Trial Statistical Analysis Plan

c14526274-02

BI Trial No.:	1245.110
Title:	A phase III randomised, double-blind trial to evaluate efficacy and safety of once daily empagliflozin 10 mg compared to placebo, in patients with chronic Heart Failure with preserved Ejection Fraction (HFpEF). EMPEROR-Preserved
Investigational Product:	Empagliflozin, BI 10773
Responsible trial statistician:	[redacted]
Date of statistical analysis plan:	21 FEB 2017 SIGNED
Version:	Final

1. TABLE OF CONTENTS

LIST OF TABLES

2. LIST OF ABBREVIATIONS

Term	Definition / description
ACE	Angiotensin Converting Enzyme
AE	Adverse event
AESI	Adverse event of special interest
ALT	Alanine transaminase
ANCOVA	Analysis of covariance
ARB	Angiotensin Receptor Blocker
ARNI	Angiotensin Receptor blocker-Neprelysin Inhibitor
ASA	Acetylsalicylic acid
AST	Aspartate transaminase
ATC	Anatomical-Therapeutic-Chemical
BI	Boehringer Ingelheim
BicMQ	Boehringer Ingelheim customized MedDRA Query
BMI	Body mass index
BP	Blood pressure
CCB	Calcium channel blocker
CEC	Clinical Event Committee
CI	Confidence interval
CIF	Cumulative Incidence Function
CKD-EPI	Chronic Kidney Disease Epidemiology Collaboration
CRF	Case report form
CRT	Cardiac resynchronisation therapy
CTP	Clinical Trial Protocol
CTR	Clinical Trial Report
CV	Cardiovascular
DBL	Data base lock
DBP	Diastolic Blood Pressure
DKA	Diabetic ketoacidosis
DM	Diabetes mellitus
DM&SM	Data Management and Statistics Manual
ECG	Electrocardiogram
eCRF	Electronic Case Report Form

Term	Definition / description
eGFR	Estimated glomerular filtration rate
EMA	European Medicines Agency
EoT	End-Of-Text
EOT	End of treatment
EQ5D	EuroQoL 5 dimensions
EudraCT	European Clinical Trials Database
FPG	Fasting plasma glucose
FU	Follow-up
GI	Gastrointestinal
HbA_{1c}	Glycosylated haemoglobin
HCRU	Health Care Resource Utilisation
HDL-C	High density lipoprotein cholesterol
HF	Chronic Heart Failure
HHF	Hospitalisation for heart failure
HLT	High level term
hsTnT	high-sensitivity troponin T
HR	Hazard ratio
ICD	Implantable Cardioverter Defibrillator
ICH	International Conference on Harmonisation
IPV	Important protocol violation
IRT	Interactive Response Technology
ITT	Intention-to-treat
IVRS	Interactive voice response system
JFM	Joint frailty model
KCCQ	Kansas City Cardiomyopathy Questionnaire
LAVD	Left ventricular assist device
LDL-C	Low density lipoprotein cholesterol
LLT	Low level term
LVEF	Left ventricular ejection fraction
LVOT	Last value on treatment
LTFU	Lost to follow-up

Term	Definition / description
Max	Maximum
MedDRA	Medical Dictionary for Regulatory Activities
MI	Myocardial infarction
Min	Minimum
MMRM	Mixed models repeated measures
MRA	Mineralocorticoid Receptor Antagonist
Non-CV	Non-cardiovascular
NT pro-BNP	N-terminal pro b-type natriuretic peptide
NYHA	New York Heart Association Classification
OC-AD	Observed Case including data after discontinuation
OC-OT	Observed Case on-treatment
PK	Pharmacokinetic
PAOD	Peripheral Arterial Occlusive Disease
PT	Preferred term
Q1	Lower quartile
Q3	Upper quartile
PDMAP	Project data management and analysis plan
RS	Randomized set
SAE	Serious adverse event
SAS	Statistical Analysis System
SBP	Systolic blood pressure
SCR	Screened set
SD	Standard deviation
SE	Standard error
SGLT-1	Sodium-glucose co-transporter 1
SGLT-2	SGLT-2 Sodium-glucose co-transporter 2
SI	Système international d'unités
SMQ	Standardized MedDRA query
SOC	System organ class
TBILI	Total bilirubin
TS	Treated set
TS-FU	Treated set – Follow-up

Term	Definition / description
TSAP	Trial statistical analysis plan
UACR	Urine Albumin Creatinine Ratio
ULN	Upper limit of normal
UTI	Urinary tract infection
WHO DD	World Health Organization Drug Dictionary

3. INTRODUCTION

As per the International Conference on Harmonisation (ICH) E9 guidance (1), the purpose of this document is to provide a more technical and detailed elaboration of the principal features of the analysis described in the protocol, and to include detailed procedures for executing the statistical analysis of the primary and secondary variables and other data.

This Trial Statistical Analysis Plan (TSAP) assumes familiarity with the Clinical Trial Protocol (CTP), including Protocol Amendments. In particular, the TSAP is based on the planned analysis specification as written in CTP Section 7 “Statistical Methods and Determination of Sample Size”. Therefore, TSAP readers may consult the CTP for more background information on the study, e.g., on study objectives, study design and population, treatments, definition of measurements and variables, planning of sample size, randomization.

SAS® Version 9.4 or later will be used for all analyses.

4. CHANGES IN THE PLANNED ANALYSIS OF THE STUDY

N/A

5. ENDPOINTS

5.1 PRIMARY ENDPOINT

The primary endpoint is defined in the CTP Section 5.2.1.

For further clarification: Adjudicated CV death always includes death adjudicated as death due to undetermined cause. This is applicable throughout all analyses wherever adjudicated CV death is mentioned.

5.2 SECONDARY ENDPOINTS

5.2.1 Key secondary endpoints

Secondary endpoints are specified in the CTP Sections 5.1.2.

5.2.2 Other Secondary endpoints

Other secondary endpoints are specified in the CTP Sections 5.1.2 and 5.2.

5.3 FURTHER ENDPOINTS

Further endpoints are listed in the CTP Section 5.1.3. The definition and assessment can be found in Section 5.2.

Further endpoints added include:

- Time to non-cardiovascular (non-CV) death
- Fasting plasma glucose (FPG) change from baseline to last value on treatment (LVOT) and follow-up (FU), overall and by status of diabetes mellitus (DM)

5.4 OTHER VARIABLES

5.4.1 Kansas City Cardiomyopathy Questionnaire (KCCQ) scores

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

1. Physical Limitation

Code responses to each of Questions 1a-f as follows:

- Extremely limited = 1
- Quite a bit limited = 2
- Moderately limited = 3
- Slightly limited = 4
- Not at all limited = 5

- Limited for other reasons or did not do = <missing value>

If at least three of Questions 1a-f are not missing, then the Physical Limitation Score is calculated as follows:

Physical Limitation Score = 100*[(mean of Questions 1a-f actually answered) – 1]/4

2. Symptom Stability

Code the response to Question 2 as follows:

- Much worse = 1
- Slightly worse = 2
- Not changed = 3
- Slightly better = 4
- Much better = 5
- I've had no symptoms over the last 2 weeks = 3

If Question 2 is not missing, then the Symptom Stability Score is calculated as follows:

Symptom Stability Score = 100*[(Question 2) – 1]/4

3. Symptom Frequency

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

Question 3

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

Questions 5 and 7

- All of the time = 1
- Several times a day = 2
- At least once a day = 3

- 3 or more times a week but not every day = 4
- 1-2 times a week = 5
- Less than once a week = 6
- Never over the past 2 weeks = 7

Question 9

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

If at least two of Questions 3, 5, 7 and 9 are not missing, then the Symptom Frequency Score is calculated as follows:

S3 = [(Question 3) – 1]/4

S5 = [(Question 5) – 1]/6

S7 = [(Question 7) – 1]/6

S9 = [(Question 9) – 1]/4

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

4. Symptom Burden

Code responses to each of Questions 4, 6 and 8 as follows:

- Extremely bothersome = 1
- Quite a bit bothersome = 2
- Moderately bothersome = 3
- Slightly bothersome = 4
- Not at all bothersome = 5
- I’ve had no swelling/fatigue/shortness of breath = 5

If at least one of Questions 4, 6 and 8 is not missing, then the Symptom Burder Score is defined as follows:

Symptom Burden Score = 100*[(mean of Questions 4, 6 and 8 actually answered) – 1]/4

5. Total Symptom Score

The Total Symptom Score is defined as the mean of the following available summary scores:

Symptom Frequency Score and Symptom Burden Score

6. Self-Efficacy

Code responses to Questions 10 and 11 as follows:

Question 10

- Not at all sure = 1
- Not very sure = 2
- Somewhat sure = 3
- Mostly sure = 4
- Completely sure = 5

Question 11

- Do not understand at all = 1
- Do not understand very well = 2
- Somewhat understand = 3
- Mostly understand = 4
- Completely understand = 5

If at least one of Questions 10 and 11 is not missing, then the Self-Efficacy Score is calculated as:

Self-Efficacy Score = 100*[(mean of Questions 10 and 11 actually answered) – 1]/4

7. Quality of Life

Code responses to Questions 12, 13 and 14 as follows:

Question 12

- It has extremely limited my enjoyment of life = 1

- It has limited my enjoyment of life quite a bit = 2
- It has moderately limited my enjoyment of life = 3
- It has slightly limited my enjoyment of life = 4
- It has not limited my enjoyment of life at all = 5

Question 13

- Not at all satisfied = 1
- Mostly dissatisfied = 2
- Somewhat satisfied = 3
- Mostly satisfied = 4
- Completely satisfied = 5

Question 14

- I felt that way all of the time = 1
- I felt that way most of the time = 2
- I occasionally felt that way = 3
- I rarely felt that way = 4
- I never felt that way = 5

If at least one of Questions 12, 13 and 14 is not missing, then the Quality of Life Score is calculated as:

Quality of Life Score = 100*[(mean of Questions 12, 13 and 14 actually answered) – 1]/4

8. Social Limitation

Code responses to each of Questions 15a-d as follows:

- Severely limited = 1
- Limited quite a bit = 2
- Moderately limited = 3
- Slightly limited = 4
- Did not limit at all = 5

- Does not apply or did not do for other reasons = <missing value>

If at least two of Questions 15a-d are not missing, then the Social Limitation Score is calculated as:

Social Limitation Score = 100*[(mean of Questions 15a-d actually answered) – 1]/4

9. Overall Summary Score

The Overall Summary Score is defined as the mean of the following available summary scores:

Physical Limitation Score, Total Symptom Score, Quality of Life Score and Social Limitation Score

10. Clinical Summary Score

The Clinical Summary Score is defined as the mean of the following available summary scores:

Physical Limitation Score and Total Symptom Score

Note: references to “means of questions actually answered” imply the following.

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

(sum of the responses to those n-i questions) / (n-i)

not

(sum of the responses to those n-i questions) / n

6. GENERAL ANALYSIS DEFINITIONS

6.1 TREATMENTS

There will be 4 basic treatment phases in this trial: screening, study treatment phase (with either empagliflozin 10mg or matching placebo), post-treatment and post-study. However, during the study treatment phase, patients are allowed to go off-treatment and subsequently re-start treatment. This may happen not at all or repeatedly for a given patient.

The purpose of the definitions below is to describe all the different study/treatment intervals, in which a patient can lie during the course of the trial. Note that the term "treatment regimen" also covers the “off-treatment” time periods.

Table 6.1: 1 Treatment regimens / study intervals

Label	**Interval**	**Start date**
Screening	Screening	Date of informed consent
Placebo/ Empagliflozin 10mg	Treatment	Date of first administration of double-blind study treatment
Off-treatment (if applicable)	During Treatment interval, but not on treatment	Date of last administration of the study medication before temporarily discontinuation + 1 day
Placebo/ Empagliflozin 10mg (if applicable)	During Treatment interval, after restart of study medication	Date study medication re-started
Post-treatment	Post-treatment	Date of last administration of study drug + 1 day
Post-study	Post-study	Date of trial completion +1 day

Details on the definition of on-treatment period for different endpoints are listed in Table 6.1:2. The efficacy analyses will follow the intention-to-treat (ITT) principle in assigning patients to treatment groups, i.e. patients will be analysed as randomised.

Safety analyses will also assign patients to the treatment group as randomized.

If a patient erroneously receives the wrong trial drug, the patient will be analysed as per the randomized treatment for all analyses, since patients are to be brought back to the site and dispensed correct drug as soon as possible. Additionally, the AEs with an onset during the time of the incorrect study treatment will be listed separately.

Table 6.1: 2 Endpoint specific follow-up period for the assignment to active treatment

Endpoint	Last day of assignment to treatment phase (days after study drug stop date)
Adverse events	7
Safety laboratory measurements	3
Heart rate	1
Glycosylated haemoglobin (HbA_{1c})	7
FPG	1
Body weight	1
Creatinine and estimated glomerular filtration rate (eGFR)	1
N-terminal pro b-type natriuretic peptide (NT pro-BNP)	1
Blood pressure	1

6.2 IMPORTANT PROTOCOL VIOLATIONS (IPVS)

A protocol violation (PV) is important, if it affects the rights or safety of the study patients or if it can potentially influence the primary outcome measures for the respective patients in a way that is neither negligible nor in accordance with the study objectives.

The IPVs will be described in the clinical trial report (CTR). A listing of patients with medication code broken will be provided.

Table 6.2: 1 Important protocol violations

Category / Code		Description	Requirements	Excluded from
A		**Entrance criteria not met**		
A1		**Target indication not met**		
	A1.06	No chronic HF NYHA class II-IV	Inclusion criterion #3 violated	None
	A1.07	Conditions on ejection fraction (EF) violated	Inclusion criterion #4 violated	None
	A1.08	Conditions on NT-proBNP violated	Inclusion criterion #5 violated	None
	A1.09	Conditions on HF violated	Inclusion criterion #6 violated	None
A2		**Inclusion criteria not met**		
	A2.02	Age out of range	Inclusion criterion #1 violated	None
	A2.08	Specific inclusion criterion for women of child-bearing potential violated	Inclusion criterion #2 violated	None
A3		**Exclusion criteria not met**		
	A3.27	Patient with unstable conditions	Exclusion criteria #1, #2, #7, #11, or #16 violated	None

Table 6.2: 1 Important protocol violations (cont.)

Category / Code		Description	Requirements	Excluded from
	A3.42	Recently implanted ICD	Exclusion criterion #3 violated	None
	A3.43	Implanted CRT	Exclusion criterion #4 violated	None
	A3.29	Cardiomyopathy infiltrative diseases (e.g. amyloidosis), accumulation diseases (e.g. hemochromatosis, Morbus Fabry), muscular dystrophies, with reversible causes (e.g. stress cardiomyopathy), hypertrophic obstructive cardiomyopathy or known pericardial constriction	Exclusion criterion #5 violated	None
	A3.30	Any severe (obstructive or regurgitant) valvular heart disease, obstructive or regurgitant, or any valvular disease expected to lead to surgery during the trial in the opinion of the investigator	Exclusion criterion #6 violated	None
	A3.31	Atrial fibrillation or atrial flutter with a resting heart rate >110bpm	Exclusion criterion #8 violated	None
	A3.32	Systolic blood pressure at visit 1 or 2 out of range	Exclusion criterion #9 or #10 violated	None
	A3.06	Indication of liver disease	Exclusion criterion #12 violated	None
	A3.09	Renal insufficiency or renal impairment (assessed by eGFR)	Exclusion criterion #13 violated	None
	A3.34	Haemoglobin at visit 1 below cut-off	Exclusion criterion #14 violated	None
	A3.35	History of Ketoacidosis	Exclusion criterion #15violated	None
	A3.36	Gastrointestinal (GI) surgery or GI disorder that could interfere with study medication absorption in the investigator's opinion	Exclusion criterion #17 violated	None
	A3.37	Documented or active malignancy	Exclusion criterion #18	None
	A3.38	Life expectancy of <1 years in the opinion of the investigator	Exclusion criterion #19 violated	None
	A3.39	Intake of restricted medications or any drug considered likely to interfere with the safe conduct of the trial	Exclusion criterion #20 violated	None
	A3.40	Treatment with any SGLT-2 inhibitor or SGLT-1 and 2 inhibitor	Exclusion criterion #21 violated	None
	A3.11	Intake of an investigational drug in another trial within 30 days prior to intake of study medication in this trial	Exclusion criterion #22 violated	None
	A3.41	Known allergy or hypersensitivity to empagliflozin	Exclusion criterion #23 violated	None
	A3.13	Relevant alcohol or drug abuse and condition affected study compliance	Exclusion criterion #24 violated	None
	A3.12	Specific exclusion criterion for premenopausal women violated	Exclusion criterion #25 violated	None
	A3.14	Any other clinical condition unsafe for participation	Exclusion criterion #26 violated	None
B		**Informed consent**		
	B1	Informed consent not available/not done	Informed consent date missing or inclusion criterion #9 violated	All
	B2	Informed consent too late	Informed consent date was after Visit 1	None

Table 6.2: 1 Important protocol violations (cont.)

Category / Code		Description	Requirements	Excluded from
C		**Trial medication and randomisation**		
	C1.02	Incorrect trial medication taken	Wrong medication taken for more than 20% of the overall treatment duration Can only be finally judged after data base lock (DBL) since unblinding information is required.	None
D		**Concomitant medication**		
	D2.02	Use of prohibited medication	Use of SGLT-2 or combined SGLT-1 and 2 – inhibitors after randomization before trial termination	None

6.3 PATIENT SETS ANALYSED

The following patient sets are defined

- *Screened Set (SCR)*
  Consists of all patients screened for the trial, with informed consent given and who completed at least one screening procedure at Visit 1.

- *Randomised set (RS)*
  This patient set includes all randomised patients, whether treated or not.

- *Treated set (TS)*
  This patient set includes all patients who were dispensed study medication and were documented to have taken at least one dose of investigational treatment.

- *Treated Set-Follow-up (TS-FU)*
  All patients in the TS, for whom a follow-up visit was performed between 23 and 45 days after last intake of study medication.

6.4 SUBGROUPS

Subgroups to be considered in the analyses are provided below in Table 6.4: 1. Missing categories for subgroup variables will not be considered in the respective analysis.

If there is missing information directly from central laboratory for any of the subgroups of parameters, where data is also collected in the IRT system, then the information as transferred from IRT will be used to assign a patient to a certain category.

Table 6.4: 1 Categories of covariates for displays of baseline characteristics and subgroup analyses

Variable	Categorization	Demo-graphics*	Subgroups for Efficacy endpoints**	Safety
Age (years)	<50 50 to <65 65 to <75 75 to <85 >= 85	X		
	<50 50 to <65 65 to <75 >= 75	X	X§	X
	< 65 ≥65	X	X	
Gender	male female	X	X	X
Region	NA, LA, Europe, Asia, Other+	X	X	X
Ethnicity	Hispanic/ Latino Not Hispanic/ Latino	X	X	X
Race	White Black/ African-American Asian Other including mixed race	X	X	X
BMI (kg/m^2)	<30 ≥30	X	X	
eGFR at baseline	>=90 60 to <90 45 to <60 30 to <45 15 to <30 <15	X		
	>=90 60 to <90 45 to <60 30 to <45 <30	X	X§	X
	>=60 <60	X	X	
UACR	UACR (in mg/g): <30 >=30 to <=300 >300	X	X§	

Table 6.4: 1 Categories of covariates for displays of baseline characteristics and subgroup analyses (cont.)

Variable	Categorization	Demo-graphics*	Subgroups for Efficacy endpoints**	Safety
Diabetes at baseline	DM, no-DM	X	X	X
	DM, pre-DM, non-DM	X	X	
History of hypertension	Yes / No	X	X	
Baseline SBP	< / >= median	X	X	
Baseline BP	SBP<140 and DBP<90 vs. SBP>= 140 or DBP>=90	X	X	X (volume depletion AEs)
Atrial Fibrillation at baseline	Yes/ No	X	X	
Baseline LVEF	</>=50	X	X	
	< / >= median		X	
History of HHF	Yes / No	X	X	
Cause of HF	Ischemic / Non-ischemic	X	X	
Time since diagnosis of HF	<= 1 year, 1-5 years, > 5 years	X	X[§]	
NYHA at baseline	I/II vs. III/IV	X	X	
NT-proBNP at baseline	< / >= median	X	X	
Baseline use of ACE-inibitor, ARB or ARNi	Yes / No	X#	X	X (renal AEs and volume depletion AEs)
Baseline use of ACE-inhibitor or ARB but no ARNi	Yes / No	X#	X	
Baseline use of ARNi	Yes / No	X#	X	
Baseline use of MRA	Yes / No	X#	X	
Baseline use of diuretics	Yes / No	X#	X	X (renal AEs and volume depletion AEs)
Baseline use of loop or high-ceiling diuretics	Yes / No	X#	X	X (renal AEs and volume depletion)
Baseline use of thiazides	Yes / No	X#	X	

Table 6.4: 1 Categories of covariates for displays of baseline characteristics and subgroup analyses (cont.)

Variable	Categorization	Demo-graphics*	Subgroups for Efficacy endpoints**	Safety
Baseline use of diuretics low-ceiling excluding thiazides	Yes / No	X#	X	
Baseline use of beta-blockers	Yes / No	X#	X	
Baseline use of Ivabradine	Yes / No	X#	X	
Baseline use of CCB dihydropyridines	Yes / No	X#	X	
Baseline use of CCB non-dihydropyridines	Yes / No	X#	X	
Baseline use of digitalis	Yes / No	X#	X	
baseline hemoglobin	< / >= median		X	
hsTnT at baseline	< / >= median		X	

* The column demographics shows categories shown in the overall demographics. Demographics are not planned by subgroup.
** Subgroups planned for the primary endpoint, time to cardiovascular (CV) death, time to first HHF, HHF (first and recurrent), renal slope.
[+] Region categorization: see Table 9.1.1.
part of the presentation of baseline concomitant therapy as outlined in section Section 7.2
[§]The interaction tests for these subgroups will be trend tests taking into account ordering of the subgroups.

6.5 POOLING OF CENTRES

This section is not applicable because centre/country is not included in the statistical model.

6.6 HANDLING OF MISSING DATA AND OUTLIERS

6.6.1 Imputation methods

There will be no imputation of data for safety analyses. For patients who discontinue the trial treatment prematurely, all efforts will be made to follow patients for survival and for any other endpoints including the primary and key secondary endpoints until the end of the trial.

There will be different methods of looking at continuous longitudinal data.

Observed case on-treatment (OC-OT):
Only the available data that were observed while patients were on study medication (defined as time from first drug intake until last permanent treatment stop date plus the endpoint specific follow-up time (as defined in Table 6.1:1)) are considered.
Measurements are assigned to planned weeks according to Table 6.7: 1

Observed case including data after discontinuation (OC-AD):
All available data are considered, including values obtained on treatment or post-treatment.
Measurements are assigned to planned weeks according to Table 6.7: 1

KCCQ imputation
For endpoints of KCCQ scores, for patients who die, a worst score (score of 0) will be imputed for the score at all subsequent scheduled visits after the date of death where the score would have been assessed.

The following is used for longitudinal and time to event-data:

Multiple imputations:
A multiple imputations approach will be considered to impute missing data. Multiple imputation approaches taken are further specified in Section 7 with the planned sensitivity analyses.

6.6.2 Missing data

Adverse event data
Missing or partial date information for AEs will be replaced according to general Boehringer Ingelheim (BI) rules described in the BI guidance for handling of missing and incomplete AE dates (2).

Adverse event onset dates, including partial onset dates from clinical event committee (CEC):
In the unlikely case that only the year is documented, the day and month will be imputed as 01 Jan unless the subsequently derived date is before randomisation; in this case the date of randomisation will be used as start date.
If year and month is present the day will be imputed as first of the month unless the subsequently derived date is before randomisation; in this case the date of randomisation will be used.

Death date
For patients with a record of death captured on the electronic case report form (eCRF) with missing or only partial death date from all available sources, the death date will be derived as the latest date of any dates as of: event onset and end dates from either the AE page, or CEC adjudicated onset dates, by using also imputed AE dates, last day known to be alive + 1 day and date of trial completion.

Missing information on the date of first administration of trial drug
If the date of first drug administration is missing but the patient was randomised, the date of the first drug administration will be set to the date of randomisation. If the date of first administration is partially missing with the month and year present, the day will be set to the date of randomisation if randomisation was in the same month. If randomisation was in the month prior to the first drug administration the missing day will be imputed as the first day of the month.

Missing date of last FU for non-fatal events.

If the date of last FU for non-fatal events is missing, it will be imputed as the last of any of the following data:

- visit dates
- date if any assessment such as NYHA, blood pressure, EQ5D, KCCQ, pregnancy test, central laboratory
- adverse event/outcome event start and end dates except fatal events
- onset dates of adjudicated (confirmed and non-confirmed) events except fatal events
- drug administration dates, last drug stop date,
- date of clinical routine exam
- date of trial completion (if patient did not die)

Missing date of trial completion (=last contact or date of death)
If the date is completely missing the following rules will be applied:

- If a patient has withdrawn informed consent, this date will be imputed by the date of IC withdrawal.
- If a patient died and has not withdrawn consent, this date will be imputed by the date of death.
- If the date is incomplete with only month and year reported, the date will be imputed by the first day of the month.
- If a patient did not die, the date of trial completion will be imputed by the last date the patient was known to free of non-fatal events.

All other cases need to be assessed by the trial team on an individual basis, using the above points as guidance.

