Statistical Analysis Plan (SAP) for

Empagliflozin to elderly and obese patients with increased risk of developing heart failure: the *Empire Prevent* trial program

Administrative information

Trial registration

Clinicaltrialsregister.eu, EudraCT Number2020-005317-40.

ClinicalTrials.gov, Registration numbers:

Empire Prevent Cardiac: NCT05084235
Empire Prevent Metabolic: NCT05042973

Protocol version

Protocol version 1.8, Mar 30th 2023

SAP revisions

Revision history: None
SAP revisions in relation to interim analyses etc.: None

Roles and responsibility

Camilla Fuchs Andersen, MD[1,2]: SC member. Sub-investigator, Herlev and Gentofte Hospital
Julie Hempel Larsen, MD[3,4]: Steering Committee (SC) member. Sub-investigator, Odense University Hospital
Jacob Eifer Møller, MD[3,4]: SC member. Sponsor and primary investigator, Odense University Hospital
Morten Schou, MD[1,2]: SC member. Primary investigator, Herlev and Gentofte Hospital
Julie Lyng Forman, MSc PhD[5]: Statistical advisor.

[1]Department of Cardiology, Herlev-Gentofte University Hospital, Copenhagen, Denmark
[2]Faculty of Health and Medical Sciences, Copenhagen University, Denmark
[3]Department of Cardiology, Odense University Hospital, Denmark
[4]Faculty of Health Sciences, University of Southern Denmark, Denmark
[5]Section of Biostatistics, Copenhagen University, Denmark

Statistical Analysis Plan Approval Signature Page

Person writing the SAP:

Name: Camilla Fuchs Andersen
Position/Title: MD, PhD-student

23.01.25

Signature & Date

Approved by:

Name: Julie Hempel Larsen
Position/Title: MD, PhD-student

23.01.25

Signature & Date

Name: Julie Lyng Forman
Position/Title: MSc, PhD, statistical advisor

24.01.25

Signature & Date

Name: Prof. Jacob Eifer Møller
Position/Title: MD, DMSc, sponsor, primary investigator

Jacob Eifer Møller Digitalt signeret af Jacob Eifer Møller
Dato: 2025.01.23 20:13:41 +01'00'

Signature & Date

Name: Prof. Morten Schou
Position/Title: MD, PhD, primary investigator

Signature & Date

Table of contents

Introduction

This statistical analysis plan (SAP) is an addendum to the Empire Prevent trial design paper,[1] which elaborates on the analysis plan where we consider that supplementary information is warranted.

Aim

The aim of this clinical trial program is to investigate the cardiovascular, metabolic and renal effects of 180 days treatment with the sodium glucose co-transporter 2 (SGLT2) inhibitor empagliflozin in elderly and obese patients with increased risk of future heart failure (HF) in order to evaluate the drug's therapeutic potential for HF prevention.

Main hypotheses

We hypothesize that 180 days treatment with 10 mg of empagliflozin in elderly and obese patients with increased risk of future HF will result in the following compared to placebo:

1) an increase in exercise capacity (peak VO_2 measured in ml/min/kg) and a reduction in LV mass index (g/m^2) (*Empire Prevent Cardiac*)
2) a reduction in ventricular EAT mass (g) and eECV (L) (*Empire Prevent Metabolic*)

Secondary hypotheses

Empire Prevent Cardiac

It is hypothesized that 180 days of treatment with 10 mg empagliflozin is further associated with:

a) a reduction in left ventricular end-diastolic volume index (LVEDVI), (ml/m^2)
b) a reduction in left ventricular end-systolic volume index (LVESVI), (ml/m^2)
c) an increase in left ventricular ejection fraction (LVEF), (%)
d) an increase in daily activity level measured by an accelerometer (daily counts per minute, cpm)

Empire Prevent Metabolic:

It is hypothesized that 180 days of treatment with 10 mg empagliflozin is further associated with:

a) a reduction in pericardial adipose tissue (PeAT) volume, (cm^2)
b) a reduction in total EAT volume, (cm^2)
c) a reduction in estimated plasma volume (ePV), (L)
d) a reduction in right ventricular end-diastolic volume index (RVEDVI), (ml/m^2)

Methods

Trial design and Randomization

As described in the trial design paper: *"The Empire Prevent trial program is an investigator-initiated, prospective, double-blind, randomized, placebo-controlled clinical superiority trial program comparing the treatment effect of empagliflozin with placebo in elderly and obese patients with increased risk of developing HF. The trial program consists of 2 parallel clinical trials: the Empire Prevent Cardiac trial and the Empire Prevent Metabolic trial for which the results will be interpreted and published separately.*

Risk factors for HF are defined as at least 1 manifestation of hypertension, CV [cardiovascular] disease and/or CKD [chronic kidney disease], but no history of diabetes or HF. Eligible patients will be randomized 1:1 to 180 days treatment with either 10 mg empagliflozin daily or matching placebo." [1]

The *Empire Prevent Cardiac* trial and the *Empire Prevent Metabolic* trial are designed to include N=204 and at least N=120 to allow for 10% drop-out. Please refer to the design paper for sample size calculations.[1]

Timing of outcome assessments

All endpoints are assessed at the baseline visit (day 0) and at the study closure visit (day 180±15 days), with the following exceptions: daily activity level, cardiac magnetic resonance imaging (MRI), DXA scan and echocardiography may be performed ±21 days from the baseline and study closure visits, respectively.

End of trial

Each trial is terminated, when the last participant has finished the final visit 2 (180 days), including the cardiac MRI.

Timing of main analyses

Prior to unblinding, all data on primary and secondary endpoints are collected, and all MRI analyses of primary and secondary endpoints are completed collectively. Thus, MRI analyses of primary and secondary endpoints are blinded.

Statistical Principles

Analysis populations

- The *intention-to-treat (ITT) population* will include all randomized patients, regardless of their eligibility and regardless of early discontinuation of treatment, according to the treatment they were randomized to receive.
- The *per-protocol population* will include randomized patients that have a baseline visit and a study closure visit regarding the endpoint (primary or secondary) of interest, that have an adherence of at least 80% to study medication, and that have no major protocol deviations.
- The *safety population* will include all randomized patients who received at least one dose of the study medication and will be analysed according to the treatment they actually received.

Adherence and protocol deviations

In this study, one pill is taken every day. A medicine bottle containing 90 pills are handed out twice: 1) at the start of study medication, and 2) after 90(-15) days. After 90(-15) days, excess pills are counted. If pills are counted before 90 days, the remaining treatment period will be calculated. In these cases, excess pills will be administered along with the second pill container to ensure that the patient has enough study medication throughout the study period.

After 180 (-15) days, excess pills are returned for accountability. Adherence in percent (% adherence) will be calculated as (number of pills taken / number of pills scheduled) *100. The number of pills scheduled will be calculated as the duration of treatment (end of study medication - start of study medication). The number of taken pills will be calculated as (180 - returned pills). In case number of pills taken exceeds the number of pills scheduled, adherence will be reported as 100%. Non-adherence is defined as adherence < 80%. Adherence reported as percentage will be presented by randomization group as median [interquartile range, IQR]. Also, the number and percent of participants taking more than 80% of the prescribed study medication for the duration of the treatment will be presented by randomization group. Adherence and protocol deviations will be summarized in separate tables, see Appendix A.

Definition of protocol deviations:

- Major: Incorrect data being collected and documented, errors in applying inclusion/exclusion criteria.
- Minor: Missed follow-up visits or single examinations.

Screening data

The following summaries will be presented for all screened patients in both trials:

1) Start and end dates of recruiting,
2) The number of patients screened
3) The number of screened patients not recruited and the reasons for non-recruitment
4) The number of patients recruited.

The number of screened patients includes patients excluded before the screening visit and patients excluded between the screening visit and randomization, and these numbers will be summarized accordingly.

Eligibility

The list of in- and exclusion criteria (including full specifications and definitions of the conditions above) can be found in the protocol and in the design paper.[1] The number of ineligible patients randomized, if any, will be reported, with reasons for ineligibility.

Withdrawal/follow-up

The level of consent withdrawal will be tabulated (classified as "withdrawal from intervention only'' or "withdrawal from intervention, follow-up and data collection''). The timing of withdrawals, lost to follow-up and exclusion from analysis will be presented in a modified CONSORT flow diagram with numbers and reasons given at each time point, see Appendix B.

Baseline patient characteristics

Patients will be described both overall and separately for the two randomized groups in each trial with respect to age, sex, body mass index (BMI), race, heart rate, systolic blood pressure, medical history (including ischemic heart disease (IHD), atrial fibrillation, and new onset diabetes mellitus), NT-pro-BNP, hbA1c, estimated GFR (including mean and proportion of patients with an estimated GFR < 60 ml/min/1.73m2), and relevant medication (including long acting nitrates, lipid-lowering agents and anticoagulant agents). New onset diabetes mellitus is defined as a HbA1c equal to or above 48 mmol/mol at more than one visit in patients without a previous history of diabetes mellitus. Baseline characteristics will be summarized descriptively in a table, see Appendix C. Formal

statistical comparisons of baseline characteristics will not be performed, as recommended in the CONSORT statement.

