Statistical Analysis Plan

Protocol Title: EMpagliflozin evaluation By measuRing ImpAct on HemodynamiCs in PatiEnts with Heart Failure (EMBRACE-HF)

Protocol Number: 1245.129

ClinicalTrials.gov registration: NCT03030222

Version Number: Final

Version Date: March 8, 2020

Abbreviation or special term	Explanation
AE	Adverse Event
AFib	Atrial fibrillation
BMI	Body mass index
BNP	B-type natriuretic peptide
CRF	Case Report Form (electronic/paper)
DBP	Diastolic Blood Pressure
DM	Diabetes mellitus
EPS	Enrolled Patients Set
GFR	Glomerular filtration rate
HbA1c	Hemoglobin A1c
HDL	High-Density Lipoprotein
HF	Heart failure
KCCQ	Kansas City Cardiomyopathy Questionnaire
LVEF	Left Ventricular Ejection Fraction
ITTS	Intention to treat set
IP	Investigational Product
MCAR	Missing completely at random
MDRD	Modification of Diet in Renal Disease
NSVT	Nonsustained ventricular tachycardia
NTproBNP	N-terminal (NT)-pro hormone BNP
NYHA	New York Heart Association
PPS	Per protocol set
PAC	Premature atrial contractions
PTDV	Premature treatment discontinuation visit
PVC	Premature ventricular contractions
SAE	Severe Adverse Event
SAFS	Safety Analysis Set
SAS	Statistical Analysis Software
SBP	Systolic Blood Pressure
SD	Standard deviation
SE	Standard error

Abbreviation or special term	Explanation
SGLT-2	Sodium-glucose cotransporter 2
T2DM	Type 2 Diabetes Mellitus
VF	Ventricular fibrillation
VT	Ventricular tachycardia

Contents

1 INTRODUCTION

1.1 Study Objectives

Primary Study Objectives	To compare pulmonary artery diastolic pressure at the end of treatment period (defined as average of pulmonary artery diastolic pressure measurements between weeks 8-12) between empagliflozin and placebo.
Secondary Study Objectives	1. To compare pulmonary artery diastolic pressure between empagliflozin and placebo at each interim time point (weeks 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12) 2. To compare pulmonary artery systolic pressure at the end of treatment period between empagliflozin and placebo 3. To compare pulmonary artery systolic pressure between empagliflozin and placebo at each interim time point (weeks 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12) 4. To compare mean pulmonary artery pressure at the end of treatment period between empagliflozin and placebo 5. To compare mean pulmonary artery pressure between empagliflozin and placebo at each interim time point (weeks 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12) 6. To compare the average of measurements at 6 and 12 weeks in heart failure related health status between empagliflozin and placebo, using the Kansas City Cardiomyopathy Questionnaire (KCCQ) overall summary score. 7. To compare proportion of patients with a $\geq$ 5pts increase from baseline in the Kansas City Cardiomyopathy Questionnaire (KCCQ) overall summary score at 6 and 12 weeks between empagliflozin and placebo 8. To compare 6-minute walk test at follow-up (defined as average of measurements at 6 and 12 weeks) between empagliflozin and placebo 9. To compare NTproBNP at follow-up (defined as average of measurements at 6 and 12 weeks) between empagliflozin and placebo 10. To compare BNP at follow-up (defined as average of measurements at 6 and 12 weeks) between empagliflozin and placebo 11. To compare proportion of patients with a $\geq$ 20% decrease from baseline in NTproBNP at 6 and 12 weeks between empagliflozin and placebo. 12. To compare proportion of patients with a $\geq$ 20% decrease from baseline in BNP at 6 and 12 weeks between empagliflozin and placebo.

	13. To compare proportion of patients with both a ≥ 5pts increase from baseline in KCCQ and a ≥ 20% decrease from baseline in NTproBNP at 6 and 12 weeks between empagliflozin and placebo 14. To compare the number of diuretic medication adjustments during the treatment period (up-titration and down-titration of diuretic doses to be evaluated separately) between empagliflozin and placebo 15. To compare Hemoglobin A1c at follow-up (defined as average of measurements at 6 and 12 weeks) between empagliflozin and placebo (evaluated separately in patients with and without type 2 diabetes)
Exploratory objectives	1. To compare change in pulmonary artery diastolic pressure between week 12 and week 13 between patients originally randomized to empagliflozin or placebo 2. To compare change in mean pulmonary artery pressure between week 12 and week 13 between patients originally randomized to empagliflozin or placebo 3. To compare change in pulmonary artery systolic pressure between week 12 and week 13 between patients originally randomized to empagliflozin or placebo 4. To compare heart rate at the end of treatment period (defined as average measurements between weeks 8-12) between empagliflozin and placebo 5. To compare effects on average weekly loop diuretic dose (furosemide equivalent) between empagliflozin and placebo. 6. To compare effects on hospitalizations for heart failure between empagliflozin and placebo. 7. To compare effects on urgent outpatient heart failure visits between empagliflozin and placebo. 8. To compare effects on hospitalizations for heart failure or urgent outpatient visits for heart failure between empagliflozin and placebo 9. To compare change in NYHA Class at 6 weeks from baseline and 12 weeks from baseline between empagliflozin and placebo. 10. To compare NTproBNP, BNP, 6-minute walk test, Hemoglobin A1c, and KCCQ Overall summary Score, at 6 weeks and 12 weeks separately, between empagliflozin and placebo 11. To compare the number of medication adjustments other than diuretics (nitrates, hydralazine, ACE, ARB, b-blockers, sacubitril/valsartan) during the treatment period between empagliflozin and placebo 12. Proportion of patients that progress to diabetes during the treatment period (within the subgroup of patients without diabetes at baseline only)

1.2 Study Design

Design Configuration and Subject Population	Randomized, double-blind, placebo-controlled trial. The control group will receive placebo administered orally once daily for 12 weeks plus standard of care. The treatment group will receive empagliflozin 10 mg administered orally once daily for 12 weeks plus standard of care. A follow-up visit at week 13 will be performed to evaluate filling pressures and markers of renal function.
Treatment Groups	Empagliflozin 10 mg administered orally vs. placebo
Inclusion criteria	1. Age > 18 and <120 at the screening visit 2. Established diagnosis of heart failure (for at least 16 weeks prior to the screening visit) with either preserved (LVEF>40%) or reduced systolic function (LVEF≤40%), due to either ischemic or non-ischemic etiology, documented by an imaging modality (echocardiography, nuclear imaging, LV angiography, magnetic resonance imaging) within the past 24 months. 3. No major change in diuretic management for 48 hours prior to screening visit or 48 hours prior to randomization visit (major change defined by doubling of diuretic dose or addition of another diuretic medication) 4. NYHA class II, III or IV heart failure symptoms at the screening and randomization visit 5. Presence of previously (≥ 2 weeks prior to screening visit) implanted CardioMEMs pulmonary artery pressure monitor for a clinical indication unrelated to the study. 6. PA diastolic pressure ≥ 12 mmHg at the time of the screening visit (last measurement available prior to the screening visit). 7. Ability to provide informed consent prior to initiating screening visit procedures
Exclusion criteria	1. Decompensated heart failure (hospitalization for heart failure within the 2 weeks prior to screening) or between screening and randomization 2. History of type 1 diabetes 3. Major change in diuretic management during 48 hours prior to screening visit or 48 hours prior to randomization visit. (major change defined by doubling of diuretic dose or addition of another diuretic medications) 4. Significant variability in baseline PA diastolic pressures during screening period. Defined as changes greater than +/- 6 mmHg from average PA diastolic pressure during week 1 of the screening phase and average PA diastolic pressure during week 2 of the screening phase for those patients with an average baseline PA diastolic pressure during

	week 1 of the screening phase of <30 mmHg. If the average baseline PA diastolic pressure during week 1 of the screening phase is ≥30 mmHg, then ≥20% relative change in average PA diastolic pressure between week 1 and week 2 of the screening phase will be used to define significant variability. 5. Initiation of hydralazine, long-acting nitrates, beta blockers, ACEI/ARBs or Valsartan/sacubitril in the prior 4 weeks prior to screening 6. Estimated glomerular filtration rate (eGFR) < 30 mL/min/1.73 m2 at the screening visit 7. Admission for an acute coronary syndrome (ST-elevation MI, non-ST-elevation MI, or unstable angina), percutaneous coronary intervention, or cardiac surgery within 30 days prior to the screening visit. 8. Implantation of cardiac resynchronization therapy (CRT) device within the previous 90 days. 9. Implantation of the CardioMEMs device within the past 2 weeks. 10. Planned cardiovascular revascularization (percutaneous intervention or surgical) or major cardiac surgery (coronary artery bypass grafting, valve replacement, ventricular assist device, cardiac transplantation, or any other surgery requiring thoracotomy), or planned implantation of CRT device within the 90 days after the screening visit. 11. Participation in any interventional clinical trial (with an investigational drug or device) that is not an observational registry within the 4 weeks prior to the screening visit. 12. History of hypersensitivity to empagliflozin 13. For women of child-bearing potential: Current or planned pregnancy or currently lactating 14. Women of childbearing potential are defined as any female who has experienced menarche and who is NOT permanently sterile or postmenopausal. Post menopausal is defined as 12 consecutive months with no menses without an alternative medical cause. Women of child-bearing potential, who are sexually active, must agree to use a medically-accepted method of birth control for the duration of the study. Acceptable birth control methods include: (1) surgical sterilization (such as a hysterectomy or bilateral tubal ligation), (2) progesterone hormonal contraceptives (birth control pills or implants), (3) barrier methods (such as a condom or diaphragm) used with a spermicide, or (4) an intrauterine device (IUD). Women of child-bearing potential will have a urine pregnancy test at every clinic visit and it must be negative to continue study participation. 15. Life expectancy <1 year at the screening visit

	16. Patients who are volume depleted based upon physical examination at the time of the screening or randomization visit 17. PA diastolic pressure < 12 mmHg at the time of the screening visit (average of last four measurements available prior to the screening visit). 18. Patients currently being treated with any SGLT-2 inhibitor (dapagliflozin, canagliflozin, empagliflozin) or having received treatment with any SGLT-2 inhibitor within the 8 weeks prior to the screening visit 19. Average supine systolic BP <90 mmHg at the screening or randomization visit 20. Current documented history of bladder cancer 21. Active Gross Hematuria 22. Heart failure due to restrictive cardiomyopathy, active myocarditis, constrictive pericarditis, severe stenotic valve disease, and HOCM (hypertrophic obstructive cardiomyopathy). 23. History of heart transplant. 24. Patients on heart transplant list as 1a and 1b status

1.3 Sample Size and Power

Planned Sample Size	Approximately 30 patients will be enrolled in each arm.
Power Statement	A sample size of 28 patients for each group will achieve 80% power with α=0.05 to detect a 20% difference in diastolic pulmonary artery pressure between the empagliflozin 10 mg and placebo groups over the treatment period. The assumptions for this calculation were: 1) difference in diastolic pulmonary artery is 4, 2) between-patient standard deviation = 6, 3) Within patient correlation: 0.7, 4) diastolic PA pressure in the intervention arm of the trial will achieve a plateau at the 9th time point for treatment group.

2 GENERAL CONSIDERATIONS FOR DATA ANALYSES

2.1 Analysis Data Sets

Efficacy and safety analyses will be performed on data from the following analysis sets:

- Enrolled patients set (EPS) will consist of all patients who signed the informed consent. This data set will be used to summarize the patient disposition data.

- Modified intention to treat set (modified ITTS) is defined as all patients who have been randomized to study treatment, have received at least one dose of study medication, and have sufficient evaluable data for endpoint ascertainment during follow up. For clarity, patients with no evaluable data for a particular outcome during follow up will be excluded from the analyses of these respective endpoints. For example, for the CardioMems endpoints, patients with no measurements between 8 and12 weeks will be excluded from the analyses. In patients who undergo initiation of intravenous home inotropic support and/or receive left ventricular assist device or heart transplant during the study or who are started on a prohibited medication (open label SGLT-2is), for efficacy analyses, the data point(s) prior to these events will be used. Subjects will be analyzed according to the randomization group. The modified ITTS data set will be used for the primary, secondary, and selected exploratory efficacy endpoints.

- On-treatment set (OTS) includes all subjects in the modified ITTS. Primary and secondary endpoint measurements will be excluded for the follow-up time point(s) when subjects were temporarily or permanently off the study drug at the time the corresponding measurements were obtained. The OTS data set will be used for a sensitivity analyses for the primary, secondary, and selected exploratory efficacy endpoints.

- The per protocol set (PPS) includes all subjects in the modified ITTS who did not have any major protocol deviations. Major protocol deviations, which are detailed in section 2.2, will be determined prior to unblinding of treatment groups. Subjects will be summarized according to actual treatment received regardless of the allocated treatment. This will be used for a sensitivity analysis for the primary efficacy endpoints only.

- The safety analysis set (SAFS) includes all patients who received at least 1 dose of study medication. Throughout the safety results sections, erroneously treated patients (eg, those randomized to empagliflozin but actually given placebo) will be accounted for in the actual treatment group. If a patient received study drug from the wrong kit for only a part of the treatment duration and then switched to another, the associated actual treatment group for that patient will be the treatment group the patient had the longest exposure to. The main safety analyses will be restricted to adverse events that occurred between randomization and the 12-week visit. Adverse events that occurred between 12 weeks and 13 weeks (after discontinuation of study treatment) will be collected, and presented in a separate, supplemental analysis. The safety analysis set will be used to summarize safety data and patient demography and their baseline characteristics, and to analyze selected exploratory endpoints (heart failure hospitalizations, urgent heart failure visits, and a composite of heart failure hospitalizations or urgent heart failure visits). If there is a difference between the SAFS and modified ITTS, baseline and demography data will also be presented for the IITS.

2.2 Protocol Deviations and Major Eligibility Violations

Important Protocol Deviations (IPDs) are defined as those important deviations from the protocol that are likely to have an impact on the efficacy and/or safety of study treatments.

Protocol deviations will be reviewed in a blinded fashion by the study team prior to database lock. All decisions to exclude patients and/or data from the modified ITTS or PPS will be made prior to the unblinding of the study and agreed by the study team.

Error! Reference source not found.9.2 specifies the criteria for IPDs Patients having IPDs will be summarized and listed by treatment group and overall. The list of protocol deviations (including IPDs) will be compiled separately and finalized prior to database lock.

2.3 Strata, Covariates, and pre-specified subgroup analyses

All efficacy and safety endpoints will be analyzed in the entire cohort, and then within the subgroups of patients with and without diabetes. Analyses for the primary and secondary endpoints will be adjusted for the corresponding baseline measurements as well as history of diabetes (DM), eGFR, and age. A sensitivity analysis will be conducted for each endpoint adjusting for the following additional baseline covariates: AFib type (No AFib, persistent/permanent AFib, paroxysmal AFib) and LVEF (as continuous variable). Restricted cubic splines will be used for continuous variables to accommodate non-linear effects.

Additional sensitivity analysis will use the same model as the primary analysis, but adjust only for the corresponding baseline measurements, without additional covariates.

Additionally, the following pre-specified subgroup analyses will be performed for the primary endpoints (stratified by the baseline variables below):

- Baseline LVEF (≤40%, > 40%)
- Atrial fibrillation type (No AFib, permanent/persistent AFib, paroxysmal AFib)
- Age (<65, ≥65)
- Sex (male, female)
- Race (white, black, other)
- Baseline RAASi type (ARNI, ACE/ ARB, neither)
- Loop diuretic dose (Furosemide equivalent mean daily dose: ≤ 40 mg, >40 mg)

Subgroup analyses will be carried out by augmenting the primary analysis model with terms for subgroup and a subgroup-by-treatment interaction. Adjusted point estimates and 95% confidence intervals will be calculated for the effect of empagliflozin compared with placebo within each subgroup. An interaction p-value will be provided.

Due to low number of patients per site, site effects will not be explored.

2.4 Endpoints Definitions

Primary endpoints:

Average pulmonary artery diastolic pressure at the end of treatment period (defined as average of weekly average pulmonary artery diastolic pressure measurements between weeks 8-12)

Secondary endpoints:

1. Weekly average pulmonary artery diastolic pressure at each interim time point (weeks 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12)
2. Average pulmonary artery systolic pressure at the end of treatment period (defined as average of weekly average pulmonary artery systolic pressure measurements between weeks 8-12)
3. Weekly average pulmonary artery systolic pressure at each interim time point (weeks 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12)
4. Average mean pulmonary artery pressure at the end of treatment period (defined as average of weekly average mean pulmonary artery pressure measurements between weeks 8-12)
5. Weekly average mean pulmonary artery pressure at each interim time point (weeks 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12)
6. Heart failure related health status using the Kansas City Cardiomyopathy Questionnaire (KCCQ) overall summary score at follow-up (defined as average of measurements at 6 and 12 weeks).
7. Proportion of patients with a $\geq$ 5pts increase from baseline in the Kansas City Cardiomyopathy Questionnaire (KCCQ) at 6 and 12 weeks.
8. 6-minute walk test at follow-up (defined as average of measurements at 6 and 12 weeks)
9. NTproBNP at follow-up (defined as average of measurements at 6 and 12 weeks)
10. BNP at follow-up (defined as average of measurements at 6 and 12 weeks).
11. Proportion of patients with a $\geq$ 20% decrease from baseline in NTproBNP at 6 and 12 weeks.
12. Proportion of patients with a $\geq$ 20% decrease from baseline in BNP at 6 and 12 weeks.
13. Proportion of patients with both a $\geq$ 5pts increase from baseline in KCCQ and a $\geq$ 20% decrease from baseline in NTproBNP at 6 and 12 weeks.
14. Number of diuretic medication adjustments during the treatment period (up-titration and down-titration of diuretic doses to be evaluated separately)
15. Hemoglobin A1c at follow-up (defined as average of measurements at 6 and 12 weeks), evaluated separately in patient with and without type 2 diabetes

Exploratory Outcome Variables

1. Change in pulmonary artery diastolic pressure between week 12 and week 13
2. Change in mean pulmonary artery pressure between week 12 and week 13
3. Change in pulmonary artery systolic pressure between week 12 and week 13
4. Average heart rate from at the end of treatment period
5. Effects on average weekly loop diuretic dose (furosemide equivalent).

   The following equations will be used to convert doses to Furosemide equivalent:

   40 mg Furosemide = 20 mg Torsemide = 2 mg Bumetanide = 50 mg Ethacrynic Acid

   When patient is not on a loop diuretic, dose = 0 mg.

   For subjects with dosing schedule reported as "as needed (prn)", we will assume a twice a week dosing regimen for loop diuretic at the recorded dose.

   In addition to comparing daily loop diuretic dose between treatment arms, we will also compare the proportion of patients who had their loop diuretic dose reduced or discontinued during the treatment period.
6. Effects on hospitalizations for heart failure.

   Only events positively adjudicated by the clinical event committee will be included. Events that occurred between screening and randomization, and between 12 weeks and 13 weeks will be excluded from this analysis. Events that occurred between the week 12 and week 13 study visits will be reported in a separate, supplemental analysis.
7. Effects on urgent outpatient heart failure visits.

   Only events positively adjudicated by the clinical event committee will be included. Events that occurred between screening and randomization, and between 12 weeks and 13 weeks will be excluded from this analysis. Events that occurred between the week 12 and week 13 study visits will be reported in a separate, supplemental analysis.
8. Effects on hospitalizations for heart failure and urgent outpatient visits for heart failure

   Only events positively adjudicated by the clinical event committee will be included. Events that occurred between screening and randomization, and between 12 weeks and 13 weeks will be excluded from this analysis. Events that occurred between the week 12 and week 13 study visits will be reported in a separate, supplemental analysis.
9. Change in NYHA Class at 6 weeks from baseline and 12 weeks from baseline.

   Change in NYHA class is defined as a 3-level categorical variable: 1) Decrease, 2) No change, 3) Increase. For patients missing NYHA class at 12 weeks, value at 6 weeks will be used.
10. Effect on NTproBNP, BNP, 6-minute walk test, Hemoglobin A1c, and KCCQ overall summary score at 6 and 12 weeks
11. Number of medication adjustments other than diuretics (nitrates, hydralazine, ACE, ARB, b-blockers, sacubitril/valsartan) during the treatment period
12. Proportion of patients that progress to diabetes during the treatment period (within the subgroup of patients without diabetes at baseline only)

2.5 Multiple Testing

The primary hypothesis will be tested at the 2-sided 5% significance level. No adjustments for multiplicity will be made for secondary and exploratory endpoints.