Missing information on the date of trial medication stop

- If this date is missing and the end-of treatment visit date is available, date of trial medication stop will be imputed by the date of the end-of-treatment visit.
- If the date is incomplete with only month and year and the respective visit date is missing, the date of last drug administration will be imputed by the last day of this month. If this would be later than the date of trial completion, then the date of trial completion will be used for imputation.
- If a patient is lost-to-follow up, no date of last drug administration is reported and the date of the end-of treatment visit is not available, the date of last drug administration is set as the date of last FU for non-fatal events.
- For a patient who dies in the treatment phase with no information on the date of last drug administration, the date is set as the date of death, assuming that the patient took the medication until the date of death.

If the imputed date based on the above rules is later than the longest treatment duration based on drug supply, then trial medication stop will be imputed based on the longest treatment duration based on drug supply.

If after imputation, the date of trial completion is before the date of last drug administration, the start of the post-study period is defined as the maximum of last drug administration +1 day and the date of the trial completion +1 day.

All other cases need to be assessed by the trial team on an individual basis, trying to use the points above as guidance.

Missing information on concomitant therapy dates
For incomplete date information always the midpoint of the possible interval will be used. If only the year is present the day and month will be imputed as 01 July, if year and month is present the day will be imputed as 15. If the year is missing, the date will be considered missing.
If this leads to contradictions for the start or end date of a concomitant therapy (e.g. imputed end date before documented start date) a partial end date will be imputed as the end of the interval or a partial start date will be imputed as the start of the interval in the database to resolve this contradiction.

All other cases or conflicting cases resulting from these imputation rules need to be assessed by the trial team on an individual basis.

Missing measurement to confirm “sustained” decrease
An eGFR $(CDK\text{-}EPI)_{cr}$ reduction is considered sustained, if it is determined by two or more consecutive post-baseline central laboratory measurements separated by at least 30 days (first to last of the consecutive eGFR values). If there is no additional measurement >= 30 days after the eGFR reduction is observed and the patient dies within 60 days of this measurement, then the eGFR reduction is also considered sustained.

Pharmacokinetic (PK) variables
Missing data and outliers of PK data are handled according to (3)

6.7 BASELINE, TIME WINDOWS AND CALCULATED VISITS

For serum creatinine and values based on upon this measurement such as eGFR, baseline will be defined as the mean of all available measurements from the screening visit until start of randomised trial medication. For all other endpoints, baseline will be defined as the last available measurement before start of randomised trial medication.

Since the protocol specifies, that all measurements are taken at visit 2 before any intake of trial medication, all measurements at the first day of drug intake are analysed as before any intake of randomised trial medication.

For randomised patients without any treatment intake: For serum creatinine and values based on upon this measurement such as eGFR, baseline will be defined as the mean of all available measurements from the screening visit until and including the day of randomisation. For all other endpoints, baseline will be defined as the last available measurement before or on the day of randomisation.

Measurements taken after the first intake of randomised trial medication will be considered on-treatment values if they have been obtained up to end of the parameter specific follow-up period as defined in Table 6.7: 1 below and will be assigned to the randomised study drug for efficacy and safety analyses.

Measurements taken after the end of the endpoint specific follow-up period and after the last intake of study drug will be considered post-treatment values.

On-treatment (for OC-OT analysis) or all post-randomisation (for OC-AD analysis) efficacy and safety measurements will be assigned to visits based on time windows around the planned visit dates. These time windows are defined based on the planned number of days after the date of first administration of study drug.

The time window for the first visit after randomisation starts on the day after the first intake of study drug. The midpoint between two on-treatment visits defines the end of a time window, with the midpoint being included in the time window of the preceding visit.

Table 6.7: 1 Time windows for post-baseline efficacy measurements scheduled for each on-treatment visit for the first 3 years

			Time window (actual days after baseline)	
Visit number	Visit label	Planned days	Start	End
Endpoints assessed at each on-site visit (e.g. creatinine / eGFR)				
2	Baseline	0	NA	1
3	Week 4	28	2	56
4	Week 12	84	57	154
6	Week 32	224	155	294
8	Week 52	364	295	448
10	Week 76	532	449	616
12	Week 100	700	617	784
14	Week 124	868	785	952
16	Week 148	1036	953	1120
18	…			
FU	FU	Trt stop + 30 days	Trt stop + 23 days	Trt stop + 45 days
KCCQ				
2	Baseline	0	NA	1
3	Week 4	28	2	126
6	Week 32	224	127	294
8	Week 52	364	295	532
FU	FU	Trt stop + 30 days	Trt stop + 23 days	Trt stop + 45 days

[A] Only values taken prior to the start of treatment with randomised study drug can be considered baseline values. Time windows will be used for assignment of measurements to scheduled visits.

For examinations that are not planned at every on-treatment visit, the time windows will be defined according to the same algorithm, based on the midpoint between the planned visit day of such an examination. Examples for eGFR and KCCQ can be found in Table 6.7: 1

Only one observation per time window will be selected for analysis at an on-treatment visit – the value will be selected which is closest to the protocol planned visit day. If there are two observations which have the same difference in days to the planned day or if there are two observations on the same day, the first value will be used.

Time windows for assignment of planned off-treatment measurement

For evaluation of the off-treatment assessment 'at 30 day follow-up' only values obtained at >= 23 days to <= 45 days after last trial drug stop will be considered.

The value that is closest to the planned day of 30 days after last trial drug stop will be used. If there are 2 values equally close, the later value will be used.

6.8 CALCULATION OF TIME TO EVENT

This section describes the calculation of the time to event and the time that patients without an event are in the study (under risk).

Consistent with that approach, the respective time intervals determine the start and end for the derivation of occurrence of a specific event.

For those patients with an event, the time to event is calculated as:

<date of event> - < start date> + 1

For those patients without an event, the time at risk is calculated as:

<date of censoring> - < start date > + 1

6.8.1 Start date

In general, the time to event will be derived from the date of randomisation.

If study drug administration happened before calling IVRS, the date of first drug administration will be used as start date.

For the following endpoints (analysed as occurrence or time to first event), the date of first drug intake will be used as start date:

- AE analyses acc. to Section 7.8.1
- Endpoints purely based on laboratory measurements, that include a relation to baseline (such as change decrease from baseline >= 40%, doubling vs baseline, etc.)

Please note, that for composite endpoints, that include component(s) using randomisation date and other component(s) using first drug intake date as start date, the time at risk for the composite will start with date of randomisation (which may be earlier). For the individual components, the component specific start date will be used.

6.8.2 Date of event

For adjudicated events, the date determined by the adjudication committee will be used; this can be different from the investigator reported date.

For the endpoints of time to CV death, time to all-cause mortality and time to non CV death the respective death date will be used rather than time to the first onset of the fatal AE.

For composite outcomes, e.g. time to adjudicated HHF or adjudicated CV death, the earliest onset date of the corresponding components will be used. For the component of CV death or other death components, date of death will always be used rather than the onset date also for composite outcomes.

For endpoints, where myocardial infarction (MI) and stroke are included as a fatal and non-fatal component, the onset of the event is considered for the derivation of time to first occurrence, not the date of death. For time to CV death the date of death is used for a fatal MI or fatal stroke.

The time to first occurrence type of endpoints based on laboratory data including endpoints including the requirement a "sustained" measurement are determined by the date of the first measurement that fulfils this condition.

For events with multiple possible episodes, such as HHF or all-cause hospitalisation, the onset date of the first episode will be used unless noted otherwise. The same applies to time-to-AE analysis.

For analysis of recurrent events, time at risk for the next event will start at the day after the end date of an event. For HHF and all-cause hospitalization the end date will be the day of discharge. If date of discharge is missing then it will be assumed that the discharge was the day after hospitalization unless the AE is marked as fatal, in which case date of death will be used.

6.8.3 Censoring

The underlying principle is that the censoring date should be the last date a patient was known to be free of an endpoint event (e.g. free of each component of HHF + CV death).

Patients without occurrence of a specific endpoint (composite endpoint or individual components) will be considered censored at the last date, the patient was known to be free of the event. For non-fatal events this is the last date the patient could be followed up for all non-fatal events as documented in the eCRF or imputed in as per Section 6.6.2..

Censoring is considered independent from study drug intake.

All-cause mortality

A patient, without the event will be censored at the latest of

- Date of study completion
- Last AE onset date or last AE end (if complete-imputed dates not used)
- Last date known alive from the vital status page

Endpoints of any cause-specific death, e.g. CV death

The same censoring rule as in all-cause mortality applies, and in addition, date of death if died from other causes than the one specified in the endpoint.

Endpoints based on laboratory data only

Patients who already fulfil the respective condition at baseline or without post-baseline laboratory measurements are not considered in the number of patients at risk for this endpoint.

If a baseline laboratory measurement is not available for the parameter of interest, it is assumed that the patient did not experience the condition corresponding to the endpoint at baseline and the patient is included in the patients at risk for this endpoint. Patients without an event and available post-baseline laboratory measurements will be considered censored at the date of last laboratory sampling of the corresponding parameter. Patients with missing baseline laboratory required to derive a change from baseline and patients without laboratory data following the baseline measurement will be censored on the date of randomisation.

Composite endpoints

Only patients that are included in the analyses for all components of the composite endpoint will be included in the analysis of the composite.

Of those, a patient with at least one event in any of the components of the composite will be considered to have an event and the date of the first event will be used for the composite endpoint. A patient without an event will be considered censored at the earliest of all censoring dates of the component endpoints.

Censoring for analyses up to trt stop + x days

For any analyses until a certain number of x days after treatment discontinuation (e.g. sensitivity analyses until 30 days after treatment discontinuation), censoring time will be the minimum of the censoring time as described above and treatment discontinuation + x days. Patients with an event after treatment discontinuation + x days will be censored at treatment discontinuation + x days.

7. PLANNED ANALYSIS

For End-Of-Text (EoT) tables, the set of summary statistics is: N / Mean / SD / SE / Min / Q1 (lower quartile)/ Median / Q3 (upper quartile)/ Max. The 1st and 99th percentiles might be substituting minimum and maximum in tables with open-ended values to safeguard against implausible extremes.

Geometric means and ranges will be added to the presentation or replace the presentation of mean and standard deviation for parameters which rather follow a log-normal distribution than a normal distribution.

Statistical parameters will be displayed to a defined number of decimal places as specified in the BI guideline "Reporting of Clinical Trials and Project Summaries" (4)

Tabulations of frequencies for categorical data will include all possible categories and will display the number of observations in a category as well as the percentage (%) relative to the respective treatment group (unless otherwise specified, all patients in the respective patient set whether they have non-missing values or not). Percentages will be rounded to one decimal place. The category missing will be displayed only if there are actually missing values.

Disposition of the patient population participating in the trial will be analyzed by treatment groups and presented in the clinical trial report as a frequency-distribution. The number of patients participating (screened, randomized, screened but not randomized, etc.) in the study by region, country and, for treated patients, centre, will also be analyzed by treatment group and presented as a frequency distribution.

Disposition as required for reporting for the trial in EudraCT will be provided. Enrolment will be summarized by country and by age group for reporting in EudraCT. (see (13)).

Number of patients lost to follow up (no information on vital status after start of study closure) and number of patients lost to follow up for the primary endpoint (no information on primary endpoint after start of study closure) will be summarised.

The frequency of patients with IPVs will be presented by treatment group for the randomized set. The frequency of patients in different analysis sets will also be analyzed for each treatment group.

7.1 DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Demographics and baseline characteristics will be analyzed based on RS. Standard descriptive analysis and summary tables will be presented. These summary tables will include description of subgroup variables detailed in Section 6.4. Descriptive analysis of the following variables measured at baseline will be presented: Age, height, body mass index (BMI), time since diagnosis, HbA_{1c}, systolic blood pressure (SBP), diastolic blood pressure (DBP), weight, eGFR, UACR, NT-pro BNP, low density lipoprotein cholesterol (LDL-C) and high density lipoprotein cholesterol (HDL-C).

A summary of the number of patients in each randomisation stratum per treatment planned vs. actual will also be shown. The planned information will be based upon the data received from

the interactive voice response system (IVRS) provider. Analyses will be based on actual information collected via the CRF / central laboratory, not via IVRS.

7.2 CONCOMITANT DISEASES AND MEDICATION

Only descriptive statistics are planned for this section of the report using the randomised set. Concomitant medication use will be summarised by Anatomical-Therapeutic-Chemical classification level 3 (ATC3) and preferred name. Summaries will be presented for concomitant therapies taken at randomisation and separately those taken at or after randomisation. Separate summaries of use of heart failure-related drugs (e.g. ARNi, beta-blockers, ivabradine, diuretics, ACE-inhibitors, ARBs, MRAs, digitalis), anticoagulants, acetylsalicylic acid (ASA), or lipid lowering drugs at randomisation will be presented. Use of devices at randomisation will also be summarized at randomisation.

Concomitant diseases will be summarised by system organ class and preferred term. Relevant medical history by treatment group will also be presented. Both summaries will be presented using the randomised set.

7.3 TREATMENT COMPLIANCE

Only descriptive statistics are planned for this section of the report. The number and percentage of patients with overall compliance will be reported. Overall compliance will be calculated as a weighted average of reported compliance. The sum of all reported compliance over the planned visits will be divided by the total duration. The treated set of patients will be considered.

7.4 PRIMARY ENDPOINT

7.4.1 Primary analysis

As the primary endpoint, time to the first event of adjudicated HHF or adjudicated CV death will be reported in days. The primary analysis will be based on RS, using all data available until trial completion, including the data after end of treatment.

The primary endpoint will be displayed using cumulative incidence function (CIF) curves and expressed as the hazard ratio with associated two-sided 95% confidence intervals (CIs) and two-sided CIs based available alpha-level for the analysis. The alpha levels for the interim analysis and final analysis will follow the Hwang, Shih and De Cani α-spending function as specified in Section 7.4 of the CTP, which are expected to be 0.001 and 0.0248 (one-sided) when the interim analysis occurs at the time of 60% information.

Estimator and corresponding confidence intervals will not be corrected for interim analysis.

The primary endpoint will be analysed using Cox regression, with factors of treatment (empagliflozin, placebo), region (North America, Latin America, Europe, Asia and “other” including India, South Africa, Turkey and Australia), baseline status of diabetes (diabetes, prediabetes, no diabetes), age (continuous), gender, left ventricular ejection fraction (LVEF) (continuous) and eGFR $(CKD\text{-}EPI)_{cr}$ at baseline (continuous). Since the stratification factors are included in the model as covariates, no stratified Cox regression will be used.

Breslow's method will be used for dealing with ties.

The individual relevant components of the composite will be summarized descriptively. In this descriptive analysis of the relevant component for the composite, CV death after HHF will not be counted. In all other analyses of CV death alone defined in this document, all CV deaths will be counted, disregarding any earlier events.

7.4.2 Sensitivity analyses

The following sensitivity analyses will be conducted:

- A Cox model including only treatment as covariate, not adjusting for any other variables.
- The same Cox regression will be performed on the TS, including only any events up to 30 days after treatment discontinuation.
- For patients who are without primary event and lost to follow up before trial completion, the treatment specific incidence rates for empagliflozin and placebo will be used to impute the primary events in a multiple imputations framework. The primary model will be applied to the imputed datasets. It is planned to perform 100 imputations. Rubin's rules will used to summarize the log hazard ratios and the result will be back-transformed to show a hazard ratio with confidence interval.
- The endpoint will be evaluated based on investigator reported events.
- A competing risk model by Fine-Gray will be explored, including the same set of covariates as in the primary analysis, sub-distribution hazard ratios will be provided (5)

A Kaplan-Meier curve of time to censoring for primary endpoint will be presented in order to assess whether there was differential censoring. For this analysis an event will count as censoring and a censoring (including censoring due to the competing event of non-CV death) will count as event.

7.4.3 Proportional hazards assumption violated

The proportional hazards assumption will be explored by plotting log (-log (survival function)) against the log of time by treatment group and checked for parallelism. The interaction of treatment with log of time will be included in the model described above for an exploratory analysis. Further, Schoenfeld residuals for each covariate and treatment will be plotted against time and log (time).

In case the proportionality assumption is violated for treatment, an attempt will be undertaken to identify groups of patients for which the proportionality assumption holds and a stratified Cox regression will be performed. The HR and corresponding CIs will be obtained from the stratified Cox model.

In addition a piecewise Cox model assuming proportional hazards in a series of consecutive time intervals as proposed by Collett (6) will be investigated.

7.4.4 Subgroup analyses

Subgroup variables will be explored as described in Section 6.4 for the primary endpoint. The HR between the two treatments along with 95% CI and the p-value for test of treatment equality within each category of the subgroup as well as the p-value for the subgroup-by-treatment interaction will be estimated by the Cox proportional hazard model including the same covariates as in the primary analysis of primary endpoint, the subgroup variable if not part of the covariates of the primary analysis model, and subgroup-by-treatment interaction. If the subgroup variable is a categorization of a continuous covariate, this covariate will be dropped from the subgroup model. A forest plot will be presented with the estimated HR and the two-sided 95% CI for each subgroup category. The CIF plots will also be presented for each subgroup category.

For those subgroups marked as § in Table 6.4.1. the interaction p-values will be conducted as trend tests, taking into account that the subgroup categorizations are ordered.

If there are less than 14 patients with events in one subgroup, then this subgroup will not be included in the model. If this leaves only one subgroup, the subgroup analysis will not be conducted.

For the continuous covariates LVEF, eGFR and age, the influence of the covariate will also be investigated on a continuous scale. For this purpose the continuous covariate will be added to the model if not already included and the interaction term of the continuous covariate and treatment will additionally be included into the model. The hazard ratio depending on the continuous covariate will be plotted and the interaction p-value will be reported.

7.5 SECONDARY ENDPOINTS

7.5.1 Key secondary endpoints

Occurrence of adjudicated HHF (first and recurrent)

Hospitalisation for heart failure will be analysed by a joint frailty model (JFM) that accounts for the dependence between recurrent HHF and CV death (7) The primary analysis will be based on all data available until trial completion, including the data after end of treatment.

Define $T_{i0} = 0$ and let $T_{i1}, T_{i2},\ldots, T_{iNi}$ be the recurrent event times for person i, where N_i is the number of recurrent HHF events before $X_i = \min(C_i, D_i)$, the minimum of an independent censoring time C_i and a dependent CV death time D_i. The JFM is defined through the hazard functions for the recurrent event process and CV death

$$r_i(t \mid \omega_i) = \omega_i \exp\{\beta_1 z_i\} r_0(t)$$
$$\lambda_i(t \mid \omega_i) = \omega_i^{\alpha} \exp\{\beta_2 z_i\} \lambda_0(t)$$

The recurrent heart failure hospitalisations hazard function for the i-th patient conditional on the patient specific random frailty, ω_i, is given by r_i and is proportional to the baseline intensity function, r_0. The conditional hazard function for time to CV death for patient i is given by, λ_i, with the baseline hazard given by λ_0, and β_1, β_2 are p x 1 vectors of regression coefficients associated with vectors of covariates z_i. The same covariates as for the analysis of the primary endpoint will be used, e.g. ß.1=treatment (empagliflozin, placebo), ß.2=region (North America, Latin America, Europe, Asia and "other" including India, South Africa, Turkey and Australia), ß.3=baseline status of diabetes (diabetes, prediabetes, no diabetes) etc.

Patient specific independent random effects are denoted by ω_i and are assumed to follow a gamma distribution with mean 1 and variance θ. The correlation of the recurrent events is quantified by θ, with higher values corresponding to greater within-patient correlation and also greater between-patient variability. The parameter α determines the relationship between the recurrent heart failure hospitalisations and time to CV death. When $\alpha < 0$, higher frailty will result in a greater risk of recurrence and lower risk of terminal event (i.e. a negative correlation between the frailties), and when $\alpha > 0$, higher frailty will result in a greater risk of recurrence and is associated with a higher risk of CV death (i.e. a positive correlation between the frailties).

Let t_{ij} and x_i be the observed recurrent event times and follow-up, respectively. Denote by δ_{ij} and Δ_i, the indicator of the recurrent event at time t_{ij} and the indicator of CV death at time x_i, respectively. The likelihood for person i is then given by the following:

$$L_i = \int_{\omega_i} \prod_{j=1}^{N_i} \left[\omega_i r_i(t_{ij})\right]^{\delta_{ij}} \exp\left\{\int_0^{x_i} \omega_i r_i(t)dt\right\} \left[\omega_i^{\alpha} \lambda_i(x_i)\right]^{\Delta_i} \exp\left\{\int_0^{x_i} \omega_i^{\alpha} \lambda_i(t)dt\right\} f_{\theta}(\omega_i)d\omega_i.$$

Adopting piecewise constant hazards for the recurrent events and CV death allows estimation of the likelihood by Gaussian quadrature. The implementation of Gaussian quadrature techniques is incorporated into Proc NLMIXED of SAS 9.4. SAS Code as given in (9) will be used.

The joint model gives two distinct hazard ratios:

$HR_{HHF} = \exp\{\beta_{11}\}$ is the hazard ratio associated with the effect of treatment on the recurrent event rate of HHF, and $HR_{CVD} = \exp\{\beta_{21}\}$ is the CV death hazard ratio.

Estimates and 95%-CI for the hazard ratios and for α (relationship between the recurrent heart failure hospitalisations and CV death) will be given.

The following sensitivity analyses will be conducted

- Based on the TS, including only any events up to 30 days after treatment discontinuation
- Instead of CV-death, jointly model HHF with all-cause mortality as the terminal events

- The endpoint will be additionally evaluated based on investigator reported events for HHF and CV death.

- A parametric joint gamma-frailty model will model the recurrent event component using a Poisson distribution and model the CV mortality component using a log-logistic distribution, conditional on the frailty parameter. Individual frailties are again assumed to follow a Gamma distribution. Thus HHF rates follow a negative binomial distribution and times to CV death follow a Lomax distribution (see (8))

In case the semi-parametric joint modelling cannot converge numerically with the existing SAS procedures, the parametric joint gamma-frailty model as described above may be used instead for the confirmatory analysis.

The number of HHF events per patients will be summarized descriptively. Additionally a negative binomial model will be fitted to the data of recurrent HHF. This will be done once including only treatment as covariate and once including all covariates as the primary model. Rate ratio and confidence intervals of both models will be reported.

The mean cumulative incidence will be displayed for adjudicated recurrent HHF.

Subgroup analyses will be explored as outlined in Section 6.4 for adjudicated recurrent HHF. For subgroup analyses the term of subgroup (if not already part of the model) and subgroup by treatment will be added to the model of the recurrent event. If the subgroup variable is a categorization of a continuous covariate, this covariate will be dropped from the subgroup model.

eGFR (CKD-EPI)$_{cr}$ slope of change from baseline

Slope in change from baseline of eGFR (CKD-EPI)$_{cr}$ will be analysed by a random coefficient model allowing for random intercept and random slope per patient. The model will include the factors treatment, gender, geographical region, and status of DM as fixed effects and eGFR (CKD-EPI)$_{cr}$ at baseline (continuous), LVEF (continuous), age (continuous), time and interaction of treatment by time as linear covariates and allow for randomly varying slope and intercept between patients. The model will include all on-treatment change from baseline data.

Since the slope is run on the change from baseline data, the intercept will model the acute drop, whereas the long-term effect is modelled by the slope.

A plot of individual patient slopes and separately of individual patient intercepts will be provided per treatment.

Multiple imputation will be used to handle missing data as a sensitivity analysis. All variables included in the analysis model will be included in the imputation model. After exploring the missing data mechanism and observed measurements on the blinded data additional variables may also be included in the imputation model. If the data is monotone missing, a regression model will be used for imputation. In the case of non-monotone missing, a Markov chain Monte Carlo (MCMC) step will be used to create monotone missing data in multiple datasets. It is planned to perform 100 imputations. The regression method will then be used to

complete the imputation in each dataset. For each imputed complete dataset, the analysis model described above will be used for the analysis. The slope results will be summarized using Rubin's rules.

Subgroup analyses as outlined in Section 6.4 will be explored for eGFR slope. The subgroup model will include additionally to the model described above, the subgroup if not already part of the model, subgroup by treatment and subgroup by treatment by time. If the subgroup variable is a categorization of a continuous covariate, this covariate will be dropped from the subgroup model.

7.5.2 Other Secondary endpoints

No correction for multiple hypotheses testing will be made for other secondary endpoints.

Time to first event of sustained reduction of ≥40% eGFR (CKD-EPI)$_{cr}$ or

- *(for patients with eGFR (CKD-EPI)$_{cr}$ ≥30 mL/min/1.73 m2 at baseline): sustained eGFR <15 mL/min/1.73m2*
- *(for patients with eGFR (CKD-EPI)$_{cr}$ <30 mL/min/1.73 m2 at baseline): sustained eGFR < 10 mL/min/1.73 m2*

Time to all-cause mortality

Time to first adjudicated HHF

Time to adjudicated CV death

Time to onset of DM in patients with baseline pre-DM

All time-to-event endpoints will be reported in days.

The same model and data frame as used in the primary analysis of primary endpoint will be applied to all these time-to-event endpoints.

For time to first HHF and time to CV death, the analysis will be repeated based on investigator reported events.

Change from baseline in clinical summary score (HF symptoms and physical limitations domains) of the Kansas City cardiomyopathy Questionnaire (KCCQ) at week 52

Change from baseline in clinical summary score for HF symptoms and physical limitations domains of the KCCQ at week 52 will be evaluated by a mixed model for repeated measures data including baseline score, LVEF (continuous), age (continuous) and eGFR (CKD-EPI)$_{cr}$ at baseline (continuous) as linear covariates and treatment, visit, baseline score by visit, visit by treatment, gender, geographical region and status of diabetes at baseline as fixed effects. All on-treatment data up to week 52 will be included and the analysis will be conducted on the treated set.

A sensitivity analysis will be conducted including data after discontinuation (OC-AD) on the randomized set.

Occurrence of all-cause hospitalisation (first and recurrent)

A similar joint frailty model as in the HHF will be analysed for all-cause hospitalization. Instead of CV death, all-cause mortality will be jointly modelled as the terminal events.

7.6 FURTHER ENDPOINTS

Time to event endpoints

Further time to event endpoints will generally be analysed in a Cox proportional hazards model similar to the primary analysis on RS. If the endpoint does not include any cause of death, a CIF plot with all-cause mortality as competing risk will be displayed; otherwise, a CIF plot with causes of death not included in the endpoint as competing risk will be displayed.

Continuous endpoints:

The following endpoints will be evaluated by mixed models repeated measures (MMRM) as defined in the protocol.

- HbA_{1c} change from baseline over time (OC-AD on the randomised set and OC-OT on the treated set)
- SBP, DBP change from baseline over time (OC-AD on the randomised set and OC-OT on the treated set)
- Weight change from baseline over time (OC-AD on the randomised set and OC-OT on the treated set)
- NT-pro BNP relative change from baseline over time (OC-AD on the randomised set and OC-OT on the treated set) (value will be transformed to the log scale before MMRM analysis. Estimates obtained from the model will then be back-transformed and reported on the original scale.)
- KCCQ overall summary score change from baseline at week 52 (OC-AD on the randomised set and OC-OT on the treated set) (model will include all timepoints up to week 52)
- KCCQ individual domains change from baseline at week 52 (OC-AD on the randomised set and OC-OT on the treated set) (model will include all timepoints up to week 52)
- KCCQ total symptom score change from baseline at week 52 (OC-AD on the randomised set and OC-OT on the treated set) (model will include all timepoints up to week 52)

- KCCQ based on patient-relevant outcome change from baseline at week 52 (OC-AD on the randomised set and OC-OT on the treated set) (model will include all timepoints up to week 52)

- UACR relative change from baseline over time (OC-AD on the randomised set and OC-OT on the treated set) (value will be transformed to the log scale before MMRM analysis. Estimates obtained from the model will then be back-transformed and reported on the original scale.)

- eGFR $(CKD\text{-}EPI)_{cr}$ change from baseline over time (OC-AD on the randomised set and OC-OT on the treated set)

As outlined in the protocol, the above endpoints will be analysed in a mixed model with repeated measures (MMRM), including baseline value, age, LVEF and eGFR $(CKD\text{-}EPI)_{cr}$ at baseline as linear covariates and treatment group, visit, visit by treatment interaction, baseline by visit interaction, geographical region, gender and baseline history of DM as fixed effects. An additional factor of "week reachable" for the parameter in question, which for each patient is the theoretically reachable planned measurement based on the time of randomisation will adjust for the different planned study times.

An unstructured covariance structure will be used to model the within-patient errors.

Descriptive statistics will be calculated for the value at visit and change from baseline based on OC-AD on the randomised set and OC-OT on the treated set.

Additionally subgroup analyses will be performed as follows:

HbA_{1c} change from baseline will be evaluated by status of diabetes (non-DM, pre-DM, DM).

UACR will be evaluated by UACR at baseline (<30 mg/g, ≥30mg/g to ≤300mg/g, >300 mg/g).

In order to evaluate the mean effect on eGFR after approximately 3 years, the above described MMRM models will be used in the following way: a mean effect of the timepoints week 124, 148 and 172 will be calculated. This will be done for the imputations of OC and OC-AD.

For MMRM subgroup analyses, the model will additionally include the term for subgroup and the interaction terms for subgroup by visit, subgroup by treatment and subgroup by treatment by visit in addition.

To support analysis of renal function, eGFR throughout the trial will be categorized according to the following CKD staging: All calculations for the staging of renal function will be based on the originally measured laboratory values and the upper limit of normals (ULNs) given by the laboratory, not on normalised values with BI standard reference ranges.

Table 7.6: 1 CKD staging

Stage	**eGFR (mL/min/1.73m²)**	**Description**	**Label for displays**	***Additional labels#***
1	≥90	Normal or high	≥90	≥90 (CKD 1)
2	60 to <90	Mildly decreased	60 to <90	60 to <90 (CKD 2)
3A	45 to <60	Mildly to moderately decreased	45 to <60	45 to <60 (CKD 3a)
3B	30 to <45	Moderately to severely decreased	30 to <45	30 to <45 (CKD 3b)
4	15 to <30	Severely decreased	15 to <30	15 to <30 (CKD 4)
5	<15	Kidney failure	<15	<15 (CKD 5)

A shift table from baseline to last value on treatment for eGFR $(CKD\text{-}EPI)_{cr}$ will be provided.

A summary will also be created representing the number of patients per treatment group who experienced a doubling in creatinine on treatment compared to baseline that was out of the normal range.

In cases where urine albumin values are reported to be below the quantification limits (e.g. <3 mg/L) the albumin / creatinine ratio is determined as missing and will not be replaced by estimated values.

Transitions from baseline to last value on-treatment based on the following UACR categories: normal (<30mg/g), microalbuminuria (30-<=300 mg/g and macroalbuminuria (>300 mg/g) will be presented.

For an evaluation of follow-up after discontinuation of trial medication an additional summary of the NT-proBNP, FPG, eGFR, UACR, creatinine and the KCCQ endpoints as above will be produced. This summary will include descriptive statistics for baseline, actual values and change from baseline to last on-treatment value and follow-up, and change from last on-treatment assessment to follow-up. The TS-FU patient set will the basis for this summary, but only those patients with a valid baseline, last value on treatment and follow-up value will be included in all displays. Analyses will be repeated for patients with at least 52 weeks of cumulative exposure. FPG analysis will also be conducted by DM status at baseline.

Analyses of change from baseline to LVOT and FU as outlined above will be additionally modelled, separately for LVOT and FU. An analysis of covariance (ANCOVA) model including treatment group, gender, geographical region and history of DM as fixed effect and baseline eGFR $(CKD\text{-}EPI)_{cr}$ (continuous), age (continuous), LVEF (continuous) and baseline of the endpoint (if not already included) as linear covariates will be used.

Descriptive statistics will be presented also for creatinine for the value at visit and change from baseline based on OC-AD on the randomised set and OC-OT on the treated set.

Win ratio:

An unmatched win ratio considering adjudicated CV death and adjudicated HHF will be analysed based on unmatched pairs. All patients randomised to empagliflozin will be compared to all patients randomised to placebo. Only common follow-up time will be considered for the comparison. Patients on empagliflozin are considered to have "won" the comparison if either the other patient has died due to CV causes while the patient on empagliflozin was still alive, or if both patients did not die due to CV causes, then if the other patient had more occurrences of HHF. The number of comparisons won is noted as N_W. Patients on empagliflozin are considered to have "lost" the comparison if the empagliflozin patient died due to a CV cause while the patient on placebo was still alive, or if both patients did not die due to a CV cause, then if the patient on empagliflozin had more occurrences of HHF. The number of comparisons lost is noted as N_L. The win ratio is N_W / N_L.

The rules for winning and losing follow Rogers 2014 (8) and analysis of the unmatched win ratio will be conducted as described in Pocock 2012 (10) .

Other types of further endpoints

For other types of endpoints e.g. change in NYHA class at week 52, HCRU and EQ5D, descriptive statistics will be provided.

For NYHA class a shift table will also be provided for changes from baseline over time and for change from baseline to EOT and FU.

Pharmacokinetic analysis

Descriptive statistics of trough concentrations of empagliflozin will be presented.

7.7 EXTENT OF EXPOSURE

There will be three methods of calculating exposure:

a. First intake to last intake of study drug, including off-treatment periods
b. First intake to last intake of study drug, excluding off-treatment periods
c. Overall observational period (randomisation until end of follow-up for vital status, see censoring for all-cause mortality in Section 6.8.3)

Descriptive statistics tables with mean, standard deviation (SD), median and range of the number of days a patient was on treatment will be provided. These tables will also provide the sum-total of the time (in years) that all patients were on treatment.

Frequency tables of number and percentage of patients belonging to categorical ranges of exposure weeks will be provided as well. Following are the categories of exposure-ranges (in weeks): 0 to 12 weeks, >12 to 26 weeks, >26 to 52 weeks, >52 to 78 weeks, >78 to 104

weeks, >104 to 156 weeks, >156 weeks. Categorical ranges may be adapted based on the actual duration of the study.

7.8 SAFETY ANALYSIS

The safety analysis will be based on the treated set (TS), treatment will be evaluated as randomised.

The AE analysis will include all adverse events (including outcome events as reported by the investigator).

While tables will generally display results by randomised treatment, listings will reflect whether a measurement/AE occurred on or off treatment.

7.8.1 Adverse events

Unless otherwise specified, the analyses of adverse events will be descriptive in nature and will be based on the number of patients with AEs and not on the number of AEs.

AEs will be coded using the most recent version of the Medical Dictionary for Regulatory Activities (MedDRA) coding dictionary.

Definitions of Boehringer Ingelheim customized MedDRA Queries (BIcMQs) for new MedDRA versions are maintained by the BI dictionary maintenance group.

For analysis multiple AE occurrence data on the CRF will be collapsed into an AE provided that all of the following applies:

- All AE attributes are identical (including LLT, intensity, action taken, therapy required, seriousness, reason for seriousness, relationship, outcome, AE of special interest and also additional information of specific AEs or AESIs such as source of sepsis (urinary tract or not) or type of genital infection (fungal balanitis or vulvovaginitis versus other than fungal balinitis or vulvovaginitis)).
- The occurrences were time-overlapping or time-adjacent (time-adjacency of 2 occurrences is given if the second occurrence started on the same day or on the day after the end of the first occurrence).

For further details on summarization of AE data, please refer to (2), (11).

7.8.1.1 Assignment of AEs to treatment

The analysis of adverse events will be based on the concept of treatment emergent adverse events. That means that all adverse events occurring between first drug intake till 7 days after last drug intake will be assigned to the randomised treatment. All adverse events occurring before first drug intake will be assigned to ‘pre-treatment’ and all adverse events occurring after last drug intake + 7 days with be assigned to ‘post-treatment’.

In Section 15.3 general AE analyses tables will present only on-treatment AEs (applying the rule of 7 days for assignment as on-treatment) for the treatments groups (exceptions for cancer events, hepatic events, lower limb amputations and bone fractures as well as

adjudicated events see below). When looking at BIcMQs or standardized MedDRA queries (SMQs) and including all AEs up to termination of the trial, the time at risk will match the time at risk of non-fatal CV events (see section 6.8).

Appendix 16.1.9.2 will include an analysis (overall summary table, frequency of AEs by system organ class (SOC) / preferred term (PT), frequency of serious adverse events (SAEs) by SOC/PT) where AEs and SAEs are assigned to the following phases: Screening, each treatment group, post-treatment for each treatment group.
The tables presenting frequency of AEs by SOC/PT and frequency of SAEs by SOC/PT will be repeated in Section 16.1.9.2 with treatment-specific post-treatment phase included, hereby also incidence rates for the post-treatment phase will be presented.

For listings, AEs will be assigned to one of the treatment phases of Screening, Placebo, Empa 10, Placebo post-treat, Empa 10 post-treat, post-study.

7.8.1.2 Analysis of other significant AEs

Other significant AEs will be reported and summarised according to ICH E3 (12) criterion. Thus, AEs classified as 'other significant' will include those non-serious adverse events with 'action taken = study drug permanently / temporarily discontinued'.

7.8.1.3 AE summaries

An overall summary of adverse events will be presented.

The frequency of patients with adverse events will be summarised by treatment, primary system organ class and preferred term. Incidence rates as defined in Section 7.8.1.8 will generally be included. AEs will also be reported by intensity (without incidence rates). Separate tables will be provided for patients with other significant adverse events, for patients with adverse events of special interest (AESIs), for patients with serious adverse events, for patients with AEs leading to discontinuation, and for patients with drug-related AEs.

Overall AE summaries, AEs by SOC and PT, SAEs and AEs leading to discontinuation will additionally be investigated by subgroups as outlined inTable 6.4: 1.

AEs leading to death will be summarized up to end of the trial, also separately for those adjudicated as CV and those adjudicated non-CV cause.

The system organ classes will be sorted according to the standard sort order specified by the European Medicines Agency (EMA), preferred terms will be sorted by frequency (within system organ class).

Additionally, the following analyses will be reported in Appendix 16.1.9.2 for disclosure on EudraCT and clinicaltrials.gov:

- Frequency [N(%)] of subjects with non-serious adverse events occurring with incidence in preferred term greater than 5% by treatment,
- Adverse Events per arm for disclosure on EudraCT by treatment"
- Non-serious Adverse Events for disclosure on EudraCT by treatment

- Serious Adverse Events for disclosure on EudraCT by treatment

For further details, see also (13).

7.8.1.4 Adverse events of special interest (AESIs)

Hepatic injury

Adverse events reported as AEs of special interest relating to hepatic injury as specified in the protocol will be summarised.

Additionally Hepatic AEs will be summarized based on an SMQ based definition. From SMQ Drug related hepatic disorders (20000006) the following narrow sub-SMQs will be used:

- Narrow sub-SMQ Liver related investigations, signs and symptoms (20000008)
- Narrow sub-SMQ Cholestasis and jaundice of hepatic origin (20000009)
- Narrow sub-SMQ Hepatitis, non-infectious (20000010)
- Narrow sub-SMQ Hepatic failure, fibrosis and cirrhosis and other liver damage-related conditions (20000013)

A table with frequencies of patients with these AEs by treatment, primary SOC and preferred term will be provided. This presentation will be repeated by DM status at baseline (DM vs no DM). Hepatic SAEs and hepatic AEs leading to disc based on the above SMQ definition will be presented.

In addition to the '7-day-on-treatment approach', a '30-day-on-treatment approach' will be presented for the overall hepatic adverse events based on the SMQ definition.

Patients with hepatic injury will be listed.

For presentations on adjudicated hepatic events, refer to Section 7.8.1.6.

Decreased renal function

Adverse events reported as AEs of special interest relating to decreased renal function as specified in the protocol will be summarised.

A frequency tables of patients with AEs related to decreased renal function by treatment, primary SOC and preferred term will additionally be provided based on the narrow standardized MedDRA query (SMQ) Acute renal failure (20000003).

This presentation will be repeated by the subgroups as outlined in Table 6.4: 1. SAEs and AEs leading to disc based on the SMQ Acute renal failure (20000003) will be presented.

In addition, frequency tables will be produced for patients with elevated creatinine ≥ 2 x baseline and > 1 x upper limit of normal (ULN).

Patients with decreased renal function will be listed.

Ketoacidosis

A frequency tables of patients with AEs related to ketoacidosis will be presented by treatment, primary SOC and preferred term for investigator reported cases and for the broad and narrow BIcMQ definition of diabetic ketoacidosis.

For the narrow BIcMQ diabetic ketoacidosis (DKA), SAEs and AEs leading to discontinuation will be presented.

For presentations on adjudicated events, refer to Section 7.8.1.6

Patients with DKA based on the narrow and broad BIcMQ or investigator reported ketoacidosis will be listed.

Events leading to lower limb amputation

A frequency table of patients with AEs leading to lower limb amputation as identified by the investigator by treatment, primary SOC and preferred term will be provided.

A separate tables for AEs leading to lower limb amputation which are leading to discontinuation will be presented.

For events leading to lower limb amputations in addition to the '7-day-on-treatment approach' all adverse events that occurred between first study drug intake up to study end will be presented (following censoring rules like non-fatal outcome events). For both approaches, SAEs will be presented.

Lower limb amputations (up to study end) will additionally be summarised by level of amputation, reason for amputation, history of PAOD, previous amputation and status of DM at baseline (DM / no-DM).

Patients with lower limb amputation will be listed.

7.8.1.5 Specific AEs

Hypoglycaemic events

The investigator will record for each AE whether it represents a hypoglycaemic event and, if so, record additional information to assess the intensity of the hypoglycaemic event. On the basis of this information the hypoglycaemic event will be will be categorised as follows:

- asymptomatic hypoglycaemia: event not accompanied by typical symptoms of hypoglycaemia but with a measured plasma glucose concentration ≤ 3.9 mmol/L (≤ 70 mg/dL)
- documented symptomatic hypoglycaemia with a measured plasma glucose concentration ≥ 3.0 mmol/L and ≤ 3.9 mmol/L (≥ 54 mg/dL and ≤ 70 mg/dL): event accompanied by typical symptoms of hypoglycaemia
- documented symptomatic hypoglycaemia with a measured plasma glucose concentration < 3.0 mmol/L (< 54 mg/dL): event accompanied by typical symptoms of hypoglycaemia but no need for external assistance

- severe hypoglycaemic episode: event requiring the assistance of another person to actively administer carbohydrate, glucagon or other resuscitative actions
- symptomatic hypoglycaemia and plasma glucose concentration > 3.9 mmol/L (70 mg/dL)
- symptomatic hypoglycaemia and plasma glucose concentration not measured

Confirmed hypoglycaemic adverse event are defined as hypoglycaemic adverse events that had a plasma glucose concentration ≤ 70 mg/dL or required assistance.

Different tables will be shown for (i) patients with investigator defined asymptomatic or symptomatic hypoglycaemia, and (ii) patients with confirmed hypoglycaemic adverse events, i.e. hypoglycaemic adverse events that had a plasma glucose concentration ≤ 70 mg/dL or required assistance.

Subgroup analyses on confirmed events with respect to age category, renal function and diabetes background (no DM, patients with pre DM, patients with DM) will be performed.

Time to the onset of the first confirmed hypoglycaemia will be displayed using a cumulative incidence function.

In addition the number of patients with hypoglycaemia according to BIcMQ will be presented.

Patients with hypoglycaemic events will be listed.

UTI and genital infections

The following additional specific adverse events will also be assessed and be tabulated by treatment group:

- Genital infections (narrow BIcMQ list and investigator assessment)
- UTI (narrow BIcMQ and investigator assessment)

Genital infections based on investigator assessment will additionally be summarised by type (fungal balanitis or fungal vulvovaginitis, genital infection other than fungal balanitis or fungal vulvovaginits), intensity (mild, moderate or severe), time to onset of first episode (within the first 3 months of treatment or after), how the event was treated (no treatment, therapy assigned, hospitalisation), whether leading to discontinuation of treatment, and the number of episodes per patient.

UTIs based on investigator assessment will additionally be summarised by intensity (mild, moderate or severe), time to onset of first episode (within the first 3 months of treatment or after), anatomical location (upper UTI, lower UTI), how the event was treated (no treatment, therapy assigned, hospitalisation), whether leading to discontinuation of treatment, and the number of episodes per patient.

In the number of episodes analysis of UTI and genital infection AEs will be collapsed within each SSC regardless of preferred term with the collapsing following the description at the start of Section 7.8.1.

For UTIs based on the narrow BIcMQ the subgroups as outlined in Table 6.4: 1 will be presented. The same will be done for genital infections based on the narrow BIcMQ.

Complicated urinary tract infections defined as serious adverse events of narrow BIcMQ UTI, all events of sub-BIcMQ Pyelonephritis, all events of PT Urosepsis will be presented.

Complicated genital infection: defined as all serious events using the narrow BIcMQ Genital infection will also be presented.

UTIs leading to discontinuation based on the narrow BIcMQs will be presented, the same will be repeated for genital infections leading to discontinuation based on the narrow BIcMQ.

Cumulative incidence functions will also be created for time to onset of the first UTI and for time to onset of the first genital infections, both based on the respective narrow BIcMQ.

Patients will UTIs or genital infections will be listed.

Pyelonephritis and sepsis

The following specific adverse event will also be tabulated by treatment group:

- Acute Pyelonephritis (based on investigator assessment): patient incidence overall and by gender
- Pyelonephritis (based on the narrow sub-SMQ): patient incidence overall and by gender
- Sepsis (based on investigator assessment): patient incidence overall and by source of infection (UTI or not UTI)

Patients with pyelonephritis or sepsis will be listed.

Bone fracture events:

Frequency tables of patients with bone fracture by treatment, primary system organ class (SOC) and preferred term will be provided (based on BIcMQ and investigator reporting). Investigator reported fractures will be reported by type of fracture (traumatic vs. non-traumatic).

For bone fractures based on the BIcMQ the subgroups as outlined in Table 6.4: 1 will be presented. Separate tables for bone fractures based on the BIcMQ, which are serious and those which are leading to discontinuation will be presented.

For overall bone fracture based on the BIcMQ in addition to the '7-day-on-treatment approach' all adverse events that occurred between first study drug intake up to study end will be presented (following censoring rules like non-fatal outcome events).

Patients with bone fractures will be listed.

Cancer events:

Cancer will be based on the following Sub-SMQs of SMQ Malignancies (20000090):

- Sub-SMQ Malignant or unspecified tumors (20000091)
- Sub-SMQ Malignancy related conditions (20000092), excluding the PT "Acanthosis nigricans".

Presentation will be done ordered by HLT.

Frequency tables of patients with cancer by treatment, high level term and preferred term will be provided.

For cancer in addition to the '7-day-on-treatment approach' all adverse events that occurred between first study drug intake up to study end will be presented (following censoring rules like non-fatal outcome events).
There will be an additional analysis including all patients who had a minimum cumulative study drug exposure of 6 months (excluding treatment gaps). All AEs starting from date when 6 months cumulative exposure was reached up to individual day of trial completion will be shown (following censoring rules like non-fatal outcome events).

Patients with cancer will be listed.

Volume depletion

Volume depletion will be based on the BIcMQ.

A frequency table of patients with volume depletion by treatment, primary SOC and preferred term will be provided.

For volume depletion events the subgroups as outlined in Table 6.4: 1 will be presented. Separate tables for volume depletion events, which are serious and those which are leading to discontinuation will be presented.

Patients with volume depletion will be listed.

A cumulative incidence function will be used to display the time to first volume depletion event.

For the analysis of laboratory data, refer to Section 7.8.2.

Venous embolic and thrombotic adverse events:

Venous embolic and thrombotic adverse events will be evaluated on the narrow SMQ.

A frequency table of patients with venous embolic and thrombotic adverse events by treatment, primary SOC and preferred term will be provided.

The subgroups as outlined in Table 6.4: 1 will be presented. Separate tables for venous embolic and thrombotic adverse events, which are serious and those which are leading to discontinuation will be presented.

Patients with venous embolic and thrombotic adverse events will be listed.

Increased urination

Frequency of increased urination will be summarised by primary SOC and preferred term based on the project-defined PT list.

7.8.1.6 Events qualifying for external adjudication by the CEC and Hepatic External Adjudication and Adjudication of ketoacidosis

An independent external CEC regularly reviews events and evaluates whether pre-specified criteria for these adjudication endpoints are met. Details on composition of the CEC, responsibilities and clinical event definitions are provided in the separate CEC Charter.

Adjudication assessments will be incorporated into the database.

Details of the adjudication process are described in the CEC charter.

Cardiological/neurological adverse events:
Frequency tables will be provided based on SOC and preferred for events qualifying for adjudication.

The number of patients with confirmed events per event type and breakdown of event subtype will be presented. This will be done for all CEC confirmed events.

A frequency table contrasting local vs. central (adjudicated) assessments will also be generated.

A listing will be provided, that shows the trigger events and result of adjudication.

Hepatic adverse events:
Frequency tables summarizing the relatedness and severity will be provided, including a listing showing the trigger events and adjudication results. This will include all adjudicated events up to the end of the study.

Ketoacidosis:
Frequency tables summarizing the adjudication results will be provided, including a listing showing the trigger events and adjudication results. This will include all adjudicated events up to the end of the study.

7.8.1.7 AEs while patients taking wrong medication

A listing will be provided for AEs that occurred while a patient was taking the wrong medication. Off-drug is not viewed as wrong medication.

7.8.1.8 Adverse event incidence rates

In addition to the frequency tabulations, time-adjusted adverse event analyses will be performed for all AEs, investigator defined drug-related AEs, AEs leading to discontinuation, other significant AEs, serious AEs, and adverse events of special interest by SOC, respectively HLT, and PT.

The time at risk in patient years for on-treatment phase is derived as follows:

Patients with AE:
time at risk (AE) in days = date of start of AE with specified PT/SOC/HLT – study treatment start date + 1

Patients without AE:
time at risk (AE) in days = end date of time at risk – study treatment start date + 1, where end date of time at risk is the minimum of date of last study drug intake + x days and date of death, if applicable.

The standard approach will be x=7 days, but for certain AESIs in addition other approaches will be used.

Total AE-specific time at risk per treatment group is then derived as:
Time at risk (AE) [years] = Sum of time at risk [days] over all patients in a treatment group/ 365.25

For ‘each row of a table’ (e.g. displaying an SOC), time at risk is calculated using start of first AE summarized in this row; e.g. for patient with AE in a specified SOC, time at risk = date of start of AE with specified PT in this SOC – start of study treatment + 1.

The AE incidence rate per 100 patient years can then be calculated as follows:

Incidence rate per 100 patient years (pt-yrs) = 100 * number of patients with AE / time at risk (AE) [years].

In a similar way the time at risk and incidence rate for the post-treatment period is derived. Hereby the start date is the start date of the post-treatment phase instead of the study treatment start date.

7.8.2 Laboratory data

For continuous safety laboratory parameters standardized and normalized values will be derived as well as the differences to baseline. The process of standardization and normalization as well as standard analyses for safety laboratory data are described in the BI guidance for the Display and Analysis of Laboratory Data (14). All analyses considering multiple of times upper limit of normal (ULN) will be based on original and not normalized data.

Laboratory parameters will be shown in conventional units. Where conventional units differ from SI units, analyses will be repeated on SI units in the appendix.

The analyses of laboratory data will be descriptive in nature and will be based on BI standards [see Data Management and Statistics Manual (DM&SM): Display and Analysis of Laboratory Data (14)].

Baseline for safety laboratory parameters will be the last available measurement before the start of randomised trial medication. Laboratory measurements taken up to 3 days after the last administration of randomised study drug will be considered as on-treatment.
Default settings will be used for repeated values (using worst value).

Laboratory values will be compared to their reference ranges and frequency tables will be provided for the number of patients within and outside the reference range at baseline and the last measurement on treatment. Descriptive statistics will be provided by treatment group for baseline, last value on-treatment and for changes from baseline to last value on treatment. Frequency tables will summarise the number of patients with potentially clinically significant abnormalities as defined for the current XLAB macro.

Only patients with at least one available on-treatment value will be included in the analysis of an individual laboratory parameter. All individual data will be presented in listings.

To support analyses of liver related adverse drug effects, patients with Aspartate transaminase (AST) and/or Alanine transaminase (ALT) ≥3xULN with concomitant or subsequent Total Bilirubin (TBILI≥2xULN) in a 30 day period after AST/ALT elevation are of special interest. In addition, of these cases, it will be considered whether the alkaline phosphatase (AP) is less than 2 x ULN (maximum value in a 30 day period after AST/ALT elevation) or not. The start of the 30 day time span is triggered by each liver enzyme elevation above the defined thresholds. Patients who fulfil one or two of the criteria for ALT/AST or total bilirubin elevations above and have no information available for the remaining parameter(s) within the 30 day time window will not be listed under "ALT and/or AST ≥ 3xULN with TBILI ≥ 2xULN".

In addition ALT/AST will be used to investigate elevated liver enzymes:

- ALT/AST ≥ 3 x ULN
- ALT/AST ≥ 5 x ULN
- ALT/AST ≥ 10 x ULN
- ALT/AST ≥ 20 x ULN

Frequency tables of patients with elevated liver enzymes defined by ALT and/or AST, total bilirubin and AP combinations will be provided. A scatter plot of peak ALT against peak total bilirubin will be presented with reference lines for 3 x ULN ALT and 2 x ULN total bilirubin, including an indicator for treatment received. Details on patients with elevated liver enzymes will be listed.

For the following parameters: total cholesterol, HDL-C, LDL-C, trigylcerides, haemoglobin and haematocrit the time course of changes will be assessed. The analysis will be performed by applying MMRM models to OC-AD data (on the randomised set) and respectively OC-OT data (on the treated set). The MMRM models, that will be used, are specified in Section 7.6. These analyses will be conducted on data before any normalization.

The parameters LDL-C/HDL-C ratio and non-HDL cholesterol will be evaluated descriptively over time based on OC-AD data (on the randomised set) and OC-OT data (on the treated set). These analyses will be conducted on data before any normalization.

For an evaluation of follow-up after discontinuation of trial medication an additional summary of haemoglobin, haematocrit, uric acid and lipid parameters will be produced. This summary will include descriptive statistics for baseline, actual values and change from baseline from last on-treatment value and follow-up, and change from last on-treatment assessment to follow-up. The TS-FU patient set will the basis for this summary, but only those patients with a valid baseline, last value on treatment and follow-up value will be included in all displays. Analyses will be repeated for patients with at least 52 weeks of cumulative exposure.

Analyses of change from baseline to LVOT and FU as outlined above will be additionally modelled, separately for LVOT and FU. An ANCOVA model including treatment group, gender, geographical region and history of DM as fixed effect and baseline eGFR $(CKD\text{-}EPI)_{cr}$ (continuous), age (continuous), LVEF (continuous) and baseline of the endpoint as linear covariates will be used.

7.8.3 Vital signs

An MMRM analysis for heart rate over time will be provided based on OC-AD on the randomised set and OC-OT on the treated set imputations. The model will follow the MMRM analysis described in Section 7.6.

7.8.4 Electrocardiogram (ECG)

Clinically relevant abnormalities found at physical examination or ECG at Visit 2 will be considered to have already existed prior to signing of informed consent and therefore should be considered baseline conditions instead of adverse events, unless there is good reason to assume that they first appeared after signing of consent.

Outcomes of ECGs will be part of the reporting of medical history or AE reporting. Categorical findings as collected in the eCRF will also be summarized descriptively..

7.8.5 Others

Frequency of pregnancies and pregnancy outcomes will be listed by treatment.

Results of the Modified Rankin Scale will be summarized descriptively.

8. REFERENCES

1	*CPMP/ICH/363/96:* "Statistical Principles for Clinical Trials", ICH Guideline Topic E9, Note For Guidance on Statistical Principles for Clinical Trials, current version
2	*001-MCG-156_RD-01:* "Handling of missing and incomplete AE dates", current version; IDEA for CON.
3	*001-MCS 36-472*: "Standards and processes for analyses performed within Clinical Pharmacokinetics/Pharmacodynamics", current version; IDEA for CON.
4	*001-MCG-159*: "Reporting of Clinical Trials and Project Summaries", current version; IDEA for CON.
5	R16-3839 Beyersmann J, Allignol A, Schumacher M Competing risks and multistate models with R. New York: Springer (2012)
6	R07-4680 Collett D Modelling survival data in medical research Chapman and Hall/CRC (2003)
7	R16-2229 Rogers JK, Yaroshinsky A, Pocock SJ, Stokar D, Pogoda J Analysis of recurrent events with an associated informative dropout time: application of the joint frailty model. Stat Med 35 (13), 2195 - 2205 (2016)
8	R16-4909 Rogers JK, Pocock SJ, McMurray JJV, Granger CB, Michaelson EL, Ostergren J, et al Analysing recurrent hospitalizations in heart failure: a review of statistical methodology, with application to CHARM-Preserved. Eur J Heart Fail 16, 33 - 40 (2014)
9	Li Lu, Chenwei Liu, Analysis of Correlated Recurrent and Terminal Events Data in SAS®, NESUG 2008
10	R16-4813 Pocock SJ, Ariti CA, Collier TJ, Wang D The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. Eur Heart J 33, 176 - 182 (2012)
11	*001-MCG-156*: "Handling and summarization of adverse event data for clinical trial reports and integrated summaries", current version; IDEA for CON.
12	*CPMP/ICH/137/95*: “Structure and Content of Clinical Study Reports”, ICH Guideline Topic E3; Note For Guidance on Structure and Content of Clinical Study Reports, current version
13	*001-MCG-159_RD-09*: “Additional Analysis Requirements for Disclosing Data from Clinical Trials - Display Templates", current version; IDEA for CON.
14	*001-MCG-157*: "Handling, Display and Analysis of Laboratory Data", current version; IDEA for CON.
15	*001-MCG-741*: "Clinical subgroup analyses for local and regional Populations in Asia - Clinical Bridging Study Waiver (BSW) and Descriptive Subgroup Analysis (SGA) Reports” , current version; IDEA for CON.

9. ADDITIONAL SECTIONS

9.1 REGIONS AND COUNTRIES

Countries will be assigned to regions following the assignment of the IRT system, which is outlined in Table 9.1:1. Listed countries include currently planned backup countries.

Table 9.1: 1 Regions and countries

Region	Country
Asia	China Japan Korea Singapore
Europe	Belgium Bulgaria Czech Republic Denmark Germany Hungary Italy Netherlands Poland Spain Romania Sweden UK
Latin America	Argentina Brazil Colombia Mexico
North America	Canada US
Other	India Australia South Africa Turkey

9.2 CONCOMITANT MEDICATION

Definitions of medication groups (such as ARBs, diuretics) will be based on World Health Organization Drug Dictionary (WHO DD) and will be stored in the PDMAP.

9.3 ADDITIONAL SUB-GROUP ANALYSIS FOR REGIONAL SUBMISSIONS

Disposition and demographics of the subpopulation for patients from USA and subgroup analyses for patient from the USA vs non-USA will be included in the appendix of the CTR. Efficacy endpoints evaluated will be primary endpoint, time to cardiovascular (CV) death, time to first HHF, HHF (first and recurrent) and renal slope. Safety will be summarized for patients from the USA.
Additional country or region-specific analyses will be conducted for patients from Mexico, East-Asia (China, Japan and Korea), China, Japan, Korea and India to be included into the country-specific submission documents as also outlined in (15). Primary endpoint, time to cardiovascular (CV) death, time to first HHF, HHF (first and recurrent) and renal slope will be presented. Main adverse event overviews, disposition, demographics will be presented.

9.4 INTERIM ANALYSES

An interim analysis will be conducted by the DMC as outlined in the CTP. If based on the interim analysis it is decided to stop the trial for overwhelming efficacy, all analyses belonging to the confirmatory testing strategy as described in this TSAP will be conducted on the snapshot as used for the interim analysis done by the DMC. All other analyses will be conducted on the final database lock.

10. HISTORY TABLE

Table 10: 1 History table

Version	Date (DD-MMM-YY)	Author	Sections changed	Brief description of change
Final	**21-FEB-2017**	[redacted]	None	This is the final TSAP without any modification

Trial Statistical Analysis Plan

c14526274-04

BI Trial No.:	1245.110
Title:	A phase III randomised, double-blind trial to evaluate efficacy and safety of once daily empagliflozin 10 mg compared to placebo, in patients with chronic Heart Failure with preserved Ejection Fraction (HFpEF). Including Revised protocol # 03 [c03946327-04]
Investigational Product:	Empagliflozin, BI 10773
Responsible trial statistician:	[redacted]
Date of statistical analysis plan:	12 FEB 2021 REVISED
Version:	Revised

1. TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

2. LIST OF ABBREVIATIONS

Term	Definition / description
ACE	Angiotensin Converting Enzyme
AE	Adverse event
AESI	Adverse event of special interest
ALT	Alanine transaminase
ANCOVA	Analysis of covariance
ARB	Angiotensin Receptor Blocker
ARNI	Angiotensin Receptor blocker-Neprilysin Inhibitor
ASA	Acetylsalicylic acid
AST	Aspartate transaminase
ATC	Anatomical-Therapeutic-Chemical
BI	Boehringer Ingelheim
BicMQ	Boehringer Ingelheim customized MedDRA Query
BMI	Body mass index
BP	Blood pressure
CCB	Calcium channel blocker
CEC	Clinical Event Committee
CI	Confidence interval
CIF	Cumulative Incidence Function
CKD-EPI	Chronic Kidney Disease Epidemiology Collaboration
CRF	Case report form
CRT	Cardiac resynchronisation therapy
CTP	Clinical Trial Protocol
CTR	Clinical Trial Report
CV	Cardiovascular
DBL	Data base lock
DBP	Diastolic Blood Pressure
DKA	Diabetic ketoacidosis
DM	Diabetes mellitus
DM&SM	Data Management and Statistics Manual
ECG	Electrocardiogram
eCRF	Electronic Case Report Form

Term	Definition / description
eGFR	Estimated glomerular filtration rate
EMA	European Medicines Agency
EoT	End-Of-Text
EOT	End of treatment
EQ-5D	EuroQoL 5 dimensions
EudraCT	European Clinical Trials Database
FG	Fasting glucose
FU	Follow-up
GI	Gastrointestinal
HbA_{1c}	Glycosylated haemoglobin
HCRU	Health Care Resource Utilisation
HDL-C	High density lipoprotein cholesterol
HF	Chronic Heart Failure
HHF	Hospitalisation for heart failure
HLT	High level term
hsTnT	high-sensitivity troponin T
HR	Hazard ratio
ICD	Implantable Cardioverter Defibrillator
ICH	International Conference on Harmonisation
IPD	Important protocol deviation
IRT	Interactive Response Technology
ITT	Intention-to-treat
IVRS	Interactive voice response system
JFM	Joint frailty model
KCCQ	Kansas City Cardiomyopathy Questionnaire
LVAD	Left ventricular assist device
LDL-C	Low density lipoprotein cholesterol
LLT	Low level term
LVEF	Left ventricular ejection fraction
LVOT	Last value on treatment
LTFU	Lost to follow-up
Max	Maximum

Term	Definition / description
MedDRA	Medical Dictionary for Regulatory Activities
MI	Myocardial infarction
Min	Minimum
MMRM	Mixed models repeated measures
MRA	Mineralocorticoid Receptor Antagonist
Non-CV	Non-cardiovascular
NT pro-BNP	N-terminal pro B-type natriuretic peptide
NYHA	New York Heart Association Classification
OC-AD	Observed Case including data after discontinuation
OC-OT	Observed Case on-treatment
PK	Pharmacokinetic
PAOD	Peripheral Arterial Occlusive Disease
PT	Preferred term
Q1	Lower quartile
Q3	Upper quartile
PDMAP	Project data management and analysis plan
RS	Randomized set
SAE	Serious adverse event
SAS	Statistical Analysis System
SBP	Systolic blood pressure
SCR	Screened set
SD	Standard deviation
SE	Standard error
SGLT-1	Sodium-glucose co-transporter 1
SGLT-2	Sodium-glucose co-transporter 2
SI	Système international d'unités
SMQ	Standardized MedDRA query
SOC	System organ class
TBILI	Total bilirubin
TS	Treated set
TS-FU	Treated set – Follow-up
TSAP	Trial statistical analysis plan

Term	Definition / description
UACR	Urine Albumin Creatinine Ratio
ULN	Upper limit of normal
UTI	Urinary tract infection
WHO DD	World Health Organization Drug Dictionary

3. INTRODUCTION

As per the International Conference on Harmonisation (ICH) E9 guidance (1), the purpose of this document is to provide a more technical and detailed elaboration of the principal features of the analysis described in the protocol, and to include detailed procedures for executing the statistical analysis of the primary and secondary variables and other data.

This Trial Statistical Analysis Plan (TSAP) assumes familiarity with the Clinical Trial Protocol (CTP), including Protocol Amendments. In particular, the TSAP is based on the planned analysis specification as written in CTP Section 7 “Statistical Methods and Determination of Sample Size”. Therefore, TSAP readers may consult the CTP for more background information on the study, e.g., on study objectives, study design and population, treatments, definition of measurements and variables, planning of sample size, randomization.

SAS® Version 9.4 or later will be used for all analyses.

4. CHANGES IN THE PLANNED ANALYSIS OF THE STUDY

1. In the random coefficient model used to estimate slope in change from baseline of eGFR $(CKD\text{-}EPI)_{cr}$, an additional fixed effect term of interaction of eGFR $(CKD\text{-}EPI)_{cr}$ at baseline (continuous) by time is included.

2. As per protocol the objective of the trial is "to demonstrate superiority of empagliflozin 10 mg versus placebo".
   The ITT analysis approach was chosen to as closely as possible reflect real-life conditions, disregarding any occurrences of treatment stop or restart of treatment, that may happen in clinical practice as well.
   However, it is also recognized and recommended by the Executive Steering Committee that the study-specific treatment discontinuation in the close-out period does not resemble clinical practice.
   Therefore for patients that complete the treatment period according to protocol after close-out announcement the time at risk for efficacy endpoints ends at the discontinuation of treatment and events occurring thereafter are not considered in the efficacy analyses.
   The primary estimand is now specified in Section 7.4.
   An analysis including all events after randomization will be conducted as sensitivity analysis for the primary endpoint and the key secondary endpoint of recurrent HHF. A corresponding approach will be used for continuous endpoints.

3. Mixed model for repeated measures (MMRM) for KCCQ at week 52 as well as other MMRMs were specified in condensed form to be in line with BI internal standards. Therefore, individual model terms were removed from the MMRM if already included as interaction term with treatment or visit. Additionally, the terms included in MMRMs were aligned throughout the TSAP

4. The rules defining 'win' and 'lose' for the win ratio calculation were clarified and adapted to also consider the time to first HHF event in case of a tie in the number of HHF events. The adaption was made since time to first HHF event is also considered relevant information to avoid ties in case the number of HHF events is identical. This approach is a modification of the rules applied by Rogers 2014 (8).

5. ENDPOINTS

5.1 PRIMARY ENDPOINT

The primary endpoint is the composite of time to first event of adjudicated CV death or adjudicated HHF.

For further clarification: Adjudicated CV death always includes death adjudicated as death due to undetermined cause. This is applicable throughout all analyses wherever adjudicated CV death is mentioned.

5.2 SECONDARY ENDPOINTS

5.2.1 Key secondary endpoints

Key secondary endpoints are:

- Occurrence of adjudicated HHF (first and recurrent)
- eGFR $(CKD\text{-}EPI)_{cr}$ slope of change from baseline

5.2.2 Other secondary endpoints

Other secondary endpoints are specified in the CTP Sections 5.1.2 and 5.2, and in TSAP Section 7.5.2.

5.3 FURTHER ENDPOINTS

Further endpoints are listed in the CTP Section 5.1.3, and in TSAP Section 7.6. The definition and assessment can be found in Section 5.2.

Further endpoints added include:

- Time to non-cardiovascular (non-CV) death (i.e. death cases not included in the definition of adjudicated CV death)
- Fasting glucose (FG) change from baseline to last value on treatment (LVOT) and follow-up (FU), overall and by status of diabetes mellitus (DM)
- Time to first investigator-defined CV hospitalization. Investigator-defined CV hospitalization includes AEs ticked by the investigator as “CV hospitalization (due to reasons other than heart failure, MI or stroke)” and AEs ticked as “Hospitalization for Heart failure” as well as events ticked as “Cardiovascular (CV) Death other than CV death attributed to myocardial infarction or stroke”, MI, stroke, non-fatal TIA that are reported as requiring hospitalization.
- Time to atrial fibrillation (defined as time to first reported ECG indicating atrial fibrillation or to first AE with PT atrial fibrillation).

5.4 OTHER VARIABLES

5.4.1 Kansas City Cardiomyopathy Questionnaire (KCCQ) scores

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

1. Physical Limitation

Code responses to each of Questions 1a-f as follows:

- Extremely limited = 1
- Quite a bit limited = 2
- Moderately limited = 3
- Slightly limited = 4
- Not at all limited = 5
- Limited for other reasons or did not do = <missing value>

If at least three of Questions 1a-f are not missing, then the Physical Limitation Score is calculated as follows:

Physical Limitation Score = 100*[(mean of Questions 1a-f actually answered) – 1]/4

2. Symptom Stability

Code the response to Question 2 as follows:

- Much worse = 1
- Slightly worse = 2
- Not changed = 3
- Slightly better = 4
- Much better = 5
- I've had no symptoms over the last 2 weeks = 3

If Question 2 is not missing, then the Symptom Stability Score is calculated as follows:

Symptom Stability Score = 100*[(Question 2) – 1]/4

3. Symptom Frequency

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

Question 3

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

Questions 5 and 7

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

Question 9

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

If at least two of Questions 3, 5, 7 and 9 are not missing, then the Symptom Frequency Score is calculated as follows:

S3 = [(Question 3) – 1]/4

S5 = [(Question 5) – 1]/6

S7 = [(Question 7) – 1]/6

S9 = [(Question 9) – 1]/4

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

4. Symptom Burden

Code responses to each of Questions 4, 6 and 8 as follows:

- Extremely bothersome = 1
- Quite a bit bothersome = 2
- Moderately bothersome = 3
- Slightly bothersome = 4
- Not at all bothersome = 5
- I've had no swelling/fatigue/shortness of breath = 5

If at least one of Questions 4, 6 and 8 is not missing, then the Symptom Burden Score is defined as follows:

Symptom Burden Score = 100*[(mean of Questions 4, 6 and 8 actually answered) – 1]/4

5. Total Symptom Score

The Total Symptom Score is defined as the mean of the following available summary scores:

Symptom Frequency Score and Symptom Burden Score

6. Self-Efficacy

Code responses to Questions 10 and 11 as follows:

Question 10

- Not at all sure = 1
- Not very sure = 2
- Somewhat sure = 3
- Mostly sure = 4
- Completely sure = 5

Question 11

- Do not understand at all = 1
- Do not understand very well = 2

- Somewhat understand = 3
- Mostly understand = 4
- Completely understand = 5

If at least one of Questions 10 and 11 is not missing, then the Self-Efficacy Score is calculated as:

Self-Efficacy Score = 100*[(mean of Questions 10 and 11 actually answered) – 1]/4

7. Quality of Life

Code responses to Questions 12, 13 and 14 as follows:

Question 12

- It has extremely limited my enjoyment of life = 1
- It has limited my enjoyment of life quite a bit = 2
- It has moderately limited my enjoyment of life = 3
- It has slightly limited my enjoyment of life = 4
- It has not limited my enjoyment of life at all = 5

Question 13

- Not at all satisfied = 1
- Mostly dissatisfied = 2
- Somewhat satisfied = 3
- Mostly satisfied = 4
- Completely satisfied = 5

Question 14

- I felt that way all of the time = 1
- I felt that way most of the time = 2
- I occasionally felt that way = 3
- I rarely felt that way = 4
- I never felt that way = 5

If at least one of Questions 12, 13 and 14 is not missing, then the Quality of Life Score is calculated as:

Quality of Life Score = 100*[(mean of Questions 12, 13 and 14 actually answered) – 1]/4

8. Social Limitation

Code responses to each of Questions 15a-d as follows:

- Severely limited = 1
- Limited quite a bit = 2
- Moderately limited = 3
- Slightly limited = 4
- Did not limit at all = 5
- Does not apply or did not do for other reasons = <missing value>

If at least two of Questions 15a-d are not missing, then the Social Limitation Score is calculated as:

Social Limitation Score = 100*[(mean of Questions 15a-d actually answered) – 1]/4

9. Overall Summary Score

The Overall Summary Score is defined as the mean of the following available summary scores:

Physical Limitation Score, Total Symptom Score, Quality of Life Score and Social Limitation Score

10. Clinical Summary Score

The Clinical Summary Score is defined as the mean of the following available summary scores:

Physical Limitation Score and Total Symptom Score

Note: references to “means of questions actually answered” imply the following.

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

(sum of the responses to those n-i questions) / (n-i)

not

(sum of the responses to those n-i questions) / n

Patient-preferred outcome is derived from response to the question about which domain is the most difficult for the patient to cope with at baseline:

- Loss of functional abilities: Physical Limitation Score
- Symptoms due to your heart failure: Total Symptom Score
- Your concern about how to manage your heart failure: Self-Efficacy Score
- Suffering from depression, anxiety, emotional liability because you have heart failure: Quality of Life Score
- Adapting your life to heart failure: Social Limitation Score

5.4.2 LVEF

Values of LVEF will be considered valid, when appropriate methodology was used. When methodology is missing, LVEF will be disregarded.

6. GENERAL ANALYSIS DEFINITIONS

6.1 TREATMENTS

There will be 4 basic treatment phases in this trial: screening, study treatment phase (with either empagliflozin 10mg or matching placebo), post-treatment and post-study. However, during the study treatment phase, patients are allowed to go off-treatment and subsequently re-start treatment. This may happen not at all or repeatedly for a given patient.

The purpose of the definitions below is to describe all the different study/treatment intervals, in which a patient can lie during the course of the trial. Note that the term "treatment regimen" also covers the “off-treatment” time periods.

Table 6.1: 1 Treatment regimens / study intervals

Label	**Interval**	**Start date**
Screening	Screening	Date of informed consent
Placebo/ Empagliflozin 10mg	Treatment	Date of first administration of double-blind study treatment
Off-treatment (if applicable)	During Treatment interval, but not on treatment	Date of last administration of the study medication before temporarily discontinuation + 1 day
Placebo/ Empagliflozin 10mg (if applicable)	During Treatment interval, after restart of study medication	Date study medication re-started
Post-treatment	Post-treatment	Date of last administration of study drug + 1 day
Post-study	Post-study	Date of trial completion +1 day

Details on the definition of on-treatment period for different endpoints are listed in Table 6.1:2. The efficacy analyses will follow the intention-to-treat (ITT) principle in assigning patients to treatment groups, i.e. patients will be analysed as randomised.

Safety analyses will also assign patients to the treatment group as randomized.

If a patient erroneously receives the wrong trial drug, all subsequent medication packs dispensed to the patient will still be for the treatment group to which the patient was randomized. Therefore the adverse events will be analyzed as per randomized treatment, which is expected to reflect the prevailing treatment.

In the exceptional case that a patient took the wrong treatment, adverse events may occur while being on the wrong treatment. Analyses of this data are described in Section 7.8.1.6.

Table 6.1: 2 Endpoint specific assignment to the on-treatment phase

Endpoint	Last day of assignment to the on-treatment phase (days after last intake of study medication)
Adverse events	7
Safety laboratory measurements, including Urine Albumin Creatinine Ratio (UACR)	3
Heart rate	1
Glycosylated haemoglobin (HbA_{1c})	7
FG	1
Body weight	1
Creatinine and estimated glomerular filtration rate (eGFR)	1
N-terminal pro B-type natriuretic peptide (NT pro-BNP)	1
Blood pressure	1
Patient reported outcome	7

6.2 IMPORTANT PROTOCOL DEVIATIONS (IPDS)

A protocol deviation (PD) is important, if it affects the rights or safety of the study patients or if it can potentially influence the primary outcome measures for the respective patients in a way that is neither negligible nor in accordance with the study objectives.

The IPDs will be described in the clinical trial report (CTR). A listing of patients with medication code broken will be provided.

Table 6.2: 1 Important protocol deviations

Category / Code		Description	Requirements	Excluded from
A		**Entrance criteria not met**		
A1		**Target indication not met**		
	A1.06	No chronic HF or no NYHA class II-IV	Inclusion criterion #3 not met	None
	A1.07	Conditions on ejection fraction (EF) not met	Inclusion criterion #4 not met	None
	A1.08	Conditions on NT-proBNP not met	Inclusion criterion #5 not met	None
	A1.09	Conditions on HF not met	Inclusion criterion #6 not met	None
A2		**Inclusion criteria not met**		
	A2.02	Age out of range	Inclusion criterion #1 not met	None

Table 6.2: 1 Important protocol deviations (cont.)