Analysis

Endpoint definitions

Two main papers are planned for each trial, in which primary and secondary endpoints will be reported (see tables below). The trial program has several exploratory endpoints in addition to the primary and secondary endpoints. This includes, but is not restricted to, the impact of empagliflozin on left atrial maximal volume index, systolic blood pressure, glomerular filtration rate, haematocrit, haemoglobin, erythropoiesis, changes in body composition, insulin resistance and sensitivity, tissue fibrosis and biomarkers of inflammation and adipocyte function in abdominal subcutaneous adipose tissue biopsies.

All prespecified endpoints will be reported following the recommendations of the CONSORT statement with descriptive data from each randomization and estimated treatment differences of empagliflozin vs placebo presented with confidence intervals and p-values adjusted for multiple testing, see Appendix D. The full list of pre-specified endpoints in both trials are listed below.

Primary main endpoints of *Empire Prevent Cardiac:*			
Definition	**Paper**	**Assessment method**	**Confidence level**
LV mass index, g/m^2	Paper I	Cardiac MRI	97.5%
Peak VO_2, ml/min/kg	Paper II	CPET	97.5%
Secondary endpoints of *Empire Prevent Cardiac:*			
Definition	**Paper**	**Assessment method**	**Confidence level**
LVEDVI, ml/m^2	Paper I	Cardiac MRI	95%
LVESVI, ml/m^2	Paper I	Cardiac MRI	95%
Left ventricular ejection fraction, %	Paper I	Cardiac MRI	95%
Daily activity level, accelerometer counts per minute	Paper II	Accelerometry	95%
Exploratory endpoints of *Empire Prevent Cardiac:*			
Definition		**Assessment method**	**Confidence level**
Non-indexed peak VO_2, ml/min		CPET	95%
Haematocrit, volume fraction		Blood samples	95%
Height-indexed LV mass, g/m		Cardiac MRI	95%
LA maximal volume index, ml/m^2		Cardiac MRI	95%
LA emptying fraction, %		Cardiac MRI	95%
Diastolic function, E/e', peak E wave, Peak A wave, E/A ratio and E-deceleration time.		Echocardiography	95%
LV and LA Global longitudinal strain, %		Echocardiography	95%
Self-reported quality-of-life score: Physical summary score		SF-36 questionnaire	95%
Self-reported quality-of-life score: Mental summary score		SF-36 questionnaire	95%
Aortic distensibility		MRI	95%
NT-proBNP		Blood samples	95%
Troponin I, ng/L		Blood samples	95%
GWI, mmHg%		Echocardiography	95%
GWE, %		Echocardiography	95%
Diffuse Cardiac fibrosis		Cardiac MRI, T1 mapping	95%
RVEDV, ml		Cardiac MRI	95%
RVEDVI, ml/m^2		Cardiac MRI	95%
Exercise load, W		Cardiopulmonary exercise test	95%

Abbreviations: CPET, cardiopulmonary exercise test; g, gram; GWE, global work efficiency; GWI, global work index; kg, kilogram; L, litre; LA, left atrial; LV, left ventricular; LVEDVI, left ventricular end-diastolic volume index; LVESVI, left ventricular end-systolic volume index; m, meter; min, minute; ml, millilitre; mmHG, millimetres of mercury; MRI, magnetic resonance imaging; ng, nanogram; NT-pro-BNP, N-terminal pro-brain natriuretic peptide; RVEDV, right ventricular end-diastolic volume; RVEDVI, right ventricular end-diastolic volume index; SF-36, 36-item short form survey; VO_2, oxygen consumption.

Primary main endpoints of *Empire Prevent Metabolic:*			
Definition	**Paper**	**Assessment method**	**Confidence level**
Ventricular EAT mass, g	Paper I	Cardiac MRI, SAX	97.5%
eECV, L	Paper II	Calculation[a]	97.5%
Secondary endpoints of *Empire Prevent Metabolic:*			
Definition	**Paper**	**Assessment method**	**Confidence level**
PeAT volume, cm^2	Paper I	Cardiac MRI, 4CV	95%
Total EAT volume, cm^2	Paper I	Cardiac MRI, 4CV	95%
ePV, L	Paper II	Calculation[b]	95%
RVEDVI, ml/m^2	Paper II	Cardiac MRI	95%
Exploratory endpoints of *Empire Prevent Metabolic:*			
Definition		**Assessment method**	**Confidence level**
Haematocrit, volume fraction		Blood samples	95%
Haemoglobin, mmol/L		Blood samples	95%
RVEDV, ml		Cardiac MRI	95%
IL-6, pg/ml		Blood samples	95%
Hs-CRP, mg/L		Blood samples	95%
Total body adipose tissue, g		DXA scan	95%
Total body lean mass, g		DXA scan	95%
Visceral adipose tissue, g		DXA scan	95%
Matsuda Index		OGTT	95%
HOMA-IR		OGTT	95%
EPO, pg/ml		Blood samples	95%
Erythroferrone		Blood samples	95%
Hepcidin		Blood samples	95%
Transferrin		Blood samples	95%
Transferrin saturation		Blood samples	95%
Reticulocyte count		Blood samples	95%
Myocardial iron content		Cardiac MRI, T2*	95%
Measured GFR, $ml/min/1.73m^2$		DTPA clearance	95%
Estimated GFR, $ml/min/1.73m^2$		Blood samples	95%
Albuminuria, mg/l		Urine samples	95%
Albumin/creatinine ratio, mg/g		Urine samples	95%
Serum urate		Blood samples	95%
Lipoprotein A		Blood samples	95%
TNF-∝, pg/ml		Blood samples	95%
Adiponectin, ug/ml		Blood samples	95%
Leptin, ng/ml		Blood samples	95%
GDF-15		Blood samples	95%
Galectin-3		Blood samples	95%
Endotrophin (PRO-C6)		Blood samples	95%
Inflammatory biomarkers Biomarkers of adipocyte function and fibrosis		Subcutaneous adipose tissue biopsy	95%