2.6 Missing Data

For CardioMEMS sensor data (pulmonary artery diastolic pressure, pulmonary artery systolic pressure, and mean pulmonary artery pressure), patients with no measurements in the two weeks between screening and randomization and patients with no measurements at the end of the study (8-12 weeks) will be excluded from analyses. In addition, CardioMEMS sensor data (pulmonary artery diastolic pressure, pulmonary artery systolic pressure, and mean pulmonary artery pressure), will be examined by data coordinating center staff in a blinded manner prior to data lock, and measurements confirmed to be erroneous (e.g., non-physiologic), will be excluded from analyses. Repeated measure models using maximum likelihood will be used to accommodate missed measurements under the assumption that data are missing at random conditional on the covariates and on the observed sensor data.

For other endpoints that are measured at 6-week and 12-week follow-up visits (e.g. NT pro-BNP, KCCQ, etc.), a premature treatment discontinuation visit (PTDV) will be scheduled for patient who drop out of the study prematurely, whenever possible, to get the last assessment. The PTDV assessment will be carried forward to the next follow-up time point. Analyses will be conducted among patients with a baseline assessment and with at least one follow-up assessment (including the PTDV measurements) using repeated measure models.

For partially completed KCCQs, the scoring algorithm accommodates a limited number of skipped responses.

Missing data due to deaths is expected to be very low (fewer than 5 patients) because of the short follow-up period and will be ignored in the primary analysis.

2.7 Data Handling Conventions and Transformations

Diabetes duration will be calculated as (Consent date – DM diagnosis date) / 365.25. It will be rounded to the whole year.

Body mass index (BMI) will be calculated as weight (kg) / [height (m)]2. It will be displayed to 1 decimal place in listings, but will not be rounded or truncated prior to summarization

Estimated glomerular filtration rate (eGFR) will be calculated using the MDRD-4 equation: GFR in mL/min per 1.73 m2 = 175 x SerumCr-1.154 x age-0.203 x 1.212 (if patient is African American) x 0.742 (if female). Patients with missing data in Creatinine Serum, age, race, or gender will not have eGFR calculated.

If urine microalbumin value is reported as being too high for the laboratory to calculate a UACR, 5000 mg value for UACR will be used.

3 Analyses of Baseline Characteristics

Patient demographics, baseline clinical characteristics, medical histories, and baseline labs will be described overall and by treatment group. Continuous measures will be summarized by mean $\pm$ standard deviation and compared using Student's T-tests. Categorical variables will be summarized by frequency and percent and compared using χ^2or Fisher's exact tests, as appropriate.

3.1 Demographics and Baseline Characteristics

- Age
- Sex
- Race/Ethnicity

3.2 Medical History

3.2.1 Diabetes History

- Diabetes duration
- History of diabetic peripheral neuropathy
- History of diabetic autonomic neuropathy
- History of diabetic retinopathy
- History of diabetic ketoacidosis (DKA) event
- History of hyperosmolar hyperglycemic syndrome (HHS) event
- History of severe hypoglycemic event(s)
- History of amputation

3.2.2 Other Medical History

- History of heart failure
- NYHA Class
- Most Recent LVEF Assessment
- History of PAD
- History of hypertension
- History of coronary artery disease
- History of dyslipidemia
- History of angina
- History of atrial fibrillation
- History of atrial flutter
- History of MI

- History of PCI
- History of CABG
- History of ventricular tachycardia
- ICD implanted
- Permanent pacemaker implanted
- History of valve disease

3.3 Physical examination

- Body Mass Index
- Sitting Pulse and Blood Pressure
- Supine Pulse and Blood Pressure
- Standing Pulse and Blood Pressure

3.4 Lab results

- HbA1c
- BNP
- NT pro-BNP
- Glucose (mg/dL)
- BUN - urea nitrogen (mg/dL)
- Creatinine (mg/dL)
- Estimated Glomerular Filtration Rate (eGFR)
- Sodium (mmol/L)
- Potassium (mmol/L)
- Chloride (mmol/L)
- Carbon Dioxide (mmol/L)
- Calcium (mg/dL)
- Phosphate - as phosphorus (mg/dL)
- Albumin (g/dL)
- Random Urine Creatinine (mg/dL)
- Random Urine Microalbumin < 0.2 mg/dL
- Random Urine Microalbumin (mg/dL)
- Enter Urine Albumin result, if measurable.
- Random Urine Albumin/Creatinine Ratio (mcg/mg creat)

4 SUBJECT DISPOSITION

4.1 Subject Enrollment

The number and percent of subjects randomized for each site will be summarized overall and by treatment group. The denominator for the percent calculation will be number of subjects screened.

4.2 Disposition of Subjects

A summary of subject disposition will be provided overall and by treatment group, as appropriate. This summary will present the number of subjects screened, randomized, included in the safety analysis set, and the number and percent of subjects meeting the following criteria:

- Completed the study drug treatment,
- Did not complete the study drug treatment (with summary of reasons for not completing the study treatment)
- Completed the study (includes the study treatment period and any post treatment follow up period), and
- Did not complete the study, (with summary of reasons for not completing the study).

The denominator for the percent of subjects in each category will be the number of subjects in the safety analysis set.

No inferential statistics will be generated. A data listing of reasons for premature study treatment/study discontinuation will be provided.

4.3 Extent of Exposure

4.3.1 Duration of Exposure to Study Drug

Duration of exposure to study drug will be defined as (last dose date – first dose date of double-blind treatment phase + 1), regardless of temporary interruptions in study drug administration, and will be expressed in weeks (shown to one decimal place, e.g., 4.5 weeks).

Duration of exposure to study drug will be summarized using descriptive statistics (sample size, mean, standard deviation, median, first quartile (Q1), third quartile (Q3), minimum, and maximum).

Summaries will be provided by treatment group for the safety analysis set.

4.3.2 Adherence with Study Drug

- The actual number of tablets ingested will be calculated using the Study Drug Accountability CRF:

  Total number of tablets dispensed - total number of tablets returned.

- The expected number of tablets ingested will be calculated based on the Study Drug Dispensed CRF and Study Drug Accountability CRF:

  Total number of days supposed on study drug ([date of last dose of study drug - start date + 1])

- Adherence rate
  100(total number of tablets ingested) / (expected number of tablets ingested).

Adherence will be capped at 100% for summarization. Patients with unreturned bottles will be excluded from adherence calculations.

Descriptive statistics for adherence (sample size, mean, standard deviation, median, Q1, Q3, minimum, and maximum) will be provided by treatment group and overall for the safety analysis set.

5 EFFICACY ANALYSES

5.1 Analysis of the Primary Efficacy Endpoint

Analysis of the primary efficacy endpoints will be performed on the modified ITT data set, first on the entire cohort, then within the subgroups of subjects with and without diabetes, and then within other subgroups as specified in section 2.3. Several sensitivity analyses will be performed as outlined in section 2.3 for the primary efficacy endpoint; in addition, supportive analyses will be repeated using the same models on the on-treatment set (OTS) and per-protocol set (PPS).

Mean, standard deviation, median, and interquartile range (IQR) will be reported for weekly average pulmonary artery diastolic pressure at each week (weeks 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12) and the average of measurements between weeks 8 and 12, for the entire cohort and by treatment group.

A generalized linear mixed model with identity link and Normal distribution will be used to estimate the effect of treatment on the average of pulmonary artery diastolic pressure between week 8 and 12, adjusting for baseline measurement, history of DM, eGFR, and age. Assessment time t (in weeks), defined as number of days since randomization divided by 7, will be included as restricted cubic splines with 4 knots to capture the nonlinear trajectories. Random intercept and random coefficients for cubic splines will be included for subjects. The model is as follows:

$$y_{it} = \alpha_{0i} + \boldsymbol{\alpha}_{1i}\boldsymbol{f}(t) + \beta_1 Trt_i + \boldsymbol{\beta}_2\boldsymbol{f}(t) * Trt_i + X\boldsymbol{\beta} + \epsilon_{ij}$$

$$\alpha_{0i} = \beta_{00} + \gamma_i$$

$$\boldsymbol{\alpha}_{1i} = \boldsymbol{\beta}_{10} + \boldsymbol{\eta}_i$$

$\epsilon_{ij} \sim N(0, \sigma^2)$, $\begin{pmatrix} \gamma_i \\ \boldsymbol{\eta}_i \end{pmatrix} \sim N(0, \boldsymbol{\Sigma})$

where, y_{it} denotes the diastolic pressure for subject i at week t, Trt_i is a 0/1 variable denoting treatment group (empagliflozin vs. placebo), t indicates the follow-up assessment time (t, a real number between 1 and 14) and $\boldsymbol{f}(t)$ is a vector of spline basis functions, and X is the design matrix for baseline measurement, age, DM, and baseline eGFR. γ_i are subject random intercepts $\boldsymbol{\eta_i}$ are subject random coefficient for cubic spline function $\boldsymbol{f}(t)$. With this parameterization, the quantity $\beta_1 + \boldsymbol{\beta}_2\boldsymbol{f}(t)$ represents the expected effect of empagliflozin vs. placebo on pulmonary artery diastolic pressure at week t, and the quantity $\frac{\sum_{t=8}^{12}(\beta_1+\boldsymbol{\beta}_2\boldsymbol{f}(t))}{5}$ represents the expected effect on average diastolic pressure between 8 and 12 weeks. Appropriate simplification might be necessary to ensure model convergence.

A sensitivity analysis will be performed using the same model additionally adjusting for AFib type (No AFib, persistent/permanent AFib, paroxysmal AFib) and baseline LVEF.

Additional sensitivity analysis will use the same model as the primary analysis, but adjust only for the corresponding baseline measurements, without additional covariates.

Finally, another sensitivity analysis will repeat the primary analysis on the on-treatment set (OTS).

5.2 Secondary outcome variables

The following secondary outcomes will be analyzed on the intention to treat (modified ITT) data set, first on entire subject cohort and then within subgroups of subject with or without diabetes.

1. **Weekly average pulmonary artery diastolic pressure at each interim time point (weeks 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12)**

   Using the same model as that of the primary endpoint, the effect of empagliflozin vs. placebo at each interim time point will be estimated by $\beta_2 + \boldsymbol{\beta}_3\boldsymbol{f}(t)$. The estimated average pulmonary artery diastolic pressure at each time point will be obtained using Least Square Means for empagliflozin and placebo separately.

2. **Weekly average pulmonary artery systolic pressure at the end of treatment period (defined as average of pulmonary artery systolic pressure measurements between weeks 8-12)**

3. **Weekly average pulmonary artery systolic pressure at each interim time point (weeks 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12)**

4. **Weekly average mean pulmonary artery pressure at the end of treatment period (defined as average of mean pulmonary artery pressure measurements between weeks 8-12)**
5. **Weekly average mean pulmonary artery pressure at each interim time point (weeks 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12)**

   Pulmonary artery systolic pressure and mean pulmonary artery systolic pressure will be analyzed in a manner analogous to that of pulmonary artery diastolic pressure.
6. **Heart failure related health status using the Kansas City Cardiomyopathy Questionnaire (KCCQ) overall summary score at follow-up (defined as average of measurements at 6 and 12 weeks).**

   Mean, standard deviation, median, and interquartile range (IQR) will be reported for KCCQ overall summary score and its change from baseline at 6 and 12 weeks.

   A generalized linear mixed model with identity link and Normal distribution will be used to estimate the effect of treatment on average KCCQ overall summary score at 6 and 12 weeks, adjusting for baseline measurement, history of DM, eGFR, and age. Random intercept will be included for subjects. The model is as follows:

$$\mu_{it} = \beta_{0i} + \beta_1 \mathrm{Trt_i} + \beta_2 t + \beta_3 \mathrm{Trt_i} * \mathrm{t} + \mathrm{X}\boldsymbol{\beta} + \epsilon_{ij},$$

$$\beta_{0i} = \beta_{00} + \gamma_i$$

$$\epsilon_{ij} \sim \mathrm{N}(0, \sigma^2),\ \gamma_i \sim N(0, \sigma_\gamma^2),$$

   where, μ_{it} denotes the expected KCCQ overall summary score for subject i at week t, $\mathrm{Trt_i}$ is a 0/1 variable denoting treatment group (empagliflozin vs. placebo), t indicates the follow-up assessment time (week 6, t =0; week 12, t = 1), and X is the design matrix for baseline measurement, age, DM, and baseline eGFR. γ_i are subject random intercepts with variance σ_γ^2. With this parameterization, effect of empagliflozin vs. placebo is estimated by β_1 at 6 weeks and $\beta_1 + \beta_2$ at 12 weeks. The quantity $\beta_1 + \beta_3/2$ represents the expected average effect of empagliflozin vs. placebo on KCCQ summary score at 6 and 12 weeks.

   Similar analyses will be repeated for KCCQ clinical summary score.
7. **Proportion of subjects with a ≥ 5pts increase from baseline in the Kansas City Cardiomyopathy Questionnaire (KCCQ) at 6 and 12 weeks separately.**

   Unadjusted proportions of subjects with a ≥ 5 pts increase from baseline in KCCQ overall score at 6 and 12 weeks, will be reported for the treatment group and placebo group. A logistic regression model will be used to assess the effect of treatment on the proportion of subjects with a ≥ 5 pts increase. Models will be adjusted for baseline KCCQ overall summary score, history of DM, eGFR, and age. The model is specified as follows:

$$\mathrm{logit}(\mathrm{p_i}) = \beta_0 + \beta_1 \mathrm{Trt_i} + \boldsymbol{\beta}\mathrm{X}$$

where p_i denotes the expected probability for subject i with a $\geq$ 5 pts increase in KCCQ overall score, Trt_i is a 0/1 variable denoting treatment group (empagliflozin vs. placebo), X is the design matrix for log baseline NT pro-BNP, baseline KCCQ summary score, history of DM, eGFR, and age. With this parameterization, the quantity $\exp(\beta_1)$ represents the odds ratio of achieving a $\geq$ 5 pts increase in the treatment group vs. the placebo group.

Additionally, responder analysis will be conducted to report proportion of patients with deterioration (5 points or greater decrease), no change (change between -5 to 5 points), small-moderate improvement (5 to 9 points increase), moderate-large improvement (10 to 19 points increase), and very large improvement (20 points or greater increase) in KCCQ overall summary score at 6 and 12 weeks. Similar analyses will be repeated for KCCQ clinical summary score.

8. **6 minute walk test at follow-up (defined as average of measurements at 6 and 12 weeks)**

   6 minute walk test will be analyzed in a manner analogous to that of secondary outcome 6. Effect of empagliflozin vs. placebo is estimated by β_1 at 6 weeks and $\beta_1 + \beta_2$ at 12 weeks. The quantity $\beta_1 + \beta_3/2$ represents the expected average effect of empagliflozin vs. placebo on KCCQ summary score at 6 and 12 weeks

9. **NTproBNP at follow-up (defined as average of measurements at 6 and 12 weeks)**

   NTproBNP will be analyzed in a manner analogous to that of secondary outcome 6. Although Gama distribution and log link will be used to account for the skewed nature of NTproBNP.

10. **BNP at follow-up (defined as average of measurements at 6 and 12 weeks).**

    BNP will be analyzed in a manner analogous to that of NTproBNP.

11. **Proportion of subjects with a ≥ 20% decrease from baseline in NTproBNP at 6 and 12 weeks separately.**

12. **Proportion of subjects with a ≥ 20% decrease from baseline in BNP at 6 and 12 weeks separately.**

13. **Proportion of subjects with both a ≥ 5pts increase from baseline in KCCQ and a ≥ 20% decrease from baseline in NTproBNP at 6 and 12 weeks separately.**

    Outcome 11, 12, and 13 will be analyzed in a manner analogous to that of secondary outcome 7.

14. **Number of diuretic medication adjustments during the treatment period (up-titration and down-titration of diuretic doses to be evaluated separately)**

    Number of diuretic medication adjustments, up-titrations, and down-titrations will be reported using mean, standard deviation, median, and interquartile range (IQR), and will be compared using Poison regression between the treatment and placebo group.

    Proportion of patients with no diuretic medication adjustments, up-titrations, and down-titrations will be reported using mean and standard deviation, and will be compared using Student's T-test between the treatment and placebo group.

15. **Hemoglobin A1c at follow-up (defined as average of measurements at 6 and 12 weeks), evaluated separately in subject with and without type 2 diabetes**

    HbA1c will be analyzed in a manner analogous to that of secondary outcome 6.

    In addition, proportion of patients with HbA1c $\geq$ 6.5 at 6 weeks, 12 weeks, and at both 6 and 12 weeks will be reported within the subgroups of patients with and without Diabetes.

5.3 Exploratory outcome variables

1. **Change in pulmonary artery diastolic pressure between week 12 and week 13**

   Using the same model for the primary endpoint, the effect of empagliflozin vs. placebo on change between week 12 and week 13 will be estimated by $\boldsymbol{\beta}_2(\boldsymbol{f}(13) - \boldsymbol{f}(12))$

2. **Change in mean pulmonary artery pressure between week 12 and week 13**

   Will be analyzed in a manner analogous to that of pulmonary artery diastolic pressure.

3. **Change in mean pulmonary artery systolic pressure between week 12 and week 13**

   Will be analyzed in a manner analogous to that of pulmonary artery diastolic pressure.

4. **Change in average heart rate from at the end of treatment period**

   Will be analyzed in a manner analogous to that of pulmonary artery diastolic pressure.

5. **Effects on average weekly loop diuretic dose (furosemide equivalent).**

6. Daily loop diuretic dose will be analyzed in a manner analogous to that of secondary outcome 6, although appropriate distributions and link functions will be chosen for the given outcomes. Additionally, among patients who were on loop diuretic at randomization, we will also compare the proportion of patients who had loop diuretic dose reduced or discontinued (using the average reported dose at 6 weeks and 12 weeks). **Effects on hospitalizations for heart failure.**

7. **Effects on urgent outpatient heart failure visits.**

8. **Effects on hospitalizations for heart failure and urgent outpatient visits for heart failure**

   A Cox proportional hazard model will be used to assess the effect of treatment vs. placebo on the time to the first occurrence for outcomes 5, 6, and 7, adjusting for history of DM, eGFR, and age. A sensitivity analysis will be conducted additionally adjusting for AFib type (No AFib, persistent/permanent AFib, paroxysmal AFib) and baseline LVEF. A second sensitivity analysis will be performed using the same approach, but without adjustment for covariates

9. Change in NYHA Class at 6 weeks from baseline and 12 weeks from baseline.

10. **Effect on NTproBNP, BNP, 6-minute walk test, Hemoglobin A1c, and KCCQ overall summary score at 6 weeks and 12 weeks between empagliflozin and placebo.**

NTproBNP, BNP, 6-minute walk test, Hemoglobin A1c, and KCCQ overall summary score will be analysed in a manner analogous to that of secondary outcome 6. Effect of empagliflozin vs. placebo is estimated by β_1 at 6 weeks and $\beta_1 + \beta_2$ at 12 weeks.

11. **Number of medication adjustments other than diuretics (nitrates, hydralazine, ACE, ARB, b-blockers, sacubitril/valsartan) during the treatment period**
12. **Proportion of subjects that progress to diabetes during the treatment period (within the subgroup of subjects without diabetes at baseline only)**

    Subjects that progress to diabetes will be summarized by frequency and percent and compared using χ^2or Fisher's exact tests, as appropriate

5.4 Supplemental pre-specified analyses.

For all outcomes that include assessment of KCCQ overall summary score, additional sensitivity analyses will be performed using KCCQ Clinical Summary Score, Social Limitation, Total Symptom Score, Physical Limitation, Quality of Life, in a manner analogous to that of the KCCQ overall summary score.

In addition, additional subgroup analyses (in addition to stratification by DM status, using strata specified in section 2.3) may be repeated for all secondary efficacy outcomes.

Furthermore, as a supportive analysis, distributional regressions may be conducted to assess the effect of empagliflozin on variability of pulmonary diastolic and systolic pressure.

6 SAFETY ANALYSES

Safety analyses will be performed on the safety analysis set (SAF). Total number of adverse events as well as number and proportion of subjects developing adverse event(s) will be compared by treatment group. For subject level analyses multiple events will be counted once only per subject in each summary. The following safety variables will be included:

1. All cause death
2. Cardiovascular death
3. Non-fatal myocardial infarction (MI)
4. Stroke
5. Acute kidney injury (defined as doubling of serum creatinine based on the modified RIFLE criteria)
6. Adverse events (AEs).