Category / Code		Description	Requirements	Excluded from
	A2.08	Specific inclusion criterion for women of child-bearing potential not met	Inclusion criterion #2 not met	None
A3		**Exclusion criteria met**		
	A3.44	Patient with unstable conditions	Exclusion criteria #1, #2, #7, #11 or #16 met	None
	A3.42	Recently implanted ICD	Exclusion criterion #3 met	None
	A3.43	Implanted CRT	Exclusion criterion #4 met	None
	A3.29	Patients with cardiomyopathies as defined in the protocol	Exclusion criterion #5 met	None
	A3.30	Any severe (obstructive or regurgitant) valvular heart disease, obstructive or regurgitant, or any valvular disease expected to lead to surgery during the trial in the opinion of the investigator	Exclusion criterion #6 met	None
	A3.31	Atrial fibrillation or atrial flutter with a resting heart rate >110bpm	Exclusion criterion #8 met	None
	A3.33	Systolic blood pressure at visit 1 or 2 out of range	Exclusion criterion #9. #10 met	None
	A3.06	Indication of liver disease	Exclusion criterion #12 met	None
	A3.09	Renal insufficiency or renal impairment (assessed by eGFR)	Exclusion criterion #13 met	None
	A3.34	Haemoglobin at visit 1 below cut-off	Exclusion criterion #14 met	None
	A3.35	History of Ketoacidosis	Exclusion criterion #15 met	None
	A3.36	Gastrointestinal (GI) surgery or GI disorder that could interfere with study medication absorption in the investigator's opinion	Exclusion criterion #17 met	None
	A3.37	Documented or active malignancy	Exclusion criterion #18 met	None
	A3.38	Life expectancy of <1 years in the opinion of the investigator	Exclusion criterion #19 met	None
	A3.39	Intake of restricted medications or any drug considered likely to interfere with the safe conduct of the trial	Exclusion criterion #20 met	None
	A3.40	Treatment with any SGLT-2 inhibitor or SGLT-1 and 2 inhibitor	Exclusion criterion #21 met	None
	A3.11	Intake of an investigational drug in another trial within 30 days prior to intake of study medication in this trial	Exclusion criterion #22 met	None
	A3.41	Known allergy or hypersensitivity to empagliflozin	Exclusion criterion #23 met	None
	A3.13	Relevant alcohol or drug abuse and condition affected study compliance	Exclusion criterion #24 met	None
	A3.12	Specific exclusion criterion for premenopausal women not met	Exclusion criterion #25 met	None
	A3.14	Any other clinical condition unsafe for participation	Exclusion criterion #26 met	None
B		**Informed consent**		
	B1	Informed consent not available/not done	Informed consent date missing or inclusion criterion #9 not met	All
	B2	Informed consent too late	Informed consent date was after Visit 1	None

Table 6.2: 1 Important protocol deviations (cont.)

Category / Code		Description	Requirements	Excluded from
C		**Trial medication and randomisation**		
	C1.02	Incorrect trial medication taken	Wrong medication taken for more than 20% of the overall treatment duration Can only be finally judged after data base lock (DBL) since unblinding information is required.	None
D		**Concomitant medication**		
	D2.02	Use of prohibited medication	Use of SGLT-2 or combined SGLT-1 and 2 inhibitors after randomization but before EOT visit occurring at study close out.	None

6.3 SUBJECT SETS ANALYSED

The following patient sets are defined

- *Screened Set (SCR)*
  Consists of all patients screened for the trial, with informed consent given and who completed at least one screening procedure at Visit 1.

- *Randomised set (RS)*
  This patient set includes all randomised patients, whether treated or not.

- *Treated set (TS)*
  This patient set includes all patients who were dispensed study medication and were documented to have taken at least one dose of investigational treatment.

- *Treated Set-Follow-up (TS-FU)*
  All patients in the TS, for whom a follow-up visit was performed (i.e. values of planned assessments- KCCQ, EQ-5D, vital signs or lab data reported) between 23 and 45 days after last intake of study medication.

The TS-FU does not include patients, where no planned measurements were taken, which may happen in case of telephone FU visits. Patients with intake of open label SGLT-2 inhibitor between their EOT and FU visit are also excluded from the TS-FU set.

- *Pharmacokinetic Set (PKS)*
  This patient set includes all evaluable patients of the treated set who provide at least one observation for at least one further PK endpoint of Cpre,ss,N for empagliflozin.

6.4 SUBGROUPS

Subgroups to be considered in the analyses are provided below in Table 6.4: 1. Missing categories for subgroup variables will not be considered in the respective analysis.

If there is missing information for any of the subgroups, where data is also collected in the IRT system, then the information as transferred from IRT will be used to assign a patient to a certain category.

Although several subgroup analyses are prespecified in Table 6.4: 1, the subgroup analysis by Diabetes at baseline (diabetic, non-diabetic) is the medically and academically most important subgroup analysis.

Table 6.4: 1 Categories of covariates for displays of baseline characteristics and subgroup analyses

Variable	Categorization	Demo-graphics[1]	Subgroups for Efficacy endpoints[2]	Safety[3]
Diabetes at baseline	diabetic, non-diabetic	X	X	X (except ketoacidosis and hypoglycaemia)
	non-diabetic (Pre-diabetes, normal (no pre-diabetes)), diabetic (type 2 diabetes, type 1 diabetes)	X		X (Ketoacidosis and hypoglycaemia)
Age (years)	<50 50 to <65 65 to <75 75 to <85 >= 85	X		X
	< 70 ≥70	X	X	
Sex	male female	X	X	X
Region	NA, LA, Europe, Asia, Other[+]	X		*X (overall[4] only)*
Ethnicity	Hispanic/ Latino Not Hispanic/ Latino	X		*X (overall[4] only)*
Race[$]	White Black/ African-American Asian Other including mixed race	X	X	*X*
BMI (kg/m^2)	<30 ≥30	X	X	
eGFR at baseline	>=90 60 to <90 45 to <60 30 to <45 20 to <30 <20	X		
	>=90 60 to <90 45 to <60 30 to <45 <30	X		X
	>=60 <60	X	X	

Table 6.4: 1 Categories of covariates for displays of baseline characteristics and subgroup analyses (cont.)

Variable	Categorization	Demo-graphics[1]	Subgroups for Efficacy endpoints[2]	Safety[3]
UACR	UACR (in mg/g): <30 >=30 to <=300 >300	X		
Baseline SBP	< / >= median		X	
Baseline BP	SBP<140 and DBP<90 vs. SBP>= 140 or DBP>=90	X		X (volume depletion AEs and hypotension AEs)
History of Atrial Fibrillation or Atrial Flutter[$$]	Yes/ No	X	X	
Baseline LVEF	</>=50	X		
	<50, 50 to <60, >=60	X	X	
History of HHF (in the last 12 months)	Yes / No	X	X	
Time since diagnosis of HF	<= 1 year, 1-5 years, > 5 - 10 years, >10 years	X		
NYHA at baseline	I/II/III/IV	X		
	II vs. III/IV		X	
NT-proBNP	< / >= median (calculated separately for patients with/without history of atrial fibrillation or atrial flutter[$$])		X	
Baseline uric acid	Tertiles (calculated separately for males and females)	X	X	
Baseline use of ACE-inhibitor, ARB or ARNi	Yes / No	X#	X	X (renal AEs, volume depletion AEs and hypotension AEs)
Baseline use of MRA	Yes / No	X#	X	

Table 6.4: 1 Categories of covariates for displays of baseline characteristics and subgroup analyses (cont.)

Variable	Categorization	Demo-graphics[1]	Subgroups for Efficacy endpoints[2]	Safety[3]
Baseline use of diuretics	Yes / No	X#		X (renal AEs, volume depletion AEs and hypotension AEs)
Baseline use of loop or high-ceiling diuretics	Yes / No	X#		X (renal AEs, volume depletion AEs and hypotension AEs)

[1] The column demographics shows categories shown in the overall demographics. Demographics are not planned by subgroup.
[2] Subgroups planned for the primary endpoint, time to cardiovascular (CV) death, time to first HHF, HHF (first and recurrent), renal slope.
[3] X means subgroups planned for overall AE summaries, AEs by SOC and PT, SAEs, AEs leading to discontinuation, decreased renal function, UTI, genital infection, volume depletion, hypotension and bone fracture.
[4] Overall means overall AE summaries, AEs by SOC and PT, SAEs and AEs leading to discontinuation
[5] Median calculated separately for baseline atrial fibrillation or atrial flutter (determined by ECG) / no baseline atrial fibrillation and atrial flutter (determined by ECG)
[+] Region categorization: see Table 9.1.1.
[#] part of the presentation of baseline concomitant therapy as outlined in Section 7.2
[$] Due to low number of patients in certain race categories, estimates for race subgroups may be unstable.
[$$] history of atrial fibrillation or atrial flutter means atrial fibrillation or atrial flutter reported in any ECG before treatment intake or history of atrial fibrillation or atrial flutter reported as medical history.

Additional subgroup analyses for specific further endpoints are defined in Section 7.6

6.5 POOLING OF CENTRES

This section is not applicable because centre/country is not included in the statistical model.

6.6 HANDLING OF MISSING DATA AND OUTLIERS

6.6.1 Imputation methods

There will be no imputation of data for safety analyses. For patients who discontinue the trial treatment prematurely, all efforts will be made to follow patients for survival and for any other endpoints including the primary and key secondary endpoints until the end of the trial.

There will be different methods of looking at continuous longitudinal data.

Observed case on-treatment (OC-OT):
Only the available data that were observed while patients were on study medication (defined as time from first drug intake until last permanent treatment stop date plus the endpoint specific follow-up time (as defined in Table 6.1:2) are considered.
Measurements are assigned to planned weeks according to Table 6.7: 1

Observed case including data after treatment discontinuation (OC-AD):
All available data are considered, including values obtained on treatment or post-treatment.
Measurements are assigned to planned weeks according to Table 6.7: 1.

As described in Section 4, for patients that completed the treatment phase as planned (primary reason for end of treatment being completion of treatment period according to protocol), measurements after last study drug intake should not be considered. Therefore for those patients, any data after the endpoint specific follow-up time (as defined in Table 6.1:2) will not be included in the OC-AD analysis.

KCCQ imputation
If not otherwise noted, for endpoints of KCCQ scores, for patients who die, a worst score (score of 0) will be imputed for the score at all subsequent scheduled visits after the date of death where the score would have been assessed.

Multiple imputations:
A multiple imputations approach will be considered to impute missing data. Multiple imputation approaches taken are further specified in Section 7 with the planned sensitivity analyses.

Imputation of missing covariates in multivariate Cox regression models and for recurrent event analyses
To avoid excluding patients from Cox regression analyses due to missing covariates, the overall population median of the corresponding variable will be imputed for continuous covariates and the most frequent category will be imputed for categorical covariates.
No imputation will be done for covariates included in treatment by subgroup interaction terms.

6.6.2 Missing data

Adverse event data
Missing or partial date information for AEs will be replaced according to general Boehringer Ingelheim (BI) rules described in the BI guidance for handling of missing and incomplete AE dates (2).

Partial onset dates from clinical event committee (CEC):
In the unlikely case that only the year is documented, the day and month will be imputed as 01 Jan unless the subsequently derived date is before randomisation; in this case the date of randomisation will be used as start date.
If year and month is present the day will be imputed as first of the month unless the subsequently derived date is before randomisation; in this case the date of randomisation will be used.

Death date
For patients with a record of death captured on the electronic case report form (eCRF) with missing or only partial death date from all available sources, the death date will be derived as the latest date of any dates as of: event onset and end dates from either the AE page, or CEC adjudicated onset dates, by using also imputed AE dates, last day known to be alive + 1 day, range of possible days based on partial death date and date of trial completion.

Missing information on the date of first administration of trial drug
If the date of first drug administration is missing but the patient was randomised, the date of the first drug administration will be set to the date of randomisation. If the date of first administration is partially missing with the month and year present, the day will be set to the date of randomisation if randomisation was in the same month. If randomisation was in the month prior to the first drug administration the missing day will be imputed as the first day of the month.

Missing date of trial completion (=last contact or date of death)
If the date is completely missing the following rules will be applied:

- If a patient has withdrawn informed consent, this date will be imputed by the date of IC withdrawal.
- If a patient died and has not withdrawn consent and is not lost to follow-up, this date will be imputed by the date of death.
- If a patient did not die, the date of trial completion will be imputed by the last date the patient was known to free of non-fatal events. This is defined as the maximum of
    - the last date the patient could be followed up for all non-fatal events as documented in the eCRF
    - last onset or end date of an AE
    - Onset of an adjudicated event
    - End of treatment
    - last visit date (NYHA class, EQ-5D, KCCQ, pregnancy test, vital signs, ECG, or central laboratory reported)

In case of a partially missing date, if the imputed date is before the first day (after the last day) of the month/year given as partial date, the first day (last day) of the month/year will be used.

All other cases need to be assessed by the trial team on an individual patient basis, using the above points as guidance.

Missing information on the date of trial medication stop

If the date is partially or completely missing, use the minimum of the following dates:

- End of treatment visit date if available
- Date of death
- Trial completion (last contact date)
- Longest extrapolated treatment duration (assuming 1 tablet/day)
- (in case for partially missing date) Last day of the year/month given as partial date

In case of a partially missing date, if the imputed date is before the first day of the month/year given as partial date, the first day of the month/year will be used.
All other cases need to be assessed by the trial team on an individual patient basis, using the points above as guidance.

Missing information on concomitant therapy dates
For incomplete date information generally the midpoint of the possible interval will be used. If only the year is present the day and month will be imputed as 01 July, if year and month is present the day will be imputed as 15. If the year is missing, the date will be considered missing.
If this leads to contradictions for the start or end date of a concomitant therapy (e.g. imputed end date before documented start date) a partial end date will be imputed as the end of the interval or a partial start date will be imputed as the start of the interval in the database to resolve this contradiction.
If the medication is reported as being taken at visit 2 (baseline) and the imputed start/end date contradicts this then the start/end date is imputed to the earliest/latest date respectively if this can resolve the contradiction.
Similarly if the medication is reported as not taken at visit 2 (baseline) and the imputed start/end date contradicts this then the start/end date is imputed to the baseline date/ 1 day prior to baseline respectively if this can resolve the contradiction.

All other cases or conflicting cases resulting from these imputation rules need to be assessed by the trial team on an individual patient basis.

Missing measurement to confirm “sustained” decrease
An eGFR $(CDK\text{-}EPI)_{cr}$ reduction is considered sustained, if it is determined by two or more consecutive post-baseline central laboratory measurements separated by at least 30 days (first to last of the consecutive eGFR values). If there is no additional measurement >= 30 days after the eGFR reduction is observed and the patient dies within 60 days of this measurement, then the eGFR reduction is also considered sustained.

Pharmacokinetic (PK) variables
Missing data and outliers of PK data are handled according to (3)

6.7 BASELINE, TIME WINDOWS AND CALCULATED VISITS

For serum creatinine and values based on upon this measurement such as eGFR, baseline will be defined as the mean of all available measurements from the screening visit until start of randomised trial medication. For all other endpoints, baseline will be defined as the last available measurement before start of randomised trial medication.

Since the protocol specifies, that all measurements are taken at visit 2 before any intake of trial medication, all measurements at the first day of drug intake are assumed to qualify as baseline assessments.

For randomised patients without any treatment intake: For serum creatinine and values based on upon this measurement such as eGFR, baseline will be defined as the mean of all available measurements from the screening visit until and including the day of randomisation. For all other endpoints, baseline will be defined as the last available measurement before or on the day of randomisation.

Measurements taken after the first intake of randomised trial medication will be considered on-treatment values if they have been obtained up to end of the parameter specific follow-up period as defined in Table 6.7: 1 below and will be assigned to the randomised study drug for efficacy and safety analyses.

Measurements taken after the end of the endpoint specific follow-up period after the last intake of study drug will be considered post-treatment values.

On-treatment (for OC-OT analysis) or all post-randomisation (for OC-AD analysis) efficacy and safety measurements will be assigned to visits based on time windows around the planned visit dates. These time windows are defined based on the planned number of days after the date of first administration of study drug.

The time window for the first visit after randomisation starts on the day after the first intake of study drug. The midpoint between two on-treatment visits defines the end of a time window, with the midpoint being included in the time window of the preceding visit.

Table 6.7: 1 Time windows for post-baseline efficacy measurements scheduled for each on-treatment visit for the first 3 years

			Time window (actual days after baseline)	
Visit number	Visit label	Planned days	Start	End
Endpoints assessed at each on-site visit (e.g. creatinine / eGFR)				
2	Baseline[A]	1	NA	1
3	Week 4	29	2	57
4	Week 12	85	58	155
6	Week 32	225	156	295
8	Week 52	365	296	449
10	Week 76	533	450	617
12	Week 100	701	618	785
14	Week 124	869	786	953
16	Week 148	1037	954	1121
18	…			
FU	FU	Trt stop + 30 days	Trt stop + 23 days	Trt stop + 45 days
KCCQ				
2	Baseline[A]	1	NA	1
4	Week 12	85	2	154
6	Week 32	225	155	295
8	Week 52	365	296	435
EOT	EOT / Post week 52	NA	436	Trt stop
FU	FU	Trt stop + 30 days	Trt stop + 23 days	Trt stop + 45 days

[A] Only values taken prior to the start of treatment with randomised study drug can be considered baseline values. Time windows will be used for assignment of measurements to scheduled visits.

For examinations that are not planned at every on-treatment visit, the time windows will be defined according to the same algorithm, based on the midpoint between the planned visit day of such an examination. Examples for eGFR and KCCQ can be found in Table 6.7: 1

Only one observation per time window will be selected for analysis at an on-treatment visit – the non-missing value will be selected which is closest to the protocol planned visit day. If there are two observations which have the same difference in days to the planned day or if there are two observations on the same day, the first value will be used.

Baseline definition for concomitant therapies

Concomitant medication taken at baseline is any medication with start date 'continued' or before date of first study medication intake (randomisation date will be used for patients not treated) and end date continued on or after date of first study medication intake (randomisation date will be used for patients not treated).

Time windows for assignment of planned off-treatment measurement

For evaluation of the off-treatment assessment 'at 30 day follow-up' only values obtained at >= 23 days to <= 45 days after last trial drug stop will be considered.

The value that is closest to the planned day of 30 days after last trial drug stop will be used. If there are 2 values equally close, the later value will be used.

6.8 CALCULATION OF TIME TO EVENT

This section describes the calculation of the time to event and the time that patients without an event are in the study (under risk).

Consistent with that approach, the respective time intervals determine the start and end for the derivation of occurrence of a specific event.

For those patients with an event, the time to event is calculated as:

<date of event> - < start date> + 1

For those patients without an event, the time at risk is calculated as:

<date of censoring> - < start date > + 1

6.8.1 Start date

In general, the time to event will be derived from the date of randomisation.

If study drug administration happened before calling IVRS, the date of first drug administration will be used as start date.

For the following endpoints (analysed as occurrence or time to first event), the date of first drug intake will be used as start date:

- AE analyses according to Section 7.8.1
- Endpoints purely based on laboratory measurements, that include a relation to baseline (such as change decrease from baseline >= 40%, doubling vs baseline, etc.)

Please note, that for composite endpoints, that include component(s) using randomisation date and other component(s) using first drug intake date as start date, the time at risk for the composite will start with date of randomisation (which may be earlier). For the individual components, the component specific start date will be used.

6.8.2 Date of event

For adjudicated events, the date determined by the adjudication committee will be used; this can be different from the investigator reported date.

For the endpoints of time to CV death, time to all-cause mortality and time to non CV death the respective death date will be used rather than time to the first onset of the fatal AE.

For composite outcomes, e.g. time to adjudicated HHF or adjudicated CV death, the earliest onset date of the corresponding components will be used. For the component of CV death or other death components, date of death will always be used rather than the onset date also for composite outcomes.

For endpoints, where myocardial infarction (MI) and stroke are included as a fatal and non-fatal component, the onset of the event is considered for the derivation of time to first occurrence, not the date of death. For time to CV death the date of death is used for a fatal MI or fatal stroke.

The time to first occurrence type of endpoints based on laboratory data including endpoints including the requirement a “sustained” measurement are determined by the date of the first measurement that fulfils this condition.

For events with multiple possible episodes, such as HHF or all-cause hospitalisation, the onset date of the first episode will be used unless noted otherwise. The same applies to time-to-AE analysis.

For efficacy endpoint analyses (endpoints described in sections 7.4-7.6):

As described in Section 4, the follow-up period will not be included in the time at risk for efficacy endpoints. Therefore events later than the date of treatment discontinuation (if reason for discontinuation was completion) will not be counted.

For the primary endpoint and the key secondary endpoint of recurrent HHF, additionally there will be sensitivity analysis including all events up to individual trial completion including the FU period.

6.8.3 Censoring

The underlying principle is that the censoring date should be the last date a patient was known to be free of an endpoint event (e.g. free of each component of HHF + CV death).

For all endpoints except all-cause mortality and cause-specific death, patients without occurrence of a specific endpoint (composite endpoint or individual components) will be considered censored at the individual day of trial completion.

The individual day of trial completion will be the latest of:

- the last date the patient could be followed up for all non-fatal events as documented in the eCRF
- last onset of an AE or date of death
- onset dates of adjudicated events
- end of treatment
- last visit date (NYHA class, EQ-5D, KCCQ, pregnancy test, vital signs, ECG, or central laboratory- reported)

Censoring is considered independent from study drug intake.

All-cause mortality

A patient, without the event will be censored at the latest of

- Individual day of trial completion (without the restrictions defined above for patients with withdrawn consent or lost to follow-up)
- Last date known alive from the vital status page

Endpoints of any cause-specific death, e.g. CV death

The same censoring rule as in all-cause mortality applies, and in addition, date of death if died from other causes than the one specified in the endpoint.

Endpoints based on laboratory data only

Patients who already fulfil the respective condition at baseline are not considered in the number of patients at risk for this endpoint.

If a baseline laboratory measurement is not available for the parameter of interest, it is assumed that the patient did not experience the condition corresponding to the endpoint at baseline and the patient is included in the patients at risk for this endpoint. Patients without an event and available post-baseline laboratory measurements will be considered censored at the date of last laboratory sampling of the corresponding parameter. Patients with missing baseline laboratory required to derive a change from baseline and patients without laboratory data following the baseline measurement will be censored on the date of randomisation.

Composite endpoints

Only patients that are included in the analyses for all components of the composite endpoint will be included in the analysis of the composite.

Of those, a patient with at least one event in any of the components of the composite will be considered to have an event and the date of the first event will be used for the composite

endpoint. A patient without an event will be considered censored at the earliest of all censoring dates of the component endpoints.

For efficacy endpoint analyses (endpoints described in Section 7.4-7.6):

As described in Section 4, for patients that completed the treatment phase as planned (primary reason for end of treatment being completion of treatment period according to protocol) the follow-up period will not be included in the time at risk for efficacy endpoints. Therefore for those patients, the minimum of the treatment discontinuation date and the above described dates will be used for censoring, except if specifically noted otherwise (for the sensitivity analysis). For patients that did not complete the treatment phase, the dates defined above will be used.

For the primary endpoint and the key secondary endpoint of recurrent HHF, additionally there will be a sensitivity analysis considering all events up to individual trial completion including the FU period.

Censoring for analyses up to trt stop + x days

For any analyses until a certain number of x days after treatment discontinuation (e.g. sensitivity analyses until 30 days after treatment discontinuation), censoring time will be the minimum of the censoring time as described above and treatment discontinuation + x days. Patients with an event after treatment discontinuation + x days will be censored at treatment discontinuation + x days.

7. PLANNED ANALYSIS

For End-Of-Text (EoT) tables, the set of summary statistics is: N / Mean / SD / SE / Min / Q1 (lower quartile)/ Median / Q3 (upper quartile)/ Max. The 1st and 99th percentiles might be substituting minimum and maximum in tables with open-ended values to safeguard against implausible extremes.

Geometric means and ranges will be added to the presentation or replace the presentation of mean and standard deviation for parameters which rather follow a log-normal distribution than a normal distribution.

Statistical parameters will be displayed to a defined number of decimal places as specified in the BI guideline "Reporting of Clinical Trials and Project Summaries" (4)

Tabulations of frequencies for categorical data will include all possible categories and will display the number of observations in a category as well as the percentage (%) relative to the respective treatment group (unless otherwise specified, all patients in the respective patient set whether they have non-missing values or not). Percentages will be rounded to one decimal place. The category missing will be displayed only if there are actually missing values.

Disposition of the patient population participating in the trial will be analysed by treatment groups and presented in the clinical trial report as a frequency-distribution.

Disposition as required for reporting for the trial in EudraCT will be provided. Enrolment will be summarised by country and by age group for reporting in EudraCT. (see (13)).

Number of patients lost to follow up (no information on vital status after start of study closure) and number of patients lost to follow up for the primary endpoint (no information on primary endpoint after start of study closure) will be summarised.

The reason for not randomising screened patients will be summarized descriptively.

The frequency of patients with IPDs will be presented by treatment group for the randomised set. The frequency of patients in different analysis sets will also be analyzed for each treatment group.

Descriptive statistics on impact of COVID-19 on study visits as well as study medication intake will be provided.

7.1 DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Demographics and baseline characteristics will be analyzed based on RS. Demographics will be repeated on the TS if the analysis sets differ by more than 1%. Standard descriptive analysis and summary tables will be presented. These summary tables will include description of subgroup variables detailed in Section 6.4. Descriptive analysis of the following variables measured at baseline will be presented: Age, body mass index (BMI), time since diagnosis, systolic blood pressure (SBP), diastolic blood pressure (DBP), heart rate, weight, eGFR, UACR, NT-pro BNP, LVEF, hemoglobin and troponin T. HbA_{1c} will be presented for patients with diabetes at baseline. NTproBNP will be shown for patients with or without atrial fibrillation/ atrial flutter at baseline.