Abbreviations: DXA, dual-energy X-ray absorptiometry; EAT, epicardial adipose tissue; eECV, estimated extracellular volume; EPO, erythropoietin; ePV, estimated plasma volume; g, gram;

HOMA-IR, homeostatic model assessment for insulin resistance; hs-CRP, high sensitive C-reactive protein; IL-6, interleukin 6; l, litre; LA, left atrial; LV, left ventricular; m, meter; min, minute; mg, milligram; ml, millilitre; MRI, magnetic resonance imaging; PeAT, pericardial adipose tissue; RVEDV, right ventricular end-diastolic volume; RVEDVI, right ventricular end-diastolic volume index; SAX, short axis; SF-36, 36-item short form survey; TNF-∝, tumour necrosis factor alpha; VO_2, oxygen consumption; 4CV, 4-chamber view.

[a]Calculated from body surface area based on the Dubois method: $(8116.6 \times [0.007184 \times height\ (cm)^{0.725} \times weight\ (kg)^{0.425}]) - 28.2$

[b]Calculated by the Strauss formula: $100\ x\ \frac{haemoglobin\ (baseline)}{haemoglobin\ (180\ days)}\ x\ \frac{1-haematocrit\ (180\ days)}{1-haematocrit\ (baseline)} - 100$

Statistical analysis

Analyses of the primary and secondary endpoints have been described in the design paper:

*"The primary endpoints will be analyzed using a constrained linear mixed model including follow-up time, the constrained time*treatment interaction, age at baseline, sex, and site of randomization as fixed effects and with an unstructured covariance pattern to account for repeated measurements on each study participant. The estimated treatment differences for the primary endpoints in both parts of the trial program (trial I [Empire Prevent Cardiac]: peak VO_2 and LV mass; and trial II [Empire Prevent Metabolic]: EAT and eECV, respectively, after 180 days follow-up) will be reported with 97.5% confidence intervals and Bonferroni-adjusted P-values. An adjusted P-value ≤.05 (equivalent to an unadjusted P-value ≤.025, ...) will be considered statistically significant.*

The secondary endpoints will be analyzed using a constrained linear mixed model similar to that of the primary analysis. Missing data will be handled implicitly by maximum likelihood estimation in the linear mixed model. Estimated treatment differences will be reported with 95% confidence intervals, unadjusted P-values, and P-values adjusted for multiple testing using the method of Benjamini and Hochberg which controls the false discovery rate. P-values from the two trials will be adjusted separately. An adjusted P-value ≤.05 will be considered statistically significant." [1]

Analysis of both primary and secondary endpoints will be done for the *ITT population*. Supplementary sensitivity analyses will be provided using the *per-protocol population* (for further details, please see the "Per-protocol analyses" section below). Two main papers are planned for each trial, and correction for multiplicity will be performed collectively in turn across both main papers for all secondary endpoints within each trial. The outline for the results of each paper is presented in Appendix D.

Exploratory endpoints will be analyzed using a constrained linear mixed model similar to the analyses of primary and secondary endpoints as outlined above. Analyses will be performed using the *ITT population.* Results will be reported with 95% confidence intervals and unadjusted p-values. It will be noted in the results that since p-values are not adjusted for multiple testing spurious findings cannot be ruled-out and apparent findings will have to be confirmed by independent studies.