    - Adverse events of special interest
        - Diabetic Ketoacidosis
        - Volume Depletion Event (defined as hypotension, syncope, orthostatic hypotension or dehydration)
        - Severe Hypoglycemic Event
        - Acute Kidney Injury

    - Non-traumatic Lower Limb Amputations
- Drug Adverse Event
- Serious Adverse event
    - Resulted in death
    - In-patient hospitalization or prolonging of existing in-patient hospitalization
    - Persistent or significant disability
    - Life-threatening
    - Congenital anomaly/birth defect
    - Important medical event
- Other adverse events
    - Urgent outpatient heart failure visits

Safety analyses will be restricted to adverse events that occurred between randomization and 12 weeks. Adverse events occurred between 12 weeks and 13 weeks will be presented separately in a supplemental analysis.

7 REFERENCES

1. Green CP1, Porter CB, Bresnahan DR, Spertus JA., Development and Evaluation of the Kansas City Cardiomyopathy Questionnaire: A New Health Status Measure for Heart Failure, J Am Coll Cardiol. 2000 Apr;35(5):1245-55
2. Hunger,M et all, Longtitudinal beta regression models for analyzing health-related quality of life scores over time Medical Research Methodology, 2012; 12:144

8 SOFTWARE

All analyses will be performed using SAS 9.4 or higher.

9 APPENDICES

9.1 Scoring and Interpreting the KCCQ

There are 10 summary scores within the KCCQ, which are calculated as follows:

A. Physical Limitation

The Physical Limitation score corresponds to questions 1a through 1f. Responses to questions 1a through 1f should be coded numerically as follows:

1 = Extremely Limited
2 = Quite a bit Limited
3 = Moderately Limited
4 = Slightly Limited

5 = Not at all Limited
6 = Limited for other reasons or did not do the activity

If the responses to questions 1a through 1f are not values 1, 2, 3, 4 or 5 then the response is set to missing. Note that a response of 6 (Limited for other reasons or did not do the activity) is treated as a missing value. If at least three responses to questions 1a-1f are not missing, then the physical limitation score is computed by calculating the mean response and standardizing the result as follows:

$$\text{Physical Limitation} = 100*(\text{Mean Response} - 1)/4$$

B. Symptom Stability

The Symptom Stability score corresponds to question 2. Responses to question 2 should be coded numerically as follows:

1 = Much Worse
2 = Slightly Worse
3 = Not Changed
4 = Slightly Better
5 = Much Better
6 = I've had no symptoms over the last 2 weeks

If the response is 6 (no symptoms over last 2 weeks) then set the response to 3 (not changed). If question 2 is not missing then the symptom stability score is computed by standardizing the result as follows:

$$\text{Symptom Stability} = 100*(\text{Response} - 1)/4$$

C. Symptom Frequency

The Symptom Frequency score corresponds to questions 3, 5, 7 and 9. The responses should be coded sequentially (1, 2, 3…) in order of increasing health status as follows:

<u>Question 3</u>

1 = Every Morning
2 = 3 or more times per week, but not every day
3 = 1-2 times a week
4 = Less than once a week
5 = Never over the past 2 weeks

<u>Questions 5 and 7</u>

1 = All of the time
2 = Several times per day
3 = At least once a day
4 = 3 or more times per week, but not every day
5 = 1-2 times per week
6 = Less than once a week
7 = Never over the past 2 weeks

Question 9

1 = Every night
2 = 3 or more times a week, but not every day
3 = 1-2 times a week
4 = Less than once a week
5 = Never over the past 2 weeks

If two or more responses are missing then symptom frequency cannot be computed and will be missing. Otherwise, the symptom frequency is computed by calculating the mean of the standardized responses and multiplying by 100 as follows:

$$\text{Symptom Frequency} = 100*\text{Mean}((Q3 - 1)/4, (Q5 - 1)/6, (Q7 - 1)/6, (Q9 - 1)/4)$$

D. Symptom Burden

The Symptom Burden score corresponds to questions 4, 6 and 8. The responses should be coded numerically as follows:

1 = Extremely Bothersome
2 = Quite a bit Bothersome
3 = Moderately Bothersome
4 = Slightly Bothersome
5 = Not at all Bothersome
6 = I've had no swelling (fatigue, shortness of breath)

If a response is 6 (none) then set the response to 5 (not at all). If at least one response is present then symptom burden score is computed by calculating the mean response and standardizing the result as follows:

$$\text{Symptom Burden} = 100*(\text{Mean Response} - 1)/4$$

E. Total Symptom Score

The total symptom score is calculated as the mean of the symptom frequency score and symptom burden score.

F. Self-Efficacy

The Self-Efficacy score corresponds to questions 10 and 11. Responses to questions 10 and 11 should be coded sequentially (1, 2, 3, 4, 5) in order of increasing health status, with 1 denoting the response associated with the lowest health status. If at least one question response is present then the self-efficacy score may be computed by standardizing the mean response as follows:

$$\text{Self-Efficacy} = 100*(\text{Mean Response} - 1)/4$$

G. Quality of Life

The Quality of Life score corresponds to questions 12, 13 and 14. Responses to questions 12, 13 and 14 should be coded sequentially (1, 2, 3, 4, 5) in order of increasing health status, with 1 denoting the response associated with the lowest health status. If at least one question response is present then the quality of life score may be computed by standardizing the mean response as follows:

$$\text{Quality of Life} = 100*(\text{Mean Response} - 1)/4$$

H. Social Limitation

The Social Limitation score corresponds to questions 15a through 15d. These responses should be coded numerically as follows:

1 = Severely Limited
2 = Limited Quite a bit
3 = Moderately Limited
4 = Slightly Limited
5 = Did Not Limit at All
6 = Does not apply or did not do for other reasons

If the responses to questions 15a through 15d are not values 1, 2, 3, 4 or 5 then the response is set to missing. Note that a response of 6 is treated as a missing value. If at least two question responses are present then the social limitation score may be computed by standardizing the mean response as follows:

Social Limitation = 100*(Mean Response – 1)/4

I. Clinical Summary Score

The clinical summary score is calculated as the mean of the physical limitation score and total symptom score.

J. Overall Summary Score

The overall summary score is calculated as the mean of the physical limitation score, total symptom score, quality of life score and social limitation score.

9.2 Protocol deviations

Number	Important Protocol Deviations Criteria	Major Impact on Analysis (MIA)/Other	Exclusion level
1. Did not fulfil eligibility criteria – inclusion criteria:			
1.1	Ability to provide informed consent prior to initiating screening visit procedures	MIA	Complete exclusion from PPS and MODIFIED ITTS
1.2	Age > 18 and < 120 at the screening visit	Other	No exclusion
1.3	Established diagnosis of heart failure	Other	To be decided on a case by case basis
1.4	No major change in diuretic management for 48 hours prior to screening visit or 48 hours prior to randomization visit (major change defined by doubling of diuretic dose or addition of another diuretic medication)	Other	To be decided on a case by case basis