A summary of the number of patients in each randomisation stratum per treatment will also be shown. The information will be based upon the data received from the interactive voice response system (IVRS) provider. Analyses will be based on actual information collected via the CRF / central laboratory, not via IVRS.

7.2 CONCOMITANT DISEASES AND MEDICATION

Only descriptive statistics are planned for this section of the report using the randomised set. Concomitant medication use will be summarised by Anatomical-Therapeutic-Chemical classification level 3 (ATC3) and preferred name. Summaries will be presented for concomitant therapies taken at baseline and separately those taken after baseline. Separate summaries of use of heart failure-related drugs (e.g. ARNi, beta-blockers, ivabradine, diuretics, ACE-inhibitors, ARBs, MRAs, cardiac glycosides), anticoagulants, acetylsalicylic acid (ASA), or lipid lowering drugs at baseline and newly introduced after baseline will be presented. Use of devices and other non-medication therapy at baseline and newly introduced after baseline will also be summarized. Changes of diuretic therapy over time will be summarized.

Concomitant diseases will be summarised by system organ class and preferred term. Relevant medical history by treatment group will also be presented. Both summaries will be presented using the randomised set.

7.3 TREATMENT COMPLIANCE

Only descriptive statistics are planned for this section of the report. The number and percentage of patients with overall compliance will be reported. Overall compliance will be calculated as a weighted average of reported compliance. The sum of all reported compliance over the planned visits will be divided by the total duration. The treated set of patients will be considered.

7.4 PRIMARY ENDPOINT

The primary estimand of this trial is the hazard ratio of the time to first event of adjudicated HHF or adjudicated CV death between empagliflozin 10 mg and placebo in the population of patients with heart failure with preserved ejection fraction. The primary comparison will be made regardless of changes of treatment (including discontinuation of trial medication) until completion of the planned treatment phase. For clarification, this excludes events and time at risk after the protocol-specified treatment discontinuation for patients that complete the treatment period according to protocol.

7.4.1 Primary analysis of the primary endpoint(s)

As the primary endpoint, time to the first event of adjudicated HHF or adjudicated CV death will be reported in days. The primary analysis will be based on RS, using all data available until completion of the planned treatment phase, including the data after end of treatment for patients not completing the treatment phase as planned.

The primary endpoint will be displayed using cumulative incidence function (CIF) curves and expressed as the hazard ratio with associated two-sided 95% confidence intervals (CIs) and two-sided CIs based available alpha-level for the analysis.

For the interim analysis and final analysis, the following Hwang, Shih and De Cani α-spending function at information fraction t_k with parameter $\gamma = -8$ will be used:

$$\alpha^*(\gamma, t_k) = \min\left\{\alpha, \quad \alpha\frac{1-e^{-\gamma t_k}}{1-e^{-\gamma}}\right\} = \min\left\{0.025, \quad 0.025\frac{1-e^{8t_k}}{1-e^{8}}\right\}$$

The alpha levels are expected to be 0.001 and 0.0248 (one-sided) for the interim and final tests respectively, when the interim analysis occurs at the time of 60% information.

Estimator and corresponding confidence intervals will not be corrected for interim analysis.

The primary endpoint will be analysed using Cox regression, with factors of treatment (empagliflozin, placebo), region (North America, Latin America, Europe, Asia and "other" including India, South Africa and Australia), baseline status of diabetes (diabetes, prediabetes, no diabetes), age (continuous), sex, left ventricular ejection fraction (LVEF) (continuous) and eGFR $(CKD\text{-}EPI)_{cr}$ at baseline (continuous). Since the stratification factors are included in the model as covariates, no stratified Cox regression will be used.

Breslow's method will be used for dealing with ties.

The individual relevant components of the composite will be summarized descriptively. In this descriptive analysis of the relevant component for the composite, CV death after HHF will not be counted. In all other analyses of CV death alone defined in this document, all CV deaths will be counted, disregarding any earlier events.

A hierarchical testing procedure will be followed for the assessment of the primary and key secondary endpoints. For all endpoints, superiority of empagliflozin vs. placebo will be evaluated using a two-sided test.

The tests will be performed in the following hierarchical order:

1. Time to first event of adjudicated CV death or adjudicated HHF
2. Occurrence of adjudicated HHF (first and recurrent)
3. eGFR $(CKD\text{-}EPI)_{cr}$ slope of change from baseline

Starting from step 1, if the null hypothesis of no difference is rejected, and the result is more favourable for empagliflozin, superiority is concluded in the tested endpoint, and the overall type I error is preserved for the test in the next step. If at any step the null hypothesis of no difference is not rejected, subsequent tests are conducted in an exploratory manner.

The overall type I error rate will be preserved at a level of 0.05 (2-sided). The type I error rate used at the final analysis will be influenced by the pre-planned interim analysis.

In the final analysis after the evaluation of recurrent HHF, alpha will be split into 0.001 to be used for the analysis of eGFR slope, and the rest will be transferred to the meta-analyses.

In case the trial is finished early at the time of interim analysis, using $\alpha_{interim}$ for the primary and key secondary endpoints in the testing hierarchy according to the α-spending function above, the following α-split will be used for the eGFR slope analysis and the meta-analysis:

0.1 * $\alpha_{interim}$ will be used for the eGFR slope analysis and

0.9 * $\alpha_{interim}$ will be transferred to the meta-analyses

In both the interim and final analyses, if the slope analysis is successful, the alpha of this branch will then be transferred to the meta-analyses.

The testing hierarchy is summarized in Figure 7.4.1: 1.

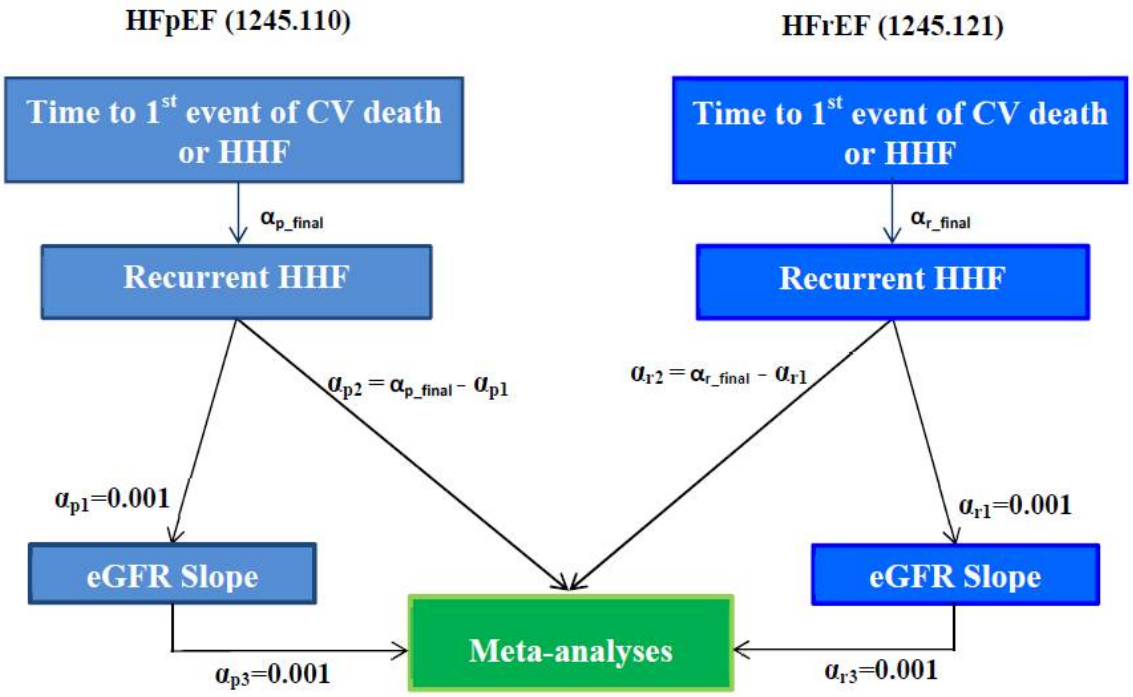

Figure 7.4.1: 1 Hierarchical analysis of trial in HFpEF (1245.110) and the parallel trial in HFrEF (1245.121) showing the alpha-spending at the final analysis

7.4.2 Sensitivity analyses

The following sensitivity analyses will be conducted:

- A Cox model including only treatment as covariate, not adjusting for any other variables.

- The same Cox regression as the primary analysis will be performed on the TS, with observation period up to 30 days after treatment discontinuation. Total number of patients with primary events occurring during the 30 days after treatment discontinuation will also be tabulated.

- For patients who are without primary event and lost to follow up before trial completion, the treatment specific incidence rates for empagliflozin and placebo for retrieved drop-outs will be used to impute the primary events in a multiple imputations framework. The primary model will be applied to the imputed datasets. It is planned to perform 100 imputations. Rubin's rules will be used to summarize the log hazard ratios and the result will be back-transformed to show a hazard ratio with confidence interval.

- The endpoint will be evaluated based on investigator defined events (same model as the primary primary analysis). Investigator defined HHF includes AEs ticked by the investigator as "Hospitalisation due to heart failure". Investigator defined CV death includes AEs ticked by the investigator as "CV death other than CV death attributed to myocardial infarction or stroke", "Undetermined cause of death" as well as HHF, CV hospitalization, MI or stroke if the event was fatal.

- A competing risk model by Fine-Gray will be explored, including the same set of covariates as in the primary analysis, sub-distribution hazard ratios will be provided (5).

- The primary analysis will be repeated to include all events up to individual trial completion including the follow-up period.

- The primary analysis will be repeated to include only confirmed primary events which meet the CDISC guidance criteria for hospitalization for heart failure (i.e., confirmed by adjudication committee, but excluding those missing a physical sign or laboratory test, or both)

- A Cox model including additional prognostic covariates, Log(NTProBNP), baseline atrial fibrillation/atrial flutter, and HHF in last 12 months, in addition to the covariates in the primary model.

- COVID-19 sensitivity analyses will be conducted:

    - The primary analysis will be repeated to include all events up to cutoff dates prior to COVID-19 outbreak – cutoff (censoring) dates are 30Nov2019 for China and 31Dec2019 for all other countries.

    - The same Cox regression model as for the primary endpoint will be fitted, however, adding an additional time dependent covariate for COVID-19 outbreak (time dependent covariate being 0 before outbreak date defined above and 1 thereafter)

- The same Cox regression model as for the primary endpoint will be fitted, however, adding an additional time dependent covariate for COVID-19 outbreak as well as a treatment by COVID-19 outbreak interaction term

- The primary analysis will be repeated to include all events up to 7 days prior a reported SARS-CoV-2 infection (based on broad scope SARS-CoV-2 infections)

A Kaplan-Meier curve of time to censoring for primary endpoint will be presented in order to assess whether there was differential censoring. For this analysis, a primary endpoint event will be counted as censoring and a censoring (including censoring due to the competing event of non-CV death) will be counted as an event.

7.4.3 Proportional hazards assumption violated

The proportional hazards assumption will be explored by plotting log (-log (survival function)) against the log of time by treatment group and checked for parallelism. The interaction of treatment with log of time will be included in the model described above for an exploratory analysis. Further, Schoenfeld residuals for each covariate and treatment will be plotted against time and log (time).

In case the proportionality assumption is violated for treatment, an attempt will be undertaken to identify groups of patients for which the proportionality assumption holds and a stratified Cox regression will be performed. The HR and corresponding CIs will be obtained from the stratified Cox model.

In addition a piecewise Cox model assuming proportional hazards in a series of consecutive time intervals as proposed by Collett (6) will be investigated.

7.4.4 Subgroup analyses

Subgroup variables will be explored as described in Section 6.4 for the primary endpoint. The HR between the two treatments along with 95% CI and the p-value for test of treatment equality within each category of the subgroup as well as the p-value for the subgroup-by-treatment interaction will be estimated by the Cox proportional hazard model including the same covariates as in the primary analysis of primary endpoint, the subgroup variable if not part of the covariates of the primary analysis model, and subgroup-by-treatment interaction. If the subgroup variable is a categorization of a continuous covariate, this covariate will be dropped from the subgroup model. A forest plot will be presented with the estimated HR and the two-sided 95% CI for each subgroup category. The CIF plots will also be presented for each subgroup category.

For subgroup analyses of baseline LVEF and baseline uric acid, the interaction p-values will be calculated using trend tests, taking into account that the subgroup categorizations are ordered. Assuming the difference between the adjacent subgroup is the same, each subgroup is coded as numeric value ordinally and fitted into the model as numeric covariate. The model also includes terms of subgroup variable and subgroup-by-treatment interaction.

If there are less than 14 patients with events in one subgroup, then this subgroup will not be included in the model. If this leaves only one subgroup, the subgroup analysis will not be conducted.

For the continuous covariates baseline LVEF, baseline eGFR and age, the influence of the covariate will also be investigated on a continuous scale. For this purpose the continuous covariate will be added to the model if not already included and the interaction term of the continuous covariate and treatment will additionally be included into the model. The hazard ratio depending on the continuous covariate will be plotted and the interaction p-value will be reported.

7.5 SECONDARY ENDPOINTS

7.5.1 Key secondary endpoints

Occurrence of adjudicated HHF (first and recurrent)

Hospitalisation for heart failure will be analysed by a joint frailty model (JFM) that accounts for the dependence between recurrent HHF and CV death (7). The primary analysis will be based on all data available until completion of the planned treatment phase, including the data after end of treatment for patients not completing the treatment phase as planned.

Define $T_{i0} = 0$ and let $T_{i1}, T_{i2},\ldots, T_{iNi}$ be the recurrent event times for person i, where N_i is the number of recurrent HHF events before $X_i = \min(C_i, D_i)$, the minimum of an independent censoring time C_i and a dependent CV death time D_i. The JFM is defined through the hazard functions for the recurrent event process and CV death

$$r_i(t \mid \omega_i) = \omega_i \exp\{\beta_1 z_i\} r_0(t)$$
$$\lambda_i(t \mid \omega_i) = \omega_i^{\alpha} \exp\{\beta_2 z_i\} \lambda_0(t)$$

The recurrent heart failure hospitalisations intensity function for the i-th patient conditional on the patient specific random frailty, ω_i, is given by r_i and is proportional to the baseline intensity function, r_0. The conditional hazard function for time to CV death for patient i is given by λ_i, with the baseline hazard given by λ_0, and β_1, β_2 are p x 1 vectors of regression coefficients associated with vectors of covariates z_i. The same covariates as for the analysis of the primary endpoint will be used, e.g. $ß_{.1}$=treatment (empagliflozin, placebo), $ß_{.2}$=region (North America, Latin America, Europe, Asia and "other" including India, South Africa, and Australia), $ß_{.3}$=baseline status of diabetes (diabetes, prediabetes, no diabetes) etc.

Patient specific independent random effects are denoted by ω_i and are assumed to follow a gamma distribution with mean 1 and variance θ. The correlation of the recurrent events is quantified by θ, with higher values corresponding to greater within-patient correlation and also greater between-patient variability. The parameter α determines the relationship between the recurrent heart failure hospitalisations and time to CV death. When $\alpha < 0$, higher frailty will result in a greater risk of recurrence and lower risk of terminal event (i.e. a negative

correlation between the frailties), and when $\alpha > 0$, higher frailty will result in a greater risk of recurrence and is associated with a higher risk of CV death (i.e. a positive correlation between the frailties).

Let t_{ij} and x_i be the observed recurrent event times and follow-up, respectively. Denote by δ_{ij} and Δ_i, the indicator of the recurrent event at time t_{ij} and the indicator of CV death at time x_i, respectively. The likelihood for person i is then given by the following:

$$L_i = \int_{\omega_i} \prod_{j=1}^{N_i} \left[\omega_i r_i(t_{ij})\right]^{\delta_{ij}} \exp\left\{\int_0^{x_i} \omega_i r_i(t)dt\right\} \left[\omega_i^{\alpha} \lambda_i(x_i)\right]^{\Delta_i} \exp\left\{\int_0^{x_i} \omega_i^{\alpha} \lambda_i(t)dt\right\} f_{\theta}(\omega_i)d\omega_i.$$

Adopting piecewise constant hazards for the recurrent events and CV death allows estimation of the likelihood by Gaussian quadrature. The implementation of the used adaptive Gaussian quadrature techniques is incorporated into Proc NLMIXED of SAS 9.4. SAS Code following the strategy outlined in (9) will be used.

The size of the pieces for the piecewise constant hazards can be different and are determined separately for the terminal as well as recurrent event process with 10 pieces each. The nodes are defined by the empirical deciles of the recurrent or terminal events, respectively.
The joint frailty model using the multiplicative parametrization with non-normal random effects will be fitted using a likelihood-reformulation method (10).

To improve convergence of the model, linear covariates (in case no interaction with treatment is modelled) will be standardized prior inclusion into the analysis and starting values for the model parameters will be determined using the following procedure:

1) An exponential model will be fitted for the terminal event and a poisson regression model for the recurrent event process including the same covariates that are included in the final joint frailty model to get initial starting values for all parameters.
2) A simplified model without random effect ω which is otherwise equal (regarding covariates, baseline hazards) to the joint frailty model is fitted using the values from step (1) as starting values for the parameters. The estimated coefficients from the simplified model will then be used as starting values for the parameters of the piecewise-constant joint frailty model.

The joint model gives two distinct hazard ratios:

$HR_{HHF} = \exp\{\beta_{11}\}$ is the hazard ratio associated with the effect of treatment on the recurrent event rate of HHF, and $HR_{CVD} = \exp\{\beta_{21}\}$ is the CV death hazard ratio.

Estimates and 95% CI for the hazard ratios and for α (relationship between the recurrent heart failure hospitalisations and CV death) will be given.

The following sensitivity analyses will be conducted

- Based on the TS, including only any events up to 30 days after treatment discontinuation
- Instead of CV-death, jointly model HHF with all-cause mortality as the terminal events (otherwise same model as initially described)

- The endpoint will be additionally evaluated based on investigator defined events for HHF and CV death (otherwise same as initially described model).

- A parametric joint gamma-frailty model will model the recurrent event component using a Poisson distribution and model the CV mortality component using a exponential distribution, conditional on the frailty parameter. Individual frailties are again assumed to follow a Gamma distribution. Thus HHF rates follow a negative binomial distribution and times to CV death follow a Lomax distribution (see (8)) (otherwise – e.g. covariates - same as initially described model).

- The analysis will be repeated to include all events up to individual trial completion including the follow-up period (otherwise same as initially described model).

- The analysis will be repeated to include only confirmed HHF which meet the CDISC guidance criteria for hospitalization for heart failure (i.e., confirmed by adjudication committee, but excluding those missing a physical sign or laboratory test, or both).

- The analysis will be repeated to include all events up to cutoff dates prior to COVID-19 outbreak – cutoff (censoring) dates are 30Nov2019 for China and 31Dec2019 for all other countries.

- The analysis will be repeated to include all events up to 7 days prior a reported SARS-CoV-2 infection (based on broad scope SARS-CoV-2 infections)

In case the semi-parametric joint modelling cannot converge numerically with the existing SAS procedures, the parametric joint gamma-frailty model as described above may be used instead for the confirmatory analysis.

The number of HHF events per patients will be summarized descriptively. Additionally a negative binomial model will be fitted to the data of recurrent HHF. This will be done once including only treatment as covariate and once including all covariates as the primary model. Rate ratio and confidence intervals of both models will be reported.

The mean cumulative incidence will be displayed for adjudicated recurrent HHF.

Subgroup analyses will be explored as outlined in Section 6.4 for adjudicated recurrent HHF. For subgroup analyses the term of subgroup (if not already part of the model) and subgroup by treatment will be added to the model of the recurrent event. If the subgroup variable is a categorization of a continuous covariate, this covariate will be dropped from the subgroup model.

eGFR (CKD-EPI)$_{cr}$ slope of change from baseline

Slope in change from baseline of eGFR (CKD-EPI)$_{cr}$ will be analysed by a random coefficient model allowing for random intercept and random slope per patient. The model will include the factors treatment, sex, geographical region, and status of DM as fixed effects and eGFR (CKD-EPI)$_{cr}$ at baseline (continuous), LVEF (continuous), age (continuous), time and interaction of treatment by time and interaction of eGFR (CKD-EPI)$_{cr}$ at baseline

(continuous) by time as linear covariates and allow for randomly varying slope and intercept between patients. The model will include all on-treatment change from baseline data.

Since the slope is run on the change from baseline data, the intercept will model the acute drop, whereas the long-term effect is modelled by the slope.

A plot of individual patient slopes and separately of individual patient intercepts will be provided per treatment and by eGFR at baseline.

Subgroup analyses as outlined in Section 6.4 will be explored for eGFR slope. The subgroup model will include additionally to the model described above, the subgroup if not already part of the model, subgroup by treatment and subgroup by treatment by time. If the subgroup variable is a categorization of a continuous covariate, this covariate will be dropped from the subgroup model.

7.5.2 Other Secondary endpoints

Other secondary endpoints are exploratory. No correction for multiple hypotheses testing will be made.

- *Composite renal endpoint: Time to first event of chronic dialysis or renal transplant or sustained* reduction of ≥40% eGFR (CKD-EPI)$_{cr}$ or*
  - *(for patients with eGFR (CKD-EPI)$_{cr}$ ≥30 mL/min/1.73 m^2 at baseline): sustained eGFR <15 mL/min/1.73m^2*
  - *(for patients with eGFR (CKD-EPI)$_{cr}$ <30 mL/min/1.73 m^2 at baseline): sustained eGFR < 10 mL/min/1.73 m^2*

*An eGFR (CDK-EPI)cr reduction is considered sustained, if it is determined by two or more consecutive post-baseline central laboratory measurements separated by at least 30 days (first to last of the consecutive eGFR values). If there is no additional measurement >= 30 days after the eGFR reduction is observed and the patient dies within 60 days of this measurement, then the eGFR reduction is also considered sustained.

Chronic dialysis is regarded as chronic if the frequency of dialysis is twice or more per week for at least 90 days.

- *Time to first adjudicated HHF*
- *Time to adjudicated CV death*
- *Time to all-cause mortality*
- *Time to onset of DM in patients with baseline pre-DM*

All time-to-event endpoints will be reported in days.

The same model and data frame as used in the primary analysis of primary endpoint will be applied to all these time-to-event endpoints. If the endpoint does not include any cause of death, a CIF plot with all-cause mortality as competing risk will be displayed; otherwise, a CIF plot with causes of death not included in the endpoint as competing risk will be displayed. For all-cause mortality a Kaplan-Meier plot will be displayed.

For time to first HHF and time to CV death, the analysis will be repeated based on investigator defined events. In addition, the analysis of time to CV death will be repeated to include all events up to 7 days prior a reported SARS-CoV-2 infection (based on broad scope SARS-CoV-2 infections)

- *Change from baseline in clinical summary score (HF symptoms and physical limitations domains) of the Kansas City cardiomyopathy Questionnaire (KCCQ) at week 52*

Change from baseline in clinical summary score for HF symptoms and physical limitations domains of the KCCQ at week 52 will be evaluated by a mixed model for repeated measures (MMRM) including LVEF (continuous), age (continuous) and eGFR $(CKD\text{-}EPI)_{cr}$ at baseline (continuous) as linear covariates and, baseline score by visit, visit by treatment, sex, geographical region and status of diabetes at baseline as fixed effects. An additional factor of "week reachable" for the parameter in question, which for each patient is the theoretically reachable planned measurement based on the time of randomisation will adjust for the different planned study times.

All on-treatment data up to week 52 will be included and the analysis will be conducted on the treated set.

A sensitivity analysis will be conducted including data after discontinuation (OC-AD) on the randomized set.

An additional analysis will be performed for the Clinical Summary Score on OC-AD without imputing a worst score for patients who die.

- *Occurrence of all-cause hospitalisation (first and recurrent)*

A similar joint frailty model as well as negative binomial regression model as in the HHF will be analysed for all-cause hospitalization. Instead of CV death, all-cause mortality will be jointly modelled as the terminal events in the joint frailty model.

7.6 FURTHER ENDPOINTS

Time to event endpoints

Further time to event endpoints will generally be analysed in a Cox proportional hazards model similar to the primary analysis on RS. If the endpoint does not include any cause of death, a CIF plot with all-cause mortality as competing risk will be displayed; otherwise, a CIF plot with causes of death not included in the endpoint as competing risk will be displayed.

If there are less than 14 patients with events, then only descriptive statistics will be presented.

Further time to event endpoints are the following:

- Time from first to second adjudicated HHF
- Time to first all-cause hospitalisation
- Time to new onset of atrial fibrillation
- Time to adjudicated MI (fatal or non-fatal)
- Time to adjudicated stroke (fatal or non-fatal)
- Time to adjudicated TIA
- Time to first event of all-cause mortality or all cause hospitalisation
- Time to first event of adjudicated CV death or adjudicated non-fatal MI
- Time to first event of adjudicated CV death or adjudicated non-fatal stroke
- Time to first event of adjudicated CV death, adjudicated non-fatal MI or adjudicated non-fatal stroke (3-point MACE)
- Time to progression to macro albuminuria (defined as UACR >300 mg/g) from baseline for patients with baseline UACR ≤ 300 mg/g
- Time to first new onset of sustained normo− or micro albuminuria (UACR ≤ 300 mg/g) in patients with macro albuminuria at baseline
- Time to first new onset of sustained normo albuminuria (UACR < 30 mg/g) in patients with micro- or macro albuminuria at baseline
- Time to first event of composite renal endpoint[1] or adjudicated CV death
- Time to first event of composite renal endpoint[1], adjudicated CV death or adjudicated HHF
- Time to first event of composite renal endpoint[1] or all-cause mortality

- Time to first acute kidney injury (based on the preferred term)
- Time to non-cardiovascular (non-CV) death (death cases not included in the definition of adjudicated CV death)
- Time to first investigator defined CV hospitalization

[1]Composite renal endpoint defined as: *chronic dialysis or renal transplant or sustained* reduction of ≥40% eGFR (CKD-EPI)$_{cr}$ or*

- *(for patients with eGFR (CKD-EPI)$_{cr}$ ≥30 mL/min/1.73 m^2 at baseline): sustained eGFR <15 mL/min/1.73m^2*
- *(for patients with eGFR (CKD-EPI)$_{cr}$ <30 mL/min/1.73 m^2 at baseline): sustained eGFR <10 mL/min/1.73 m^2*

Chronic dialysis is regarded as chronic if the frequency of dialysis is twice or more per week for at least 90 days.

