Statistical analysis plan (SAP)

EMAIL-HF

Protocol title: Implementation of SGLT2 inhibitors in patients with heart failure through a new digital strategy (EMAIL-HF): a registry-based randomized clinical trial

ClinicalTrials.gov ID: NCT05890131. Registered 3 May 2023.

SAP version: 2.0 (November 13, 2025)

Responsible for writing the SAP: Mariam Elmegaard

Statistician: Theis Lange, Department of Public Health, University of Copenhagen, Copenhagen, Denmark.
Date and signature: Theis Lange 05-FEB-2026

Principal investigator: Morten Schou, Department of Cardiology, Herlev-Gentofte Hospital, Copenhagen, Denmark.
Date and signature: 05-feb-2026

Project lead: Mariam Elmegaard, Department of Cardiology, Herlev-Gentofte Hospital, Copenhagen, Denmark.
Date and signature: Mariam Elmegaard 05-FEB-2026

Table of Contents

1. Abbreviations

- SAP: Statistical analysis plan
- HF: Heart failure
- SGLT2: Sodium glucose transporter 2
- LVEF: Left ventricle ejection fraction
- ITT: Intention-to-treat
- CONSORT: Consolidated standards of reporting trials
- ATC: Anatomical Therapeutic Chemical Classification system.
- ICD-10: International Classification of Diseases, 10th revision.

2. Introduction

This is the statistical analysis plan (SAP) for EMAIL-HF which is a registry-based, randomized, parallel group, clinical trial including patients with a registered diagnosis of heart failure (HF), evaluating the effect of a digital strategy to accelerate the implementation of sodium-glucose co-transporter 2 (SGLT2) inhibitors (empagliflozin and dapagliflozin) in patients with HF, in addition to standard care. The SAP is an addendum to the design paper of the EMAIL-HF trial(1).

3. Methods

3. 1 Primary objective

The primary objective of this trial is to assess whether a digital strategy, including an informational letter detailing SGLT2 inhibitors as a new treatment option for HF and an invitation for clinical evaluation, can increase the implementation of SGLT2 inhibitors in patients with chronic HF.

3. 2 Secondary objective

The secondary objectives include evaluating the impact of the intervention on relevant clinical outcomes, including hospitalization for HF and non-fatal myocardial infarction, non-fatal stroke, renal failure, and all-cause mortality. Additionally, the study will assess adherence to SGLT2 inhibitor therapy and the proportion of certain disadvantaged populations initiating treatment. These

populations include elderly, individuals with severe comorbidities, non-western immigrants, those of lower socioeconomic or educational levels, residents of rural areas, and patients with mental illness. Also, a key secondary objective is to gain insights into digital inequality in healthcare, understanding how socioeconomic and demographic factors may influence access to and engagement with digital health interventions. Finally, the study will examine the health economic consequences of the digital intervention by comparing annual healthcare utilization between the intervention and control groups.

3.3 Primary endpoint

The primary outcome is the proportion of patients in the two arms initiating therapy with dapagliflozin or empagliflozin, defined as ≥1 redeemed prescription from randomization up to 6 months. See Table 1.

3.4 Secondary endpoint

The secondary outcome is the composite of HF hospitalization and all-cause mortality. See Table 1.

3.5 Exploratory endpoints

See Table 1.

4. Population

4.1 Inclusion and exclusion criteria

All patients will be identified from Danish nationwide health registries, with the following criteria:

Inclusion criteria:

- Registered diagnosis of HF within the last 10 years, irrespective of LVEF
- Adult (aged 20 years or older)
- Living in the Capital Region of Denmark or Roskilde

Exclusion criteria:

- Exemption from the Danish mandatory governmental digital letter system (Digital Post)
- Redeemed prescription of a SGLT2 inhibitor after 2015

- History of ketoacidosis
- Type 1 diabetes mellitus
- Dementia
- Chronic kidney disease in long-term dialysis
- New-onset cancer within the last year (except prostate and non-melanoma skin cancer)
- Residing in a nursing home

For a detailed definition of the inclusion and exclusion criteria, see Table 2.

4.2 Study design and randomization

EMAIL-HF is a registry-based, randomized, controlled, parallel group, investigator-blinded clinical trial including all patients with a registered diagnosis of HF, living in the Capital Region of Denmark and Roskilde, evaluating the effect of a digital strategy to increase and accelerate the implementation of SGLT2 inhibitors (e.g. empagliflozin and dapagliflozin) in eligible patients with chronic HF, in addition to standard care. Patients are identified from Danish nationwide health registries(2) and randomized 1:1 to receive a letter via the Danish mandatory governmental digital letter system (Digital Post)(3), with information on SGLT2 inhibitors as a new treatment option for HF and an invitation for medical evaluation. Depending on their residence, patients are allocated to one of six experimental sites.

Randomization was performed as simple randomization without stratification. To support operational feasibility at the participating HF clinics, randomization was conducted in monthly batches of approximately 200–1000 patients. The randomization sequence was generated using R/RStudio on a secure database server at Copenhagen University Hospital – Herlev and Gentofte. Patients allocated to the intervention arm were contacted via Digital Post within 30 days of registry identification.

4.3 End of study

The primary endpoint (initiation of SