.

The data from local and central laboratory will be presented separately.

2.3.3 Vital signs

Descriptive statistics will be provided for values and change from baseline at each assessment time point for systolic blood pressure, diastolic blood pressure, pulse, respiratory rate, body temperature, and weight.

Blood pressure and pulse during 48 hours drug infusion and 48 hours post infusion will also be listed.

2.3.4 Electrocardiogram (ECG)

An ECG substudy will be conducted at a subset of selected sites to enroll at least 500 subjects. Subjects without written informed consent for the ECG substudy before any study-specific assessment is performed will be excluded from the population.

The analyses will be performed based on subjects who have at least one ECG at baseline (ECG recordings taken before the infusion starts will be considered) and one ECG post baseline (48 hour time point) which is recorded within 60 minutes after study drug infusion ends.

The following quantitative variables will be summarized: RR interval, PR interval, QRS duration, QT interval, and corrected QT interval (QTc).

The number and percentage of subjects with the following criterion will be presented.

- QT > 500 msec
- QT > 480 msec
- QT > 450 msec
- QT > 450 msec, male only
- QT > 460 msec, female only
- QT increases from baseline > 30 msec
- QT increases from baseline > 60 msec
- PR > 200
- PR > 220 msec
- PR increases from baseline > 25%
- QRS > 110 msec and increase from baseline >=25%
- QRS > 120 msec and increase from baseline >=25%
- Heart rate >100 beats per minute and increase from baseline >=25%
- Heart rate < 50 beats per minute and decrease from baseline >=25%

QT data will be analyzed for both the Fridericia (primary) and Bazzett's (secondary) corrections.

Newly occurring abnormalities in ECG morphology will be summarized by treatment group and overall. Newly occurring is defined as no occurrence on any of the baseline ECGs with the presence of the finding on any single ECG post baseline.

A listing of all abnormalities will be provided, as well as a by-subject listing of all quantitative ECG parameters.

2.3.5 Confirmed blood pressure decreases (CBPD)

The number and percentage of subjects who experience a confirmed blood pressure decrease requiring a study drug dose adjustment during study drug administration will be provided. Among subjects who experience a confirmed blood pressure decrease event, the events will be further characterized with respect to impact on study drug administration (dose decreased or

study drug discontinued or both), the timing of event onset relative to study drug initiation, the magnitude of the blood pressure decrease at its onset and the trough blood pressure recorded, the blood pressure measurements after study drug discontinuation, outcome, treatment(s) required, and whether considered symptomatic, i.e. an AE. In addition, these summaries will be provided separately for those confirmed events that resulted in study drug dose reduction, and those that resulted in study drug discontinuation.

Course of blood pressure after the occurrence of first CBPD up to Day 5 will be presented.

To further investigate the impact of CBPD, the following analyses will be performed.

- Baseline characteristics by CBPD subgroup (yes/no).
- Concomitant medications by CBPD subgroup (yes/no).
- CV death through Day 180, WHF through Day 5 and CV death or rehospitalization due to heart failure or renal failure through Day 180 by CBPD impact on study drug administration (dose decreased or study drug discontinued or both).
- AEs associated with the occurrence of the CBPD.

2.3.6 Pregnancy test results

Any pregnancies reported will be listed.

2.3.7 Liver toxicity function and liver events

2.3.7.1 Liver toxicity function

The liver toxicity (LFT abnormality) will be identified with the SMQ 'Drug related hepatic disorders – comprehensive search' (Broad) and the laboratory parameters mainly AST (aspartate aminotransferase; also known as SGOT), ALT (alanine aminotransferase; also known as SGPT), ALP (alkaline phosphatase) and TBL (total bilirubin; conjugated (direct) and unconjugated (indirect) bilirubin).

The standard SMQ-PT table and odds ratio will be used to provide the number and percentage of subjects with hepatic disorders as mentioned in Section 2.3.1 (hepatotoxicity, SMQ 'Drug related hepatic disorders – comprehensive search')

The liver function related parameters will be shown in the laboratory standard tables (see Section 2.3.2).