*An eGFR (CDK-EPI)cr reduction is considered sustained, if it is determined by two or more consecutive post-baseline central laboratory measurements separated by at least 30 days (first to last of the consecutive eGFR values). If there is no additional measurement >= 30 days after the eGFR reduction is observed and the patient dies within 60 days of this measurement, then the eGFR reduction is also considered sustained

<u>Continuous endpoints:</u>

The following endpoints will be evaluated by mixed models repeated measures (MMRM) as defined in the protocol.

- HbA_{1c} change from baseline over time (OC-AD on the randomised set and OC-OT on the treated set)
- SBP, DBP change from baseline over time (OC-AD on the randomised set and OC-OT on the treated set)
- Weight change from baseline over time (OC-AD on the randomised set and OC-OT on the treated set)
- NT-pro BNP relative change from baseline over time (OC-AD on the randomised set and OC-OT on the treated set) (value will be transformed to the log scale before MMRM analysis. Estimates obtained from the model will then be back-transformed and reported on the original scale.)
- KCCQ overall summary score change from baseline at week 52 (OC-AD on the randomised set and OC-OT on the treated set) (model will include all timepoints up to week 52)
- KCCQ individual domains change from baseline at week 52 (OC-AD on the randomised set and OC-OT on the treated set) (model will include all timepoints up to week 52)

- KCCQ total symptom score change from baseline at week 52 (OC-AD on the randomised set and OC-OT on the treated set) (model will include all timepoints up to week 52)

- KCCQ based on patient-preferred outcome change from baseline at week 52 (OC-AD on the randomised set and OC-OT on the treated set) (model will include all timepoints up to week 52)

- UACR relative change from baseline over time (OC-AD on the randomised set and OC-OT on the treated set) (value will be transformed to the log scale before MMRM analysis. Estimates obtained from the model will then be back-transformed and reported on the original scale.)

- eGFR $(CKD\text{-}EPI)_{cr}$ change from baseline over time (OC-AD on the randomised set and OC-OT on the treated set)

As outlined in the protocol, the above endpoints will be analysed in a mixed model with repeated measures (MMRM), including age, LVEF and eGFR $(CKD\text{-}EPI)_{cr}$ at baseline as linear covariates and visit by treatment interaction, baseline by visit interaction, geographical region, sex and status of diabetes at baseline as fixed effects. An additional factor of "week reachable" for the parameter in question, which for each patient is the theoretically reachable planned measurement based on the time of randomisation will adjust for the different planned study times.

An unstructured covariance structure will be used to model the within-patient errors.

Descriptive statistics will be calculated for the value at visit and change from baseline based on OC-AD on the randomised set and OC-OT on the treated set.

Additionally subgroup analyses will be performed as follows:

HbA_{1c} change from baseline will be evaluated by status of diabetes (non-DM, pre-DM, DM).

NTproBNP change from baseline will be evaluated by baseline atrial fibrillation/atrial flutter (as determined by baseline ECG).

UACR will be evaluated by UACR at baseline (<30 mg/g, ≥30mg/g to ≤300mg/g, >300 mg/g).

In order to evaluate the mean effect on eGFR after approximately 3 years, the above described MMRM models will be used in the following way: a mean effect of the timepoints week 124, 148 and 172 will be calculated. Of the three timepoints, only those with at least 5 patients with a measurement at that timepoint will be included in the average calculation. This will be done for the imputations of OC-OT and OC-AD.

For MMRM subgroup analyses, the interaction term treatment-by-visit will be replaced by a treatment-by-subgroup-by-visit interaction term.

To support analysis of renal function, eGFR throughout the trial will be categorized according to the following CKD staging: All calculations for the staging of renal function will be based

on the originally measured laboratory values, not on normalised values with BI standard reference ranges.

Table 7.6: 1 CKD staging

Stage	**eGFR (mL/min/1.73m²)**	**Description**	**Label for displays**	***Additional labels#***
1	≥90	Normal or high	≥90	≥90 (CKD 1)
2	60 to <90	Mildly decreased	60 to <90	60 to <90 (CKD 2)
3A	45 to <60	Mildly to moderately decreased	45 to <60	45 to <60 (CKD 3a)
3B	30 to <45	Moderately to severely decreased	30 to <45	30 to <45 (CKD 3b)
4	15 to <30	Severely decreased	15 to <30	15 to <30 (CKD 4)
5	<15	Kidney failure	<15	<15 (CKD 5)

A shift table from baseline to last value on treatment for eGFR $(CKD\text{-}EPI)_{cr}$ will be provided.

In cases where urine albumin values are reported to be below the quantification limits (e.g. <3 mg/L) the albumin / creatinine ratio is determined as missing and will not be replaced by estimated values.

Transitions from baseline to last value on-treatment based on the following UACR categories: normal (<30mg/g), microalbuminuria (30-<=300 mg/g and macroalbuminuria (>300 mg/g) will be presented.

For an evaluation of follow-up after discontinuation of trial medication an additional summary of the NT-proBNP, FG, eGFR, UACR, creatinine and the KCCQ endpoints as above will be produced. This summary will include descriptive statistics for baseline, actual values and change from baseline to last on-treatment value and follow-up, and change from last on-treatment assessment to follow-up. The TS-FU patient set will the basis for this summary, but only those patients with a valid baseline, last value on treatment and follow-up value will be included in all displays. Analyses will be repeated for patients with at least 52 weeks of cumulative exposure (excluding treatment gaps). FG analysis will also be conducted by DM status at baseline (3 categories).

Analyses of change from baseline to LVOT and FU on the TS-FU outlined above will be additionally modelled, separately for LVOT and FU. An analysis of covariance (ANCOVA) model including treatment group, sex, geographical region and status of diabetes at baseline

as fixed effects and baseline eGFR (CKD-EPI)cr (continuous), age (continuous), LVEF (continuous) and baseline of the endpoint (if not already included) as linear covariates will be used.

In addition to the ANCOVA models described above, descriptive statistics will be presented also for creatinine for the value at visit and change from baseline based on OC-AD on the randomised set and OC-OT on the treated set.

In addition to the analyses defined above, KCCQ clinical summary score, clinical symptoms and physical limitations domain will be analysed in two responder analyses (1) defining an increase of 5 points or more from baseline as improvement and (2) defining a reduction of 5 points or more as deterioration.

The frequency and percentage (relative to RS) of patients with improvement as well as the frequency and percentage of patients with deterioration at week 52 will be described. Patients lost to follow-up or who withdrew consent prior week 52 as well as patients who died prior week 52 will be considered as having a deterioration (no improvement) in KCCQ.

Improvement and deterioration will be analysed in separate logistic regression models including treatment, age (continuous), eGFR (CKD-EPI)cr at baseline (continuous), region, sex, baseline LVEF (continuous), status of diabetes at baseline, and baseline value of the KCCQ domain analyzed.

Win ratio:

An unmatched win ratio considering adjudicated CV death and adjudicated HHF will be analysed based on unmatched pairs. All patients randomised to empagliflozin will be compared to all patients randomised to placebo. Only common follow-up time will be considered for the comparison. Patients on empagliflozin are considered to have “won” the comparison if either the other patient has died due to CV causes while the patient on empagliflozin was still alive, or if both patients did not die due to CV causes, then if the other patient had more occurrences of HHF, or if the number of occurrences of HHF is the same but the time to the first occurrence of HHF is longer. The number of comparisons won is noted as N_W. Patients on empagliflozin are considered to have “lost” the comparison if the empagliflozin patient died due to a CV cause while the patient on placebo was still alive, or if both patients did not die due to a CV cause, then if the patient on empagliflozin had more occurrences of HHF, or if the number of occurrences is the same but the time to the first occurrence of HHF is shorter for the empagliflozin patient. The number of comparisons lost is noted as N_L. The win ratio is N_W / N_L.

The rules for winning and losing follow a modified Rogers 2014 (8) approach also considering the time to the first HHF event in case of a tie on the number of HHF events. The analysis of the unmatched win ratio will be conducted as described in Pocock 2012 (11) .

Other types of further endpoints

For other types of endpoints e.g. occurrence of adjudicated HHF within 30 days after first adjudicated HHF, change in NYHA class at week 52, HCRU and EQ-5D, descriptive statistics will be provided.

For NYHA class a shift table will also be provided for changes from baseline over time and for change from baseline to last value on treatment and last value in study and worst value in study.

The number of patients shifting from normal (no pre-diabetes) to pre-diabetes or diabetes as well as shifts from pre-diabetes to diabetes will be summarized.

Pharmacokinetic analysis

Descriptive statistics of trough concentrations of empagliflozin will be presented for patients, where PK data is available.

7.7 EXTENT OF EXPOSURE

There will be three methods of calculating exposure:

a. First intake to last intake of study drug, including off-treatment periods
b. First intake to last intake of study drug, excluding off-treatment periods
c. Overall observational period (randomisation until end of follow-up for vital status, see censoring for all-cause mortality in Section 6.8.3)

Descriptive statistics tables with mean, standard deviation (SD), median and range of the number of days a patient was on treatment will be provided. These tables will also provide the sum-total of the time (in years) that all patients were on treatment.

Frequency tables of number and percentage of patients belonging to categorical ranges of exposure weeks will be provided as well. Following are the categories of exposure-ranges (in weeks): 0 to <12 weeks, ≥12 to <26 weeks, ≥26 to <52 weeks, ≥52 to <78 weeks, ≥78 to <104 weeks, ≥104 to <156 weeks, ≥156 weeks. Categorical ranges may be adapted based on the actual duration of the study.

7.8 SAFETY ANALYSIS

The safety analysis will be based on the treated set (TS), treatment will be evaluated as randomised.

The AE analysis will include all adverse events (including outcome events as reported by the investigator).

While tables will generally display results by randomised treatment, listings will reflect whether a measurement/AE occurred on or off treatment.

7.8.1 Adverse events

Unless otherwise specified, the analyses of adverse events will be descriptive in nature and will be based on the number of patients with AEs and not on the number of AEs.

AEs will be coded using the most recent version of the Medical Dictionary for Regulatory Activities (MedDRA) coding dictionary.

Definitions of Boehringer Ingelheim customized MedDRA Queries (BIcMQs) for new MedDRA versions are maintained by the BI dictionary maintenance group.

For analysis multiple AE occurrence data on the CRF will be collapsed into an AE provided that all of the following applies:

- All AE attributes are identical (including LLT, intensity, action taken, therapy required, seriousness, reason for seriousness, relationship, outcome, AE of special interest and also additional information of specific AEs or AESIs such as source of sepsis (urinary tract or not) or type of genital infection (fungal balanitis or vulvovaginitis versus other than fungal balinitis or vulvovaginitis)).
- The occurrences were time-overlapping or time-adjacent (time-adjacency of 2 occurrences is given if the second occurrence started on the same day or on the day after the end of the first occurrence).

For further details on summarization of AE data, please refer to (2), (12).

7.8.1.1 Assignment of AEs to treatment

The analysis of adverse events will be based on the concept of treatment emergent adverse events. That means that all adverse events occurring between first drug intake till 7 days after last drug intake will be assigned to the randomised treatment. All adverse events occurring before first drug intake will be assigned to ‘pre-treatment’ and all adverse events occurring after last drug intake + 7 days with be assigned to ‘follow-up’.

In Section 15.3 general AE analyses tables will present only on-treatment AEs (applying the rule of 7 days for assignment as on-treatment) for the treatments groups (exceptions for urinary tract malignancies, hepatic events, lower limb amputations and bone fractures as well as adjudicated events see below). When looking at BIcMQs or standardized MedDRA queries (SMQs) and including all AEs up to completion of the trial, the time at risk for each patient will be calculated up to the individual day of trial completion (see Section 6.8).

Appendix 16.1.13.1 will include an analysis (overall summary table, frequency of AEs by system organ class (SOC) / preferred term (PT), frequency of serious adverse events (SAEs) by SOC/PT) where AEs and SAEs are assigned to the following phases: Screening, each treatment group, post-treatment for each treatment group.
The tables presenting frequency of AEs by SOC/PT and frequency of SAEs by SOC/PT will be repeated in Section 16.1.13.1 with treatment-specific post-treatment phase included, hereby also incidence rates for the post-treatment phase will be presented.

For listings, AEs will be assigned to one of the treatment phases of Screening, Placebo, Empa 10, Placebo post-treat, Empa 10 post-treat, post-study.

7.8.1.2 AE summaries

An overall summary of adverse events will be presented.

The frequency of patients with adverse events will be summarised by treatment, primary system organ class and preferred term. Incidence rates as defined in Section 7.8.1.7 will generally be included. AEs will also be reported by intensity (without incidence rates). Separate tables will be provided for patients with adverse events of special interest (AESIs), for patients with serious adverse events, for patients with drug related serious AEs, for patients with fatal events, for patients with AEs leading to discontinuation, and for patients with drug-related AEs.

The frequency of patients with adverse events occurring with incidence in preferred term greater than 2% by treatment will also be presented.

Overall AE summaries, AEs by SOC and PT, SAEs and AEs leading to discontinuation will additionally be investigated by subgroups as outlined in Table 6.4: 1.

Fatal AEs will be summarized up to end of the trial, also separately for those adjudicated as CV and those adjudicated non-CV cause.

The system organ classes will be sorted according to the standard sort order specified by the European Medicines Agency (EMA), preferred terms will be sorted by frequency (within system organ class).

Additionally, the following analyses will be reported in Appendix 16.1.13.1 for disclosure on EudraCT and clinicaltrials.gov:

- Frequency [N(%)] of patients with non-serious adverse events occurring with incidence in preferred term greater than 5% by treatment,
- Adverse Events per arm for disclosure on EudraCT by treatment”
- Non-serious Adverse Events for disclosure on EudraCT by treatment
- Serious Adverse Events for disclosure on EudraCT by treatment

For further details, see also (13).

7.8.1.3 Adverse events of special interest (AESIs)

Hepatic injury

Adverse events reported as AEs of special interest relating to hepatic injury as specified in the protocol will be summarised.

Additionally Hepatic injury AEs will be summarized based on an SMQ based definition. From SMQ Drug related hepatic disorders (20000006) the following narrow sub-SMQs will be used:

- Narrow sub-SMQ Liver related investigations, signs and symptoms (20000008)

- Narrow sub-SMQ Cholestasis and jaundice of hepatic origin (20000009)
- Narrow sub-SMQ Hepatitis, non-infectious (20000010)
- Narrow sub-SMQ Hepatic failure, fibrosis and cirrhosis and other liver damage-related conditions (20000013)

A table with frequencies of patients with these AEs by treatment, primary SOC and preferred term will be provided. This presentation will be repeated by DM status at baseline (DM vs no DM). Hepatic injury SAEs and hepatic injury AEs leading to disc based on the above SMQ definition will be presented.

In addition to the '7-day-on-treatment approach', a '30-day-on-treatment approach' will be presented for the overall hepatic adverse events based on the SMQ definition.

Patients with hepatic injury will be listed.

For presentations on adjudicated hepatic events, refer to Section 7.8.1.5.

Decreased renal function

Adverse events reported as AEs of special interest relating to decreased renal function as specified in the protocol will be summarised.

A frequency table of patients with AEs of acute renal failure by treatment, primary SOC and preferred term will additionally be provided based on the narrow standardized MedDRA query (SMQ) Acute renal failure (20000003).

This presentation will be repeated by the subgroups as outlined in Table 6.4: 1. SAEs and AEs leading to disc based on the SMQ Acute renal failure (20000003) will be presented.

In addition, frequency tables will be produced for patients with elevated creatinine ≥ 2 x baseline and > 1 x upper limit of normal (ULN).

Patients with decreased renal function will be listed.

Ketoacidosis

A frequency tables of patients with AEs related to ketoacidosis will be presented by treatment, primary SOC and preferred term for investigator defined cases and for the broad and narrow BIcMQ definition of diabetic ketoacidosis.

For the narrow BIcMQ diabetic ketoacidosis (DKA), AEs leading to discontinuation and an analysis by sex and diabetes status (type 1 diabetes, type 2 diabetes, pre-diabetes and normal (no pre-diabetes)) will be presented.

For the broad BIcMQ diabetic ketoacidosis (DKA), SAEs will be presented.

For presentations on adjudicated events, refer to Section 7.8.1.5.

Patients with DKA based on the narrow and broad BIcMQ or investigator defined ketoacidosis will be listed.

Events leading to lower limb amputation

A frequency table of patients with AEs leading to lower limb amputation as identified by the investigator by treatment, primary SOC and preferred term will be provided.

A separate table for AEs leading to lower limb amputation which are leading to discontinuation will be presented.

For events leading to lower limb amputations in addition to the '7-day-on-treatment approach' all adverse events that occurred between first study drug intake up to trial completion will be presented (following censoring rules like non-fatal outcome events). For both approaches, SAEs will be presented.

Lower limb amputations (up to trial completion) will additionally be summarised by renal function, level of first amputation, reason for first amputation, history of PAOD, previous amputation and status of DM at baseline (DM / no-DM).

Patients with lower limb amputation will be listed.

7.8.1.4 Specific AEs

Hypoglycaemic events

The investigator will record for each AE whether it represents a hypoglycaemic event and, if so, record additional information to assess the intensity of the hypoglycaemic event. On the basis of this information the hypoglycaemic event will be will be categorised as follows:

- severe hypoglycaemic episode: event requiring the assistance of another person to actively administer carbohydrate, glucagon or other resuscitative actions
- documented symptomatic hypoglycaemia with a measured glucose concentration < 3.0 mmol/L (< 54 mg/dL): event accompanied by typical symptoms of hypoglycaemia but no need for external assistance
- symptomatic hypoglycaemia and glucose concentration > 3.9 mmol/L (70 mg/dL)
- symptomatic hypoglycaemia and glucose concentration not measured
- asymptomatic hypoglycaemia: event not accompanied by typical symptoms of hypoglycaemia but with a measured glucose concentration ≤ 3.9 mmol/L (≤ 70 mg/dL)

Confirmed hypoglycaemic adverse event are defined as hypoglycaemic adverse events that had a glucose concentration ≤ 70 mg/dL or required assistance.

Different tables will be shown for (i) patients with investigator defined asymptomatic or symptomatic hypoglycaemia, and (ii) patients with confirmed hypoglycaemic adverse events,

i.e. hypoglycaemic adverse events that had a glucose concentration ≤ 70 mg/dL or required assistance.

Subgroup analyses on confirmed events with respect to age category, sex, race, renal function and diabetes background (type 1 diabetes, type 2 diabetes, pre-diabetes and normal (no pre-diabetes)) will be performed.

Time to the onset of the first confirmed hypoglycaemia will be displayed using a cumulative incidence function.

In addition the number of patients with hypoglycaemia according to SMQ (20000226) will be presented.

Patients with hypoglycaemic events will be listed.

UTI and genital infections

The following additional specific adverse events will also be assessed and be tabulated by treatment group:

- Genital infections (BIcMQ 'Infections" narrow subsearch "Genital tract infections predisposed to by glucosuria" and investigator assessment)
- UTI (BIcMQ "Infections" narrow subsearch "UTI predisposed by glucosuria" and investigator assessment)

Genital infections episodes based on investigator assessment will additionally be summarised by type (fungal balanitis or fungal vulvovaginitis, genital infection other than fungal balanitis or fungal vulvovaginitis), intensity (mild, moderate or severe), how the event was treated (no treatment, therapy assigned and number of antimicrobials needed to treat) and whether leading to discontinuation of treatment. The number of episodes per patient will also be presented.

UTIs episodes based on investigator assessment will additionally be summarised by intensity (mild, moderate or severe), anatomical location (upper UTI, lower UTI, asymptomatic bacteriuria), occurrence of pyelonephritis or urosepsis, how the event was treated (no treatment, therapy assigned and number of antimicrobials needed to treat) and whether leading to discontinuation of treatment. The number of episodes per patient will also be presented.

For UTIs based on the BIcMQ narrow subsearch the subgroups as outlined in Table 6.4: 1 will be presented. The same will be done for genital infections based on the BIcMQ narrow subsearch. Additionally they will be analysed by presence/absence of history of chronic or recurrent UTI or genital infection respectively.

Complicated urinary tract infections defined as serious adverse events of BIcMQ narrow subsearch 'UTI predisposed by glucosuria', all events of sub-BIcMQ 'Renal infections predisposed by glucosuria', all events of PT Urosepsis will be presented.

Complicated genital infection: defined as all serious events using the BIcMQ narrow subsearch 'Genital tract infections predisposed to by glucosuria' and all event of the narrow

subsearch 'Complicated genital tract infections predisposed to by glucosuria' will also be presented.

UTIs leading to discontinuation based on the BIcMQ narrow subsearch will be presented, the same will be repeated for genital infections leading to discontinuation based on the BIcMQ narrow subsearch .

Cumulative incidence functions will also be created for time to onset of the first UTI and for time to onset of the first genital infections, both based on the respective BIcMQ narrow subsearch .
Patients will UTIs or genital infections will be listed.

Pyelonephritis and sepsis
The following specific adverse event will also be tabulated by treatment group:

- Acute Pyelonephritis (based on investigator assessment): patient incidence overall and by sex
- Pyelonephritis or urosepsis (based on the sub-BIcMQ 'Renal infections predisposed by glucosuria' and the PT 'Urosepsis'): patient incidence overall, by sex and by status of DM at baseline (DM / no-DM)
- Sepsis (based on investigator assessment): patient incidence overall, by source of infection (UTI, not UTI or missing) and by sex and source of infection

Patients with pyelonephritis or sepsis will be listed.

Bone fracture events:

Frequency tables of patients with bone fracture by treatment, primary system organ class (SOC) and preferred term will be provided (based on the narrow BIcMQ 'Bone fractures' and investigator definition). Investigator defined fractures will be reported separately for traumatic and non-traumatic bone fractures.

For bone fractures based on the BIcMQ the subgroups as outlined in Table 6.4: 1 will be presented. Separate tables for bone fractures based on the BIcMQ, which are serious and those which are leading to discontinuation will be presented.

For overall bone fracture based on the BIcMQ in addition to the '7-day-on-treatment approach' all adverse events that occurred between first study drug intake up to trial completion will be presented (following censoring rules like non-fatal outcome events).

Patients with bone fractures will be listed.

Urinary tract malignancies:

Urinary tract cancerogenicity will be shown based on the BIcMQ 'Malignancies' – broad sub-search 14.1 'Urinary bladder and tract malignancies' and broad sub-search 14.2 'Renal malignancies' :

Presentation will be done ordered by HLT.

Frequency tables of patients with urinary tract malignancies by treatment, high level term and preferred term will be provided.

For urinary tract malignancies in addition to the '7-day-on-treatment approach' all adverse events that occurred between first study drug intake up to study end will be presented (following censoring rules like non-fatal outcome events).
There will be an additional analysis including all patients who had a minimum cumulative study drug exposure of 6 months (excluding treatment gaps). All AEs starting from date when 6 months cumulative exposure was reached up to individual day of trial completion will be shown (following censoring rules like non-fatal outcome events).

Patients with urinary tract malignancies will be listed.

Volume depletion

Volume depletion will be based on the BIcMQ 'Volume depletion of non-haemorrhagic cause and subsequent hypotension' – narrow sub-search 2 'Volume depletion and hypotension due to dehydration'.

A frequency table of patients with volume depletion by treatment, primary SOC and preferred term will be provided.

For volume depletion events the subgroups as outlined in Table 6.4: 1 will be presented. Separate tables for volume depletion events, which are serious and those which are leading to discontinuation will be presented.

Patients with volume depletion will be listed.

A cumulative incidence function will be used to display the time to first volume depletion event.

For the analysis of laboratory data, refer to Section 7.8.2.

Hypotension:

Frequency table of patients with symptomatic hypotension as defined by the investigator on the eCRF tick box by treatment, primary system organ class and preferred term will be provided.

Symptomatic hypotension will be presented by whether the intensity of diuretic medication was reduced and by whether the intensity of non-diuretic antihypertensive therapy was reduced.

Additionally hypotension will be presented by treatment, primary system organ class and preferred term. Hypotension is defined as preferred terms of the BIcMQ 'Volume depletion of non-haemorrhagic cause and subsequent hypotension' – narrow sub-search 2 'Volume

depletion and hypotension due to dehydration' (30000090) but excluding terms of the narrow subsearch 1 'Volume depletion due to dehydration'(30000089).

For hypotension events based on the project-defined subsearch the subgroups as outlined in Table 6.4: 1 will be presented.

The presentation of hypotension events based on investigator defined symptomatic hypotension will be repeated to show only events happening in the first 30 days after treatment start.

Separate tables for hypotension (BIcMQ) events, which are serious and those which are leading to discontinuation will be presented.

Separate tables for symptomatic hypotension events, which are serious and those which are leading to discontinuation will be also presented.

Patients with symptomatic hypotension events will be listed.

Additional subgroup analyses for selected adverse events of special interest or specific adverse events and other safety topics of interest are described in Section 9.7.