Missing data

The linear mixed model has been chosen for analysis due to its implicit handling of missing data, which is statistically optimal under a missing at random assumption. As described in the design paper: *"Numbers and reasons of drop-out and failed measurements will be tabulated. In case missing data are few (≤ 5%) and random, these will be handled implicitly by maximum likelihood estimation in the linear mixed model. In case drop-out is larger than expected or differential between treatment groups, sensitivity analyses based on multiple imputations will be performed."* [1]

As of January 2025, a total of N = 191 patients have been enrolled in the trial program, of which all were planned to complete an MRI. However, due to a fire at the MRI department at Herlev less patients than expected underwent MRI (N=165). Similarly, logistical concerns led to fewer patients than planned who completed the 99mTc DTPA-clearance (only N=75 of the N=92 included patients)

Although missing data occur more frequently than 5% for the primary endpoints, it is important to notice that data are missing due to external circumstances and not due to internal factors such as good or poor outcome for each patient. The linear mixed model handles data that are missing at random optimally. Therefore, sensitivity analyses using multiple imputation are not considered relevant and will not be performed.

Subgroup analyses

Test for differences in treatment effect between predefined subgroups will be performed by adding the time*subgroup and the treatment*time*subgroup interactions in the linear mixed models from the primary analyses together with a subgroup-specific covariance for the following subgroups:

"

1) *With and without impaired glucose tolerance.*
2) *With and without new-onset T2D.*

3) *With and without ischemic heart disease.*
4) *Who are biologically female versus biologically male.*
5) *With BMI <30 kg/m^2 and ≥30 kg/m^2.*
6) *With eGFR <90 mL/min/1.73m^2 and ≥90 mL/min/1.73m^2.”* [1]

As well as for the following additional subgroups:

7) Above and below median BMI.
8) Above and below median eGFR.

Subgroup-specific estimated treatment effects with 95% confidence intervals will be reported in forest plots. To control the false discovery rate, the Benjamini and Hochberg method will be used to adjust the p-values for the test of difference in treatment effect between the subgroups for multiple comparisons.

Per-protocol analyses

To estimate the effect of the treatment in ideal circumstances, the analyses of the primary and secondary endpoints will be repeated in the *per-protocol* population. Descriptive tables will be made to compare the baseline characteristics of participants who are in the *per-protocol population* to violators in each randomization group. In case any clinically relevant differences are found, the analyses will be repeated while adjusting for these as potential confounders.

Harms and safety endpoints

For safety endpoints, we will use the *safety population*. Data will be summarized for both the empagliflozin and placebo group, see Appendix E. For binary safety endpoints, we will apply an unconditional exact test, as appropriate. In case of low counts, as is expected, we will use Barnard’s exact test.

Statistical Software

Statistical analyses will be performed using R statistical software (R Foundation for Statistical Computing, Vienna, Austria) or other commercially available statistical packages. For R analyses, the LMMstar-package will be used for linear mixed model analyses.[2]

References

1. Andersen CF, Larsen JH, Jensen J, Omar M, Nouhravesh N, Kistorp C, Tuxen C, Gustafsson F, Knop FK, Forman JL, Davidovski FS, Jensen LT, Højlund K, Køber L, Antonsen L, Poulsen MK, Schou M, Møller JE. Empagliflozin to elderly and obese patients with increased risk of developing heart failure: Study protocol for the Empire Prevent trial program. *Am Heart J*. 2024;271:84–96.

2. Ozenne B, Forman J. LMMstar: Repeated measurement models for discrete times. R Stat. Softw. . 2024.

3. Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Doré C, Williamson PR, Altman DG, Montgomery A, Lim P, Berlin J, Senn S, Day S, Barbachano Y, Loder E. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. *JAMA*. 2017;318:2337–2343.

4. Nagueh SF, Smiseth OA, Appleton CP, Byrd BF 3rd, Dokainish H, Edvardsen T, Flachskampf FA, Gillebert TC, Klein AL, Lancellotti P, Marino P, Oh JK, Popescu BA, Waggoner AD. Recommendations for the Evaluation of Left Ventricular Diastolic Function by Echocardiography: An Update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *J Am Soc Echocardiogr Off Publ Am Soc Echocardiogr*. 2016;29:277–314.

References for non-standard statistical methods

None.

References to internal documents

Data management plan (DMP): DMP, version 2, 10.09.24.

Trial Master File (TMF) and Statistical Master File (SMF): TMF, version 1, 08.07.21, including SMF.