In addition, summary tables will be provided on the number and percentage of subjects having the following criteria:

Table 2-8 Criteria for evaluating liver toxicity

Parameter	Criterion
ALT or AST	ALT or AST > 3xULN
	ALT or AST > 5xULN
	ALT or AST > 8xULN
	ALT or AST > 10xULN

Parameter	Criterion
ALT/AST & TBL	ALT or AST >3x ULN and TBL >1.5x ULN
	ALT or AST >3x ULN and TBL >2x ULN
	ALT or AST >5x ULN and TBL >2x ULN
	ALT or AST >8x ULN and TBL >2x ULN
	ALT or AST >10x ULN and TBL >2x ULN
	ALT or AST >3x ULN and TBL >2x ULN and ALP <2x ULN
TBL & ALP	TBL >1.5x ULN and ALP >2x ULN
	TBL >2x ULN and ALP >2x ULN
Isolated TBL	TBL >1.5x ULN & ALT and AST <=3x ULN and ALP <=2x ULN
	TBL >2x ULN & ALT and AST <=3x ULN and ALP <=2x ULN
	TBL >3x ULN & ALT and AST <=3x ULN and ALP <=2x ULN
Isolated ALP	ALP >1.5x ULN & ALT and AST <=3x ULN and TBL <=1.5x ULN
	ALP >2x ULN & ALT and AST <=3x ULN and TBL <=1.5x ULN
	ALP >3x ULN & ALT and AST <=3x ULN and TBL <=1.5x ULN
	ALP >5x ULN & ALT and AST <=3x ULN and TBL <=1.5x ULN

ALP: alkaline phosphatase; ALT: alanine aminotransferase; AST: aspartate aminotransferase; TBL: total bilirubin; ULN: upper limit of normal.

Plots showing the relationship between TBL and ALT or AST (eDish plot) will also be provided.

2.3.7.2 Liver events

Detailed information of liver event is collected only if subjects report liver event.

The following eCRF pages will be collected and details are listed for the reported liver event.

- Medical history possibly contributing to liver dysfunction.
- Liver event overview: details of clinical signs and symptoms, event related to study drug, other alternative diagnoses etc.
- History of alcohol use
- Drugs of abuse history
- Imaging
- Pathology
- Autoimmune
- Viral serology
- Local laboratory liver function test
- Concomitant medication Acetaminophen/Paracetamol

2.4 Interim analyses

It is planned to establish an external and Independent Data Monitoring Committee (DMC) to monitor subject safety data on a regular basis during the course of the study. For this study, it is planned that the DMC will review safety data on a regular frequency, e.g., every six months. DMC may request additional safety data review. Such safety analyses do not inflate the type I error for the primary efficacy hypothesis testing and thus require no multiplicity adjustments.

It is planned to have one formal interim efficacy analyses. The cut-off for this interim analysis is planned when the information fraction (number of confirmed CV deaths relative to the targeted total number of confirmed CV death at the study completion) of 60% is achieved (corresponding to 329 CV deaths). Some adjustment to the time of interim analysis may be made to coincide with the regular DMC meetings. One-sided tests will be performed and appropriate statistical adjustments for the interim analysis actually performed will be made to ensure the overall one-sided type I error of 0.02 (Lan-DeMets spending function to match an O'Brien-Fleming stopping boundary). For the interim analysis, the analysis dataset will comprise all subjects who were randomized before the cut-off date.

The trial may be concluded early for efficacy, if a significant difference between two treatment arms for the number of CV deaths is achieved by crossing the pre-specified boundary at interim analysis. Interim analysis will be performed by an external independent statistician at an external data analysis center, who will not be involved in the study conduct. The results will be reviewed by the independent DMC. Investigators and others who are involved in the conduct of the trial and in the analysis of the final trial results, or who have contact with study centers, will remain blinded to the treatment codes and interim analysis results until all monitoring decisions have been made and the database has been locked for final analysis.

If the DMC recommends stopping the study due to overwhelming efficacy of CV deaths, i.e., crossing the pre-specified boundary and favorable risk and benefit ratio, the multiple testing procedure in Figure 2-1 will be applied and updated as Figure 2-2 below. The significance levels for WHF, LOS and CVDR are 0.6α, 0.2α and 0.2α.

The Lan-DeMets spending function of the O'Brien-Fleming stopping boundary will then be applied to determine the nominal significance level at the interim analysis based on the one-sided alpha of 0.025.