Number	Important Protocol Deviations Criteria	Major Impact on Analysis (MIA)/Other	Exclusion level
1.5	NYHA class II or III heart failure symptoms at the screening and randomization visit	Other	To be decided on a case by case basis
1.6	Presence of previously (≥ 2 weeks prior to screening visit) implanted CardioMEMs pulmonary artery pressure monitor for a clinical indication unrelated to the study.	Other	To be decided on a case by case basis
1.7	PA diastolic pressure ≥ 12 mmHg at the time of the screening visit (last measurement available prior to the screening visit).	Other	To be decided on a case by case basis
2. Did not fulfil eligibility criteria – exclusion criteria:			
2.0	Decompensated heart failure (hospitalization for heart failure within the 2 weeks prior to screening) or between screening and randomization	Other	To be decided on a case by case basis
2.1	History of type 1 diabetes	Other	To be decided on a case by case basis
2.2	Major change in diuretic management during 48 hours prior to screening visit or 48 hours prior to randomization visit. (major change defined by doubling of diuretic dose or addition of another diuretic medications)	Other	To be decided on a case by case basis
2.3	Significant variability in baseline PA diastolic pressures during screening period.	Other	To be decided on a case by case basis
2.4	Initiation of hydralazine, long-acting nitrates, beta blockers, ACEI/ARBs or Valsartan/sacubitril in the prior 4 weeks prior to screening	Other	To be decided on a case by case basis

Number	Important Protocol Deviations Criteria	Major Impact on Analysis (MIA)/Other	Exclusion level
2.5	Estimated glomerular filtration rate (eGFR) < 30 mL/min/1.73 m2 at the screening visit	Other	To be decided on a case by case basis
2.6	Admission for an acute coronary syndrome (ST-elevation MI, non-ST-elevation MI, or unstable angina), percutaneous coronary intervention, or cardiac surgery within 30 days prior to the screening visit	Other	To be decided on a case by case basis
2.7	Implantation of cardiac resynchronization therapy (CRT) device within the previous 90 days.	Other	To be decided on a case by case basis
2.8	Implantation of the CardioMEMs device within the past 2 weeks.	Other	To be decided on a case by case basis
2.9	Planned cardiovascular revascularization (percutaneous intervention or surgical) or major cardiac surgery (coronary artery bypass grafting, valve replacement, ventricular assist device, cardiac transplantation, or any other surgery requiring thoracotomy), or planned implantation of CRT device within the 90 days after the screening visit.	Other	To be decided on a case by case basis
2.10	Participation in any interventional clinical trial (with an investigational drug or device) that is not an observational registry within the 4 weeks prior to the screening visit.	Other	To be decided on a case by case basis
2.11	History of hypersensitivity to empagliflozin	Other	To be decided on a case by case basis

Number	Important Protocol Deviations Criteria	Major Impact on Analysis (MIA)/Other	Exclusion level
2.12	For women of child-bearing potential: Current or planned pregnancy or currently lactating.	Other	To be decided on a case by case basis
2.13	Life expectancy <1 year at the screening visit	Other	To be decided on a case by case basis
2.14	Patients who are volume depleted based upon physical examination at the time of the screening or randomization visit	Other	To be decided on a case by case basis
2.15	PA diastolic pressure < 12 mmHg at the time of the screening visit (average of last four measurements available prior to the screening visit).	Other	To be decided on a case by case basis
2.16	Patients currently being treated with any SGLT-2 inhibitor (dapagliflozin, canagliflozin, empagliflozin) or having received treatment with any SGLT-2 inhibitor within the 8 weeks prior to the screening visit	Other	To be decided on a case by case basis
2.17	Average supine systolic BP <90 mmHg at the screening or randomization visit	Other	To be decided on a case by case basis
2.18	Current documented history of bladder cancer	Other	To be decided on a case by case basis
2.19	Active Gross Hematuria	Other	To be decided on a case by case basis

Number	Important Protocol Deviations Criteria	Major Impact on Analysis (MIA)/Other	Exclusion level
2.20	Heart failure due to restrictive cardiomyopathy, active myocarditis, constrictive pericarditis, severe stenotic valve disease, and HOCM (hypertrophic obstructive cardiomyopathy).	Other	To be decided on a case by case basis
2.21	History of heart transplant.	Other	To be decided on a case by case basis
2.22	Patients on heart transplant list as 1a and 1b status	Other	To be decided on a case by case basis

3. Subject developed discontinuation of investigational product criteria but dosing continued			
3.0	Subject experienced an Adverse Event which in the opinion of the Investigator, contraindicated further dosing	Other	To be reviewed on the case by case basis
3.1	Pregnancy confirmed by a positive pregnancy test or other examinations	Other	Complete exclusion from PPS
3.2	Donation of blood or bone marrow during the study period	Other	To be decided on a case by case basis
3.3	Admission for an acute coronary syndrome (ST-elevation MI, non-ST-elevation MI, or unstable angina), percutaneous coronary intervention, or cardiac surgery during the study period.	Other	To be decided on a case by case basis
4. Subject received prohibited concomitant medication but study treatment not discontinued:			
4.4.1	SGLT2	Other	Only the efficacy endpoints before the start of SGLT2 will be included in modified IITS, OTS, and PPS
5. Subject received incorrect investigational treatment/dose			
5.0	Subject took incorrect treatment or damaged Investigational Drug Kit	MIA	To be decided on a case by case basis
5.1	Major compliance issues <80% or >120% compliance with IP dosing in double blind treatment period	MIA	To be decided on a case by case basis
5.3	Subject randomized but never took IP	MIA	To be decided on a case by case basis
5.4	Subject randomized and stopped taking IP for a long period of time after the initial dose, and then resumed IP	MIA	To be decided on a case by case basis

5.5	Overdose	MIA	To be decided on the case by case basis
6. Protocol-required procedure not adhered to:			
6.0	Study data collection has not been stopped when subject decided to withdraw their consent from the study completely	Other	No exclusion from PPS Data collected after consent withdrawal to be excluded from data base
6.1	Subjects with absence of NT-proBNP data at randomization visit and having values at Screening visit.	Other	No exclusion from PPS Values from screening visit will be used to impute the values at randomization visit.
6.2	Subjects with absence of baseline NT-proBNP data at both screening and randomization visits	Other	Excluded from the NT-proBNP analyses
6.3	Subjects with no NT-proBNP values after randomization	Other	Excluded from the NT-proBNP analyses
6.4	Subjects with absence of baseline KCCQ summary score at randomization visit	Other	Excluded from the KCCQ analyses
6.5	Subjects with no KCCQ overall summary score after randomization	Other	Excluded from KCCQ analyses
6.6	Subjects with absence of baseline 6-minute walk assessment at randomization visit	Other	Excluded from the 6-minute walk analyses
6.7	Subjects with no 6-minute walk assessment after randomization	Other	Excluded from 6-minute walk analyses
6.8	Subjects with missing values in the covariates: age, gender, Hx AFib, AFib type, Hx DM, eGFR at randomization, lvef at randomization	Other	No exclusion from PPS Excluded from the first and second primary analysis

6.9	Subject with absence of CardioMEMS sensor measurements (pulmonary artery diastolic pressure, pulmonary artery systolic pressure, and mean pulmonary artery pressure) in the two weeks between screening and randomization and/or at the end of the study (8-12 weeks)		Excluded from the corresponding analyses