Standard Operating Procedures (SOPs) or documents: EMP14_1.0_SOP_eCRF, 08.07.21

Note: inspiration for this SAP came from the SAP Guidance Document: Recommended Items to Address in a Clinical Trial SAP[3]

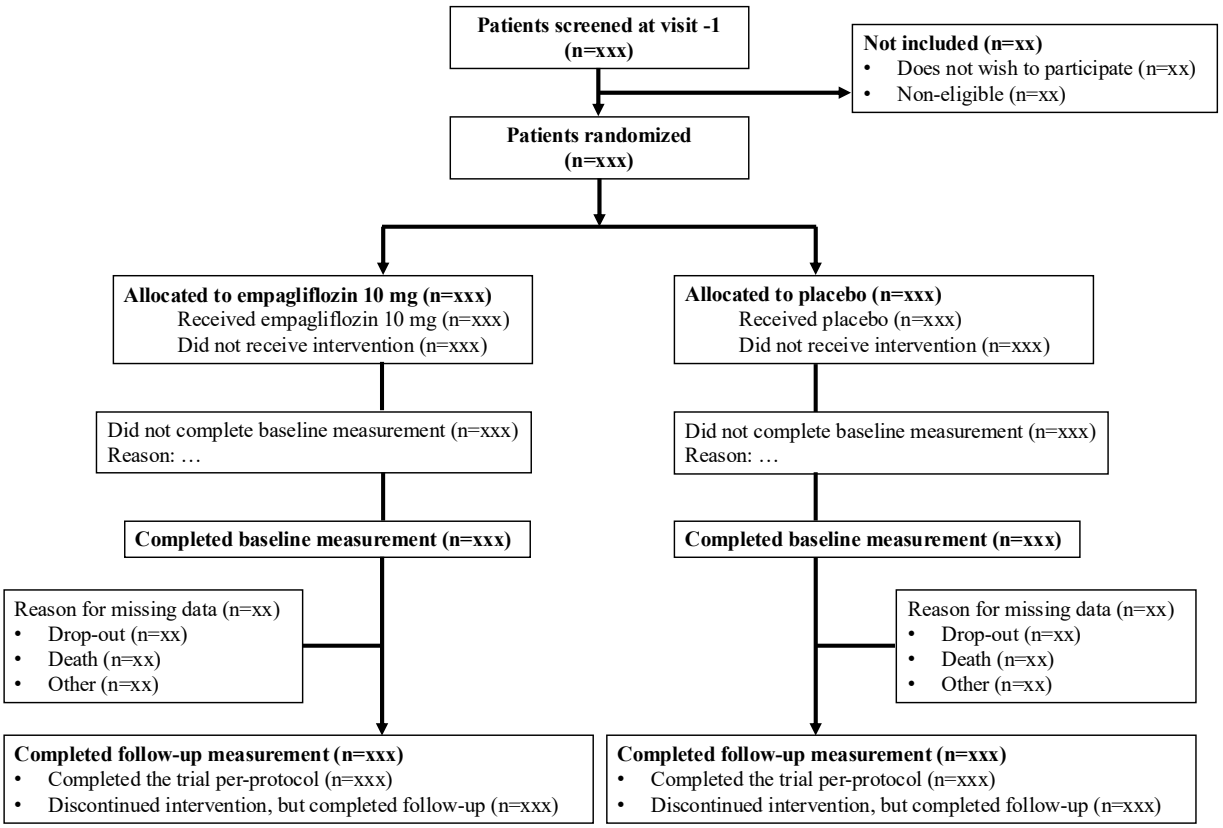
Appendix A. Study drug adherence and major and minor protocol deviations

Table A1. Adherence to treatment.	Empagliflozin (n=xx)	Placebo (n=xx)
Adherence rate (%), median (IQR)		
Patients with adherence rate > 80%, frequency (%)		

Abbreviations: IQR, interquartile range.

Table A2. Protocol deviations in the treatment period	Empagliflozin (n=xx)	Placebo (n=xx)
Major protocol deviations		
Collection of incorrect data		
Error in applying inclusion and exclusion criteria		
Minor protocol deviations		
Missed single examination		
Missed control visit		
Missed telephone visit		
Missed follow-up visit		
Other		

Appendix B. The Consolidated Standards of Reporting Trials (CONSORT) flow diagram, modified



Appendix C. Baseline characteristics

Table C1. Baseline characteristics	**Empagliflozin (n=xx)**	**Placebo (n=xx)**	**All patients (n=xx)**
Age (years), mean (SD)			
Male			
BMI (kg/m^2), median [Q1–Q3]			
Waist-to-hip ratio, median [Q1–Q3]			
Caucasian			
Risk factors for heart failure: Hypertension Ischemic heart disease Previous stroke/TCI Chronic kidney disease (eGFR 30-45 ml/min/1.73m^2)			
NYHA functional class: I II III			
Heart rate (bpm), median [Q1–Q3]			
Systolic blood pressure (mmHg), median [Q1–Q3]			
Baseline LVEF (%), median [Q1–Q3]			
Smoking status: Current smoker Previous smoker Non-smoker			
Metabolic syndrome[a]			
Diastolic dysfunction[b]			
Medical history			
Hypertension duration (months), median [Q1–Q3]			
Ischemic heart disease duration (months), median [Q1–Q3]			
Abnormal glucose tolerance: Impaired fasting glucose Impaired glucose tolerance			