The information fraction at the interim analysis for the three efficacy variables will be calculated as follows:

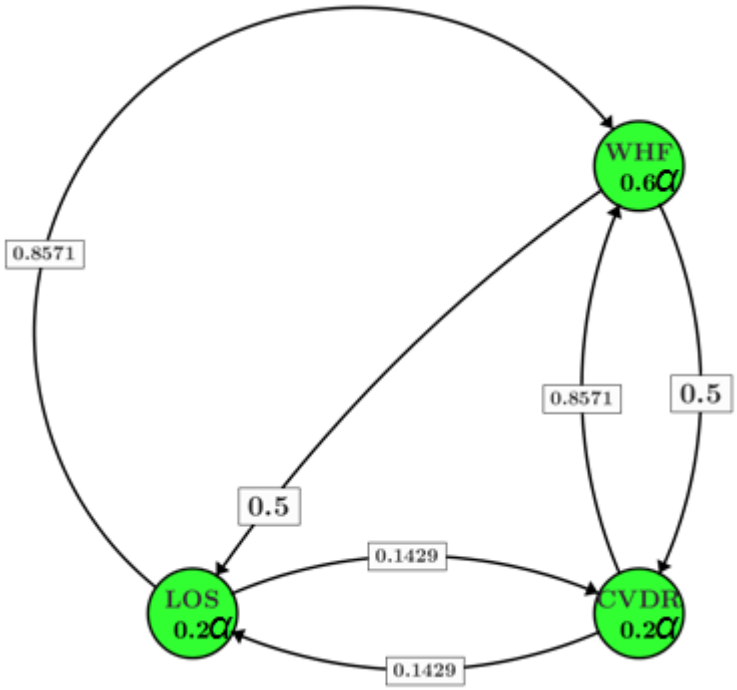
- WHF: number of randomized subjects at the interim analysis divided by 6800

  It is expected that all randomized subjects will have complete WHF information through Day 5, as all patients will be followed up at least 14 days.

- LOS and CVDR: number of randomized subjects at the interim analysis divided by 6800 times 0.9

  This assumes that 10% of the efficacy data will be missing at the interim analysis cutoff date.

Figure 2-2 Sequentially rejective multiple testing procedure if DMC recommends stopping of the study due to overwhelming efficacy in CV death



Applying the Lan-DeMets spending function of the O'Brien-Fleming stopping boundary if the DMC recommends stopping the study, assuming 70% of the randomized subjects are accrued at the interim analysis, the nominal one-sided alpha for the three efficacy variables are calculated as shown in the Table 2-9.

Table 2-9 Nominal one-sided alpha at 70% randomized subjects accrued if CVD is rejected

Efficacy variable	Updated alpha if CVD rejected	Information fraction	Nominal 1-sided alpha at interim analysis
WHF	0.6x0.025=0.015	0.7	0.0036
LOS	0.2x0.025=0.005	0.7x0.9=0.63	0.0004
CVDR	0.2x0.025=0.005	0.7x0.9=0.63	0.0004

Note if the boundary is crossed for any of the remaining hypotheses, the multiple testing procedure will reject the hypothesis and update the graph. The nominal one-sided significance level will be re-calculated. The process continues until no hypothesis can be rejected.

For subjects not discharged by interim analysis cutoff date, the LOS will be imputed as the maximum of

(1) mean value for subjects with actual LOS, including censoring at Day 60 and die in hospital

(2) LOS at the cutoff date

All cause death will be tested at alpha of 0.025 (one-sided).

2.5 Sample size and power considerations

The alpha (α) level of 5% is split to CV death at 4% and to WHF at 1%. The sample size calculation is based on the CV death endpoint. This is an event-driven trial with fixed maximum duration of 180 day follow-up period for each patient. In the sample size calculation, the 180-day CV death rate is assumed to be 9.0% in the placebo group, which is about 80% of 11.3% of all cause death observed in the placebo group in the RELAX-AHF study. [redacted]

The interim analysis will be performed when 329 CV deaths are accrued. At the interim analysis, if Z-score of 2.786 (one-sided nominal p = 0.0027) is observed, which corresponds to HR of 0.735 (at least 26.5% RRR), the DMC may recommend to stop the study due to crossing the boundary. If there is no stopping at the formal efficacy interim analysis, the final analysis will need to have a Z-score of 2.072 (one-sided nominal p = 0.0192), which corresponds to HR of 0.838 (at least 16.2% RRR), to declare a significant study. (East ® 5.4.2, Cytel Inc., Cambridge MA, 2010). [redacted]

If the observed overall event rate of the CV death endpoint during the course of the trial is much lower than expected based on the design assumptions (assessed on a blinded basis), adjustments may be made to the number of patients to be randomized in order to achieve the required number of events in the pre-planned timeframe (180 days) and thus to preserve the target study power.