7.8.1.5 Events qualifying for external adjudication by the CEC and Hepatic External Adjudication and Adjudication of ketoacidosis

An independent external CEC regularly reviews events and evaluates whether pre-specified criteria for these adjudication endpoints are met. Details on composition of the CEC, responsibilities and clinical event definitions are provided in the separate CEC Charter.

Adjudication assessments will be incorporated into the database.

Details of the adjudication process are described in the CEC charter.

Cardiological/neurological adverse events:
Frequency tables will be provided based on SOC and preferred term for events qualifying for adjudication. This will include all trigger events up to the end of the study.

The number of patients with confirmed events per event type and breakdown of event subtype will be presented. This will be done for all CEC confirmed events up to the end of the study as well as up until completion of the planned treatment phase. Information will be provided for all patients as well as for patients with SARS-CoV-2 infection defined by broad scope SARS-CoV-2 infections. For the analysis in patients with SARS-CoV-2 infection, only events with adjudicated onset date between 7 days prior infection and completion of the planned treatment phase will be shown.

A listing will be provided, that shows the trigger events and result of adjudication.

Hepatic adverse events:
Frequency tables summarizing the relatedness and severity will be provided including all

adjudicated events up to 30 days after treatment stop. A listing will show all trigger events and adjudication results.

Ketoacidosis:
Frequency tables summarizing the adjudication results will be provided including all adjudicated events up to 7 days after treatment stop. A listing will show all trigger events and adjudication results.

7.8.1.6 AEs while patients taking wrong medication

A listing will be provided for AEs that occurred while a patient was taking the wrong medication. Off-drug is not viewed as wrong medication for the listing.

An additional adverse event table that assigns the adverse events to the actual treatment taken will be presented. A patient who took both the assigned treatment and also at least one tablet of the wrong treatment will be counted as at risk in both treatment groups for the relevant time. Off-treatment periods will be counted towards the treatment taken before the off-treatment period for the table. The table will include all adverse events by SOC and PT.

7.8.1.7 Adverse event incidence rates

In addition to the frequency tabulations, time-adjusted adverse event analyses will be performed for all AEs, investigator defined drug-related AEs, AEs leading to discontinuation, serious AEs, and adverse events of special interest by SOC, respectively HLT, and PT.

The time at risk in patient years for the on-treatment phase is derived as follows:

Patients with AE:
time at risk (AE) in days = date of start of AE with specified PT/SOC/HLT – study treatment start date + 1

Patients without AE:
time at risk (AE) in days = end date of time at risk – study treatment start date + 1, where end date of time at risk is the minimum of date of last study drug intake + x days and date of death, if applicable.

The standard approach will be x=7 days, but for certain AESIs in addition other approaches will be used.

Total AE-specific time at risk per treatment group is then derived as:
Time at risk (AE) [years] = Sum of time at risk [days] over all patients in a treatment group/ 365.25

For ‘each row of a table’ (e.g. displaying an SOC), time at risk is calculated using start of first AE summarized in this row; e.g. for patient with AE in a specified SOC, time at risk = date of start of AE with specified PT in this SOC – start of study treatment + 1.

The AE incidence rate per 100 patient years will then be calculated as follows:

Incidence rate per 100 patient years (pt-yrs) = 100 * number of patients with AE / time at risk (AE) [years].

In a similar way the time at risk and incidence rate for the post-treatment period is derived. Hereby the start date is the start date of the post-treatment phase instead of the study treatment start date.

7.8.1.8 COVID-19 related analyses

The subgroup of patients with SARS-CoV-2 infection will be analysed. SARS-CoV-2 infections will be defined as narrow scope by the BIcMQ SARS-CoV-2 infections. All analyses will be repeated using a broad scope defined as BIcMQ SARS-CoV-2 infections including the preferred term “Suspected Covid-19”.

An overview of adverse events will be presented for this subgroup of patients. Additionally the number of patients with adverse events, the number of serious AEs and AEs leading to discontinuation of study treatment will be presented by treatment, primary system organ class and preferred term. Only adverse events with onset date 7 days prior SARS-CoV-2 infection until the end of the on-treatment period will be included.

A listing will be prepared presenting all SARS-CoV-2 infections.

To assess the reporting of adverse events before and after the start of the COVID-19 outbreak, the number of patients with AEs including incidence rates will be presented for both time periods. The same cutoff dates as defined in Section 7.5.1 will be used.

7.8.2 Laboratory data

Standard safety tables will not include eGFR or creatinine, as those are shown separately as described in Section 7.6.

For continuous safety laboratory parameters standardized and normalized values will be derived as well as the differences to baseline. The process of standardization and normalization as well as standard analyses for safety laboratory data are described in the BI guidance for the Display and Analysis of Laboratory Data (14). All analyses considering multiple of times upper limit of normal (ULN) will be based on original and not normalized data.

Laboratory parameters will be shown in conventional units. Where conventional units differ from SI units, analyses will be repeated on SI units in the appendix.

The analyses of laboratory data will be descriptive in nature and will be based on BI standards [see Data Management and Statistics Manual (DM&SM): Display and Analysis of Laboratory Data (14)].

Baseline for safety laboratory parameters will be the last available measurement before the start of randomised trial medication. Laboratory measurements taken up to 3 days after the last administration of randomised study drug will be considered as on-treatment.
Default settings will be used for repeated values (using worst value).

Laboratory values will be compared to their reference ranges and frequency tables will be provided for the number of patients within and outside the reference range at baseline and the last measurement on treatment. Descriptive statistics will be provided by treatment group for baseline, last value on-treatment and for changes from baseline to last value on treatment. Frequency tables will summarise the number of patients with potentially clinically significant abnormalities as defined for the current XLAB macro.

Only patients with at least one available on-treatment value will be included in the analysis of an individual laboratory parameter. All individual data will be presented in listings.

To support analyses of liver related adverse drug effects, patients with Aspartate transaminase (AST) and/or Alanine transaminase (ALT) ≥3xULN with concomitant or subsequent Total Bilirubin (TBILI≥2xULN) in a 30 day period after AST/ALT elevation are of special interest. In addition, of these cases, it will be considered whether the alkaline phosphatase (AP) is less than 2 x ULN (maximum value in a 30 day period after AST/ALT elevation) or not. The start of the 30 day time span is triggered by each liver enzyme elevation above the defined thresholds. Patients who fulfil one or two of the criteria for ALT/AST or total bilirubin elevations above and have no information available for the remaining parameter(s) within the 30 day time window will not be listed under "ALT and/or AST ≥ 3xULN with TBILI ≥ 2xULN".

In addition ALT/AST will be used to investigate elevated liver enzymes:

- ALT/AST ≥ 3 x ULN
- ALT/AST ≥ 5 x ULN
- ALT/AST ≥ 10 x ULN
- ALT/AST ≥ 20 x ULN

All liver enzyme elevations until 30 days of treatment discontinuation will be shown.

Frequency tables of patients with elevated liver enzymes defined by ALT and/or AST, total bilirubin and AP combinations will be provided. A scatter plot of peak ALT against peak total bilirubin will be presented with reference lines for 3 x ULN ALT and 2 x ULN total bilirubin, including an indicator for treatment received. Details on patients with elevated liver enzymes will be listed.

For the following parameters: uric acid, total cholesterol, HDL-C, LDL-C, triglycerides, haemoglobin and haematocrit the time course of changes will be assessed. The analysis will be performed by applying MMRM models to OC-AD data (on the randomised set) and respectively OC-OT data (on the treated set). The MMRM models, that will be used, are specified in Section 7.6. These analyses will be conducted on data before any normalization.

The parameters LDL-C/HDL-C ratio and non-HDL cholesterol will be evaluated descriptively over time based on OC-AD data (on the randomised set) and OC-OT data (on the treated set). These analyses will be conducted on data before any normalization.

For an evaluation of follow-up after discontinuation of trial medication an additional summary of haemoglobin, haematocrit, uric acid and lipid parameters will be produced. This summary will include descriptive statistics for baseline, actual values and change from

baseline to last on-treatment value and follow-up, and change from last on-treatment assessment to follow-up. The TS-FU patient set will the basis for this summary, but only those patients with a valid baseline, last value on treatment and follow-up value will be included in all displays. Analyses will be repeated for patients with at least 52 weeks of cumulative exposure (exluding treatment gaps).

Analyses of change from baseline to LVOT and FU on the TS-FU outlined above will be additionally modelled, separately for LVOT and FU. An ANCOVA model including treatment group, sex, geographical region and status of diabetes at baseline as fixed effect and baseline eGFR $(CKD\text{-}EPI)_{cr}$ (continuous), age (continuous), LVEF (continuous) and baseline of the endpoint as linear covariates will be used.

7.8.3 Vital signs

An MMRM analysis for heart rate over time will be provided based on OC-AD on the randomised set and OC-OT on the treated set imputations. The model will follow the MMRM analysis described in Section 7.6.

Note that heart rate refers to the eCRF question on pulse rate.

7.8.4 ECG

Clinically relevant abnormalities found at physical examination or ECG at Visit 2 will be considered to have already existed prior to signing of informed consent and therefore should be considered baseline conditions instead of adverse events, unless there is good reason to assume that they first appeared after signing of consent.

Outcomes of ECGs will be part of the reporting of medical history or AE reporting. Categorical findings as collected in the eCRF will also be summarized descriptively.

7.8.5 Others

Frequency of pregnancies and pregnancy outcomes will be listed by treatment.

Results of the Modified Rankin Scale will be summarized descriptively.

8. REFERENCES

1	*CPMP/ICH/363/96:* "Statistical Principles for Clinical Trials", ICH Guideline Topic E9, Note For Guidance on Statistical Principles for Clinical Trials, current version
2	*BI-KMED-BDS-HTG-0035:* "Handling of missing and incomplete AE dates", current version; KMED.
3	*001-MCS 36-472*: "Standards and processes for analyses performed within Clinical Pharmacokinetics/Pharmacodynamics", current version; IDEA for CON.
4	*BI-KMED-BDS-HTG-0045*: "Reporting of Clinical Trials and Project Summaries", current version; KMED.
5	Beyersmann J, Allignol A, Schumacher M. Competing risks and multistate models with R. New York: Springer 2012 [R16-3839]
6	Collett D. Modelling survival data in medical research. Chapman and Hall/CRC 2003 [R07-4680]
7	Rogers JK, Yaroshinsky A, Pocock SJ, Stokar D, Pogoda J. Analysis of recurrent events with an associated informative dropout time: application of the joint frailty model. Stat Med 2016; 35 (13): 2195 - 2205 [R16-2229]
8	Rogers JK, Pocock SJ, McMurray JJV, Granger CB, Michaelson EL, Ostergren J, et al. Analysing recurrent hospitalizations in heart failure: a review of statistical methodology, with application to CHARM-Preserved. Eur J Heart Fail 2014; 16: 33 - 40 [R16-4909]
9	Li Lu, Chenwei Liu, Analysis of Correlated Recurrent and Terminal Events Data in SAS®. NESUG 2008 [R20-1940]
10	Liu L, Yu Z A likelihood reformulation method in non-normal random effects models. Stat Med 2008; 27 (16): 3105–3124 [R19-4129]
11	Pocock SJ, Ariti CA, Collier TJ, Wang D The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. Eur Heart J 2012; 33: 176 - 182 [R16-4813]
12	*BI-KMED-BDS-HTG-0041*: "Analysis and Presentation of Adverse Event Data from Clinical Trials", current version; KMED.
13	*BI-KMED-BDS-QRG-0010*: “Reference Guide for the preparation of AE, SAE outputs for results disclosure on clinicaltrials.gov and EudraCT (EU-CTR)", current version; KMED.
14	*BI-KMED-BDS-HTG-0042*: "Handling, Display and Analysis of Laboratory Data", current version; KMED.
15	*001-MCS-40-113:* "Clinical Trial Report (CTR) Phase I - IV, Interim Reports and Other Associated Regulatory Documents” , current version; IDEA for CON.

9. ADDITIONAL SECTIONS

9.1 REGIONS AND COUNTRIES

Countries will be assigned to regions following the assignment of the IRT system, which is outlined in Table 9.1:1. Listed countries include currently planned backup countries.

Table 9.1: 1 Regions and countries

Region	Country
Asia	China Japan Korea Singapore
Europe	Belgium Czech Republic Germany Hungary Italy Netherlands Poland Romania Spain UK
Latin America	Argentina Brazil Colombia Mexico
North America	Canada US
Other	Australia India South Africa

9.2 CONCOMITANT MEDICATION

Definitions of medication groups (such as ARBs, diuretics) will be based on World Health Organization Drug Dictionary (WHO DD) and will be stored in the PDMAP.

9.3 DEFINITION OF DM AT BASELINE

Table 9.3: 1 DM at baseline

Diabetes status at baseline	Definition
T1DM	Patients who fulfil one of the following: - investigator-reported medical history of diabetes on the medical history page AND - type of diabetes is T1DM
T2DM	Patients who do not have T1DM but fulfil one of the following: - investigator-reported medical history of diabetes on the medical history page OR - any pre-treatment HbA1c value >= 6.5 OR - stratification via IRT in group of diabetes with either missing assessment of history of diabetes on the medical history page, or no pre-treatment measurement of HbA1c available and investigator reported ‘no’ for the medical history of diabetes

Table 9.3: 1 DM at baseline (cont.)

Diabetes status at baseline	Definition
Pre-DM	Patient who fulfil one of the following: - investigator reported 'no' for the medical history of diabetes on the medical history page AND no pre-treatment HbA1c >=6.5 and a pre-treatment HbA1c value of >=5.7 and <6.5 OR - stratification via IRT in group of pre-diabetes with either o missing assessment of history of diabetes on the medical history page and no pre-treatment HbA1c >=6.5 or o no pre-treatment measurement of HbA1c and investigator reported 'no' for the medical history of diabetes OR - missing assessment of history of diabetes on the medical history page AND all pre-treatment HbA1c <6.5 and a pre-treatment HbA1c value of >=5.7 and <6.5 and stratification via IRT in group of no diabetes
Normal (excluding pre-DM)	Patients not meeting the criteria of DM or pre-DM

9.4 ADDITIONAL SUB-GROUP ANALYSIS FOR REGIONAL SUBMISSIONS

Disposition and demographics of the subpopulation for patients from USA and subgroup analyses for patient from the USA vs non-USA will be included in the appendix of the CTR. Efficacy endpoints evaluated will be primary endpoint, time to cardiovascular (CV) death, time to first HHF, HHF (first and recurrent) and renal slope. Safety will be summarized for patients from the USA.
Additional subgroup analyses will be conducted by region, baseline eGFR (>=90, 60 to <90, 45 to <60, 30 to <45, and <30), and by ethnicity for the same endpoints (primary endpoint, time to cardiovascular (CV) death, time to first HHF, HHF (first and recurrent) and renal slope). For baseline eGFR (5 category) subgroup analysis, the interaction p-value will be calculated using a trend test.

Additional country or region-specific analyses will be conducted for patients from East-Asia (China, Japan and Korea), China, and Japan and included into the country-specific submission documents as also outlined in (15). Primary endpoint, time to cardiovascular (CV) death, time to first HHF, HHF (first and recurrent) and renal slope will be presented. Main adverse event overviews, disposition, demographics will be presented.

9.5 INTERIM ANALYSES

An interim analysis will be conducted by the DMC as outlined in the CTP. If based on the interim analysis it is decided to stop the trial for overwhelming efficacy, this analysis used by the DMC will be considered the primary analysis for the primary endpoint. This analysis will be repeated using the data of the final database lock. All other analyses including analyses of the key secondary endpoints will be conducted on the final database lock to provide most complete available data.

Only events that occurred up to one day prior the predefined interim analysis cut-off date will be considered for the analysis of the primary endpoint as well as CV death at interim. Patients without events up to this time will be censored at the predefined interim analysis cut-off date.

9.6 DETERMINATION OF SAMPLE SIZE

The trial is designed to achieve a power of 90% for a two-sided test with α=0.05 and hazard ratio 0.80. To achieve this at least 841 patients with a primary event are required.

Including the interim analysis with Hwang, Shih and De Cani α-spending function at 60% information with parameter γ = -8 will diminish the power only slightly to 89.98%.

The drop-out rate from the trial is assumed to be very low (<1% per year) and is not considered for determination of sample size.

A 10% yearly event rate in the placebo group is assumed. Assuming an accrual period of 18 months and a follow-up period of 20 months, it was planned to randomise 4126 patients to accrue 841 events.

The event rate and recruitment progress is assessed in a blinded manner during recruitment before any interim unblinding. If the accumulated blinded data suggests a slower accrual of

primary outcome events over calendar time than originally projected, then the number of randomised patients may be increased to a maximum of 6000 patients. Operationally, the recruitment period would be extended and could continue up to 6 months before the target number of events is expected to be achieved. The number of confirmed primary outcome events will not be affected by this consideration and will remain 841 events.

While monitoring the event rate in the above described manner, it was planned to include approximately 5750 patients in the trial.

9.7 FURTHER SAFETY ANALYSES

The following analyses will be provided:

PT Fall

Number of patients with fall (PT only) will be presented by baseline LVEF (<50, 50 to <60, >=60).

A cumulative incidence function will be used to display the time to first fall.

The number of patients with fall and concomitant (+/- 7 days) occurrence of volume depletion (narrow BIcMQ), hypotension (narrow BIcMQ), ventricular tachyarrthymia (SMQ) or hypoglycaemia (inv-def. or narrow SMQ) will be presented.

Acute renal failure (defined by narrow SMQ)

In addition to the analyses described in section 7.8.1.3 number of patients with acute renal failure will presented who have concomitant (defined as onset within +/- 7 days respectively +/- 30 days) of onset of volume depletion (narrow BIcMQ), PTs Dehydration or Volume depletion, or symptomatic hypotension (investigator-defined).

In addition to the subgroups as outlined in Table 6.4: 1 the number of patients with acute renal failure will be presented by baseline use of NSAIDs.

The number of patients with acute renal failure within 30 days after on-treatment start of selected medication among those who were not treated with the medication at baseline will be presented for any diuretics, loop or high-ceiling diuretics, ACE-inhibitor, ARB or ARNi, and for NSAIDs.

Ventricular Tachyarrhythmia

The number of patients with ventricular tachyarrhythmia (VT, defined by narrow SMQ Ventricular Tachyarrhythmia) will be presented. VT will also be reported by intensity (without incidence rates). Separate tables will be provided for patients for patients with serious VTs, for patients with drug related VTs, for patients with fatal VTs, for patients with VTs leading to discontinuation and by prevalence of ICD/CRT-D at baseline.

In addition the number of VT episodes per patient based on the SMQ definition and considering only the PT VT only will be presented.

Sepsis

In addition to the analyses described in section 7.8.1.4 investigator defined sepsis with source of infection Non-UTI and separately with missing source of infection will be presented by age (5 cat.), eGFR (5 cat.) and DM Status (2 cat.).

The number of patients with sepsis defined by narrow SMQ "Sepsis" excluding the PT Urosepsis will be presented overall and by age (5 cat.), sex, eGFR (5 cat.) and DM status (2 cat.).

Volume Depletion/Hypotension

Volume depletion and hypotension (both based on BIcMQ) as defined in section 7.8.1.4 will be presented additionally by baseline SBP (<110, 110 – 120, > 120).

Other topics:

For the following safety topics a frequency table by treatment, primary system organ class and preferred term and a subgroup analysis by diabetes status at baseline (2 cat.) will be

presented:

- Allergic skin reactions (defined by BIcMQ 'Skin eruptions', subsearch 1.2 'Allergic skin reactions excl. angioedema and application site reactions'(30000132))
- Increased urination (PT-based defined, stored in PDMAP)
- Thirst (PT-based defined, stored in PDMAP)
- Serum lipids increased (defined by BIcMQ 'Lipid metabolism disorders', subsearch 3 'Hyperlipidaemia' (30000131))
- Angioedema (defined by narrow BIcMQ Angioedema excl. Urticaria (30000130), overall and separately for terms including and excluding the term "urticaria")
- Hypersensitivity reactions (defined by narrow SMQ 'Hypersensitivity' (20000214))

Genital infection

Genital infection as defined in section 7.8.1.4 (narrow BIcMQ) will be presented additionally by sex and DM status (2 cat.).

Hypoglycaemia

Investigator reported confirmed hypoglycaemia (see section section 7.8.1.4 for definition) will be presented by baseline use of sulfonylureas, by baseline use of insulin, by baseline use of sulfonylureas or insulin.

10. HISTORY TABLE

Table 10: 1 History table

Version	Date (DD-MMM-YY)	Author	Sections changed	Brief description of change
Final	**21-FEB-2017**	[redacted]	None	This is the final TSAP without any modification
Revised	**23-JAN-2020**		4, 6.6.1, 6.8.2, 6.8.3, 7.4.1, 7.4.2, 7.5.1	Exclude FU period from ITT efficacy analyses and add sensitivity analysis including all events for primary endpoint and key secondary endpoint of recurrent HHF.
			4, 7.5.1	Include baseline by time interaction in the slope model
			5.3	Add time to investigator reported CV hospitalization. Definition of CV hospitalisation clarified Fasting plasma glucose was changed to fasting glucose because it was measured in serum and not plasma
			5.4, 6.3, 6.4, 7.6 and where applicable	Clarification of follow-up set, clarification of win ratio to take into account CV death, clarification of definition of atrial fibrillation at baseline vs history of atrial fibrillation, clarification of patient-preferred KCCQ score, other minor clarifications
			6.1, 7.8.1.6	Add analysis of AEs accounting for treatment switchers
			6.2 and where applicable	change important protocol “violations” to “deviations” and other minor wording changes.
			6.2	Clarifications in iPD table
			6.2, 6.3	Changes to IPD D2.02 and TS-FU set to reflect change in CTP with regards to open label SGLT-2 inhibitor use after EOT visit
			6.3, 7.6	Add PK set
			6.4	Adjustment of some subgroup analyses both for safety and for efficacy. Some subgroups added/replaced. Based on recommendation from the Executive Steering Committee, the subgroup analyses specified in the TSAP was reduced to the clinically most meaningful ones.
			6.4, 7.8.1	Update of planned subgroup analyses for safety.
			6.6, 7.5.2	KCCQ-CSS add sensitivity without imputation
			6.6.1	Imputation of missing covariates added.

Table 10: 1 History table (cont.)

Version	Date (DD-MMM-YY)	Author	Sections changed	Brief description of change
			6.6.2 and 6.8.2	Clarification of time at risk and missing data imputations
			6.7	Add definition of baseline medication
			6.8.2	Since recurrent event analyses are based on adjudicated event dates and duration of hospital stays is not adjudicated, only the onset dates of the events are considered for the analysis.
			7.1	Modified focussed baseline presentations (e.g. show HbA1c only for patients with DM, show atrial fibrillation also by AF at baseline, include LVEF , do not include height, LDL, HDL)
			7.2	Modified presentation of medication and non-medication therapy as baseline and separately newly introduced after baseline
			7.4	Clarified the primary estimand of the trial
			7.4.2	Additional sensitivity analyses added
			7.4.2	Clarified definition of investigator reported endpoints.
			7.4.4	Section on trend test removed since not applicable due to update on subgroups
			7.5.1	Joint frailty clarification on analysis details
			7.5.1	Clarify that other secondary endpoints are exploratory
			7.5.1	Add individual patient slope and intercept plots by eGFR at baseline.
			7.5.1	Multiple imputation for eGFR slope analysis was removed. Under missing at random assumption, the specified multiple imputation approach is asymptotically equivalent to the maximum likelihood estimate and therefore not required in addition.
			7.5.2	Negative binomial regression model for occurrence of all-cause hospitalisation (first and recurrent) added.
			7.5.2	Modify renal endpoint as per protocol amendment

Table 10: 1 History table (cont.)

Version	Date (DD-MMM-YY)	Author	Sections changed	Brief description of change
			7.5.2, 7.6	Specify MMRM in condensed form (do not include individual terms in case already included as interaction terms with visit or treatment) to be in line with BI internal standards on MMRMs. Align wording for MMRM term status of diabetes at baseline as well as factors to be included in MMRMs for the other secondary as well as further endpoints.
			7.6	Responder analyses (improvement and deterioration of 5 points or more from baseline) for KCCQ clinical summary score, clinical symptoms and physical limitations domain was added.
			7.6	Add NTproBNP change from baseline by atrial fibrillation or flutter at baseline
			7.6	Rules defining 'win' and 'lose' for the win ratio calculation were modified.
			7.6	Add analysis to mean change from baseline after approximately 3 years
			7.8.1	Changes in the safety analysis strategy in line with empagliflozin strategy (e.g. look specifically into urinary tract malignancy instead of overall malignancy, delete VTE, deleted other significant AEs, update of BIcMQ definitions, add hypotension)
			9.3	Clarification of diabetes definition
			9.5	Handling of cut-off date for interim analysis clarified.
			9.5	Clarification of primary analysis in case of interim stopping
			9.6	Region and ethnicity subgroup analyses added to regional analyses for US
			Throughout and section 9.6	Include wording from the protocol to the TSAP and clarify sample size
Revised	**12-FEB-2021**	[redacted]	6.2	Clarification on iPD
			6.4, 7.4.4,	Addition of baseline uric acid to subgroups and define trend test for baseline uric acid and baseline LVEF subgroup analyses
			6.8.3, 7.8.1.1	Clarifications with regard to censoring rules
			7, 7.4.2, 7.5.1 7.5.2 7.8.1.5	Addition of sensitivity analyses for COVID-19
			7.8.1.4	Insert word "narrow" for clarification

Table 10: 1 History table (cont.)

			7.8.1.8	Addition of safety analyses for patients with SARS-CoV-2 infection
			9.4	Addition of baseline eGFR (5 categories) subgroup for regional submission
			9.7	Addition of safety analyses to support labelling and special topics of interest