Table C1. Baseline characteristics	**Empagliflozin (n=xx)**	**Placebo (n=xx)**	**All patients (n=xx)**
New onset type 2 diabetes			
Dyslipidemia			
Atrial fibrillation/flutter			
Chronic obstructive pulmonary disease			
Asthma			
Sleep apnoea			
Gout			
Anaemia[c]			
Biomarkers			
Pro-BNP (pmol/L), mean (SD)			
HbA1c, median [Q1–Q3]			
eGFR (ml/min/1.73m^2), median [Q1–Q3]			
Rate of eGFR: 60-89 ml/min/1.73m^2 < 60 ml/min/1.73m^2			
Haemoglobin (mmol/L), mean (SD)			
Pharmacological treatment			
RAASi[d]			
BB			
Calcium-receptor blockers			
Loop diuretic			
Lipid-lowering drugs			
Antithrombotic treatment: Acetylsalicylic acid ADP-receptor blockers New oral anticoagulant treatment Vitamin K antagonists			
Long-acting Nitrates			

Numbers are frequency (%) unless stated otherwise. [a]Using the criteria provided by the Consensus Statement from the International Diabetes Federation 2006: Waist circumference of ≥ 94cm (male) or ≥ 80cm (female) and any two of the

following: 1) elevated triglycerides (≥ 1.7 mmol/l) or specific medical treatment for this lipid abnormality, 2) reduced HDL-cholesterol of <1.03 mmol/l (male) or <1.29 mmol/l (female) or specific medical treatment for this lipid abnormality, 3) elevated blood pressure of ≥ 130 mmHg (systolic) or ≥ 85 mmHg (diastolic) or medically treated hypertension, 4) elevated fasting plasma glucose ≥ 5.6 mmol/l (previously diagnosed T2D is not relevant for this population as this is an exclusion criterion). Patients taking either fibrates or nicotinic acid were presumed to have elevated triglycerides and reduced HDL-cholesterol. [b]Defined as >50% positive of the following criteria: 1) average E/e´>14; 2) septal e´ velocity <7 cm/s OR lateral e´ velocity <10 cm/s; 3) TR velocity >2.8 m/s; 4) LA volume index >34 ml/m^2.[4] [c]Defined in accordance with the World Health Organization criteria as haemoglobin ≤8.1 mmol/l (13 g/dl) for males and ≤7.4 mmol/l (12 g/dl) for females. [d]Either angiotensin-converting enzyme inhibitor, angiotensin receptor blocker, or sacubitril-valsartan.

Abbreviations: BMI, body mass index; bpm, beats per minute; CRT, cardiac resynchronization therapy; eGFR, estimated glomerular filtration rate; HbA1c, glycated haemoglobin; HF, heart failure; ICD, implantable cardioverter-defibrillator; Q1–Q3, first quartile–third quartile; LVEF, left ventricular ejection fraction; MRA, mineralocorticoid receptor antagonist; NT-proBNP, N-terminal pro-b-type natriuretic peptide; NYHA, New York Heart Association; pro-BNP, pro-brain natriuretic peptide; RAASi, renin-angiotensin-aldosterone system inhibition; SD, standard deviation.

Appendix D. Results tables

Table D1. Results from Empire Prevent Cardiac Paper I	Empagliflozin (n=xx)		Placebo (n=xx)		Estimated Treatment	P-values[c]	
	Baseline (mean (SD))	Δ-value[a] (mean (SD))	Baseline (mean (SD))	Δ-value[a] (mean (SD))	Difference, ETD[b] (CI)	P_{raw}	P_{adj}
Primary endpoint:							
LV mass index, g/m^2	xx (xx)	xx (xx)	xx (xx)	xx (xx)	xx (97.5% CI: xx to xx)		
Secondary endpoints:							
LVEDVI, ml/m^2	xx (xx)	xx (xx)	xx (xx)	xx (xx)	xx (95% CI: xx to xx)		
LVESVI, ml/m^2	xx (xx)	xx (xx)	xx (xx)	xx (xx)	xx (95% CI: xx to xx)		
LVEF, %	xx (xx)	xx (xx)	xx (xx)	xx (xx)	xx (95% CI: xx to xx)		

Table D2. Results from Empire Prevent Cardiac Paper II	Empagliflozin (n=xx)		Placebo (n=xx)		Estimated Treatment	P-values[c]	
	Baseline (mean (SD))	Δ-value[a] (mean (SD))	Baseline (mean (SD))	Δ-value[a] (mean (SD))	Difference, ETD[b] (CI)	P_{raw}	P_{adj}
Primary endpoint:							
Peak VO_2, ml/min/kg	xx (xx)	xx (xx)	xx (xx)	xx (xx)	xx (97.5% CI: xx to xx)		
Secondary endpoint:							
Accelerometer counts, cpm	xx (xx)	xx (xx)	xx (xx)	xx (xx)	xx (95% CI: xx to xx)		

[a]Δ-value: change from baseline to follow-up. [b]ETD: estimated treatment difference from baseline to follow-up based on a constrained linear mixed model with inherent baseline adjustment and adjustment for age, sex and randomization site as baseline covariates. [c]The P-values for the primary endpoints have been adjusted using Bonferroni correction with split alpha, and the P-values for the secondary endpoints have been adjusted using ad modum Benjamini and Hochberg to control the false discovery rate.