For the sample size of 6,800, the unconditional power for the WHF endpoint and all three key secondary endpoints under various HRs (mean difference for Length of Stay) are summarized in the tables below:

Table 2-10 Assumed effect sizes for worsening heart failure and the key secondary efficacy variables on power calculation

	Assumed HR or mean difference
Worsening Heart Failure (placebo rate: 12.2%)	0.70
	0.80
All cause death (placebo rate: 11.3%)	0.80
	0.816
	0.83
Length of Stay (mean ±SD between placebo and serelaxin: 0.9 ±9.6 day)	0.9 day
	0.7 day
CV death or HF/RF rehospitalization (placebo rate: 34.1%)	0.85
	0.88

Note: Assumed placebo rates for all cause death and WHF and mean difference (mean + SD) in LOS between serelaxin and placebo were based on RELAX AHF study; placebo event rate of CV death or HF/RF rehospitalization was estimated to be 34.1% by assuming an exponential distribution with the same hazard as that observed through Day 60 in RELAX AHF study.

Table 2-11 Powers (unconditional) of primary and key secondary variables on selected scenarios

A) All cause death

Scenarios of effect sizes	All cause death Day 180*
All-cause death (HR=0.80); WHF (HR=0.70)	83
All-cause death (HR=0.816); WHF (HR=0.70)	77
All-cause death (HR=0.82); WHF (HR=0.70)	70
All-cause death (HR=0.80); WHF (HR=0.80)	82
All-cause death (HR=0.816); WHF (HR=0.80)	76
All-cause death (HR=0.82); WHF (HR=0.80)	70

Note: The true HR for CV death is assumed to be 0.78 in all scenarios.

*Power if one or both of the two primary variables are significant.

B) CVD, WHF, LOS and CVDR

Scenarios of effect sizes	CV death Day 180	WHF Day 5	LOS	CV death or HF/RF hosp.	Reject All
WHF (HR=0.7) – LOS (0.9 days) – CVD/rehosp (HR=0.85)	82	99	93	93	77
WHF (HR=0.7) – LOS (0.9 days) – CVD/rehosp (HR=0.88)	82	99	92	79	68
WHF (HR=0.7) – LOS (0.7 days) – CVD/rehosp (HR=0.85)	82	99	78	93	69
WHF (HR=0.7) – LOS (0.7 days) – CVD/rehosp (HR=0.88)	82	99	77	77	60
WHF (HR=0.8) – LOS (0.9 days) – CVD/rehosp (HR=0.85)	82	82	87	87	68
WHF (HR=0.8) – LOS (0.9 days) – CVD/rehosp (HR=0.88)	82	82	86	73	59
WHF (HR=0.8) – LOS (0.7 days) – CVD/rehosp (HR=0.85)	82	81	73	86	60
WHF (HR=0.8) – LOS (0.7 days) – CVD/rehosp (HR=0.88)	81	81	72	72	53

Note: The true HR for CV death is assumed to be 0.78 in all scenarios.

Note: The power for the WHF endpoint is approximated based on logrank test, but the analysis will be based on Gehan-Wilcoxon test.

3 References

1. Bretz F, Maurer W, Brannath W, and Posch M (2009) A graphical approach to sequentially rejective multiple test procedures. Statistics in Medicine 28, 586-604
2. Levey AS, Coresh J, Balk E, et al (2003) National Kidney Foundation Practice Guidelines for Chronic Kidney Disease: Evaluation, Classification, and Stratification. Ann Intern Med; 139:137-147.
3. Metra M, Ponikowski P, Cotter G, Davison BA, Felker GM, Filippatos G, Greenberg BH, Hua TA, Severin T, Unemori E, Voors AA, Teerlink JR (2013). Effects of serelaxin in subgroups of patients with acute heart failure: results from RELAX-AHF. European Heart Journal; 34:3128-3136.