Abbreviations: CI, confidence interval; cpm, counts per minute; g, gram; LV, left ventricular; LVEDVI, left ventricular end-diastolic volume index; LVESVI, left ventricular end-systolic volume index; m, meter; min; minute; ml, milliliter; P_{adj}, P-value adjusted for multiple testing;; P_{raw}, unadjusted P-value; SD, standard deviation; VO_2, oxygen consumption.

Table D3. Results from Empire Prevent Metabolic Paper I	**Empagliflozin (n=xx)**		**Placebo (n=xx)**		**Estimated Treatment Difference, ETD[b] (CI)**	**P-values[c]**	
	Baseline (mean (SD))	**Δ-value[a] (mean (SD))**	**Baseline (mean (SD))**	**Δ-value[a] (mean (SD))**		**P_{raw}**	P_{adj}
Primary endpoint:							
Ventricular EAT mass, g	xx (xx)	xx (xx)	xx (xx)	xx (xx)	xx (97.5% CI: xx to xx)		
Secondary endpoints:							
PeAT volume, cm^2	xx (xx)	xx (xx)	xx (xx)	xx (xx)	xx (95% CI: xx to xx)		
Total EAT volume, cm^2	xx (xx)	xx (xx)	xx (xx)	xx (xx)	xx (95% CI: xx to xx)		

Table D4. Results from Empire Prevent Metabolic Paper II	**Empagliflozin (n=xx)**		**Placebo (n=xx)**		**Estimated Treatment Difference, ETD[b] (CI)**	**P-values[c]**	
	Baseline (mean (SD))	**Δ-value[a] (mean (SD))**	**Baseline (mean (SD))**	**Δ-value[a] (mean (SD))**		**P_{raw}**	P_{adj}
Primary endpoint:							
eECV, L	xx (xx)	xx (xx)	xx (xx)	xx (xx)	xx (97.5% CI: xx to xx)		
Secondary endpoints:							
ePV, L	xx (xx)	xx (xx)	xx (xx)	xx (xx)	xx (95% CI: xx to xx)		
RVEDVI, ml/m^2	xx (xx)	xx (xx)	xx (xx)	xx (xx)	xx (95% CI: xx to xx)		

[a]Δ-value: change from baseline to follow-up. [b]ETD: estimated treatment difference from baseline to follow-up based on a constrained linear mixed model with inherent baseline adjustment and adjustment for age, sex and randomization site as baseline covariates. [c]The P-values for the primary endpoints have been adjusted using Bonferroni correction with split alpha, and the P-values for the secondary endpoints have been adjusted using ad modum Benjamini and Hochberg to control the false discovery rate.

Abbreviations: CI, confidence interval; EAT, epicardial adipose tissue; eECV, estimated extracellular volume; ePV, estimated plasma volume; g, gram; L, liter; m, meter; ml, milliliter; P_{adj}, P-value adjusted for multiple testing; PeAT, pericardial adipose tissue; P_{raw}, unadjusted P-value; RVEDVI, right ventricular end-diastolic volume index; SD, standard deviation.

Appendix E. Adverse events of special interest, hospital admissions and deaths

Table E1. Adverse events in the treatment period	Empagliflozin (n=xx)	Placebo (n=xx)	P-value
Discontinuation of the drug			
Urinary tract infection Non-severe Requiring hospitalization			
Genital infection Non-severe Requiring hospitalization			
Hypoglycemia[a]			
Volume depletion[b]			
Acute renal failure			
Ketoacidosis			
Thromboembolic events			
Bone fracture			
Amputations			
Fournier's gangrene			
Hospitalizations Caused by adverse reactions Caused by cardiac events Other causes			
Deaths			

Numbers are frequency (%) unless stated otherwise. [a]Defined as an event where the patient requires assistance from others to actively administer carbohydrates, glucagon, or other corrective action. Neurological recovery following the corrective action(s) is considered sufficient evidence that the event was induced by low plasma glucose – even if the plasma glucose concentration is unavailable during the event. [b]Defined as dehydration, hypovolemia, or hypotension requiring contact with the health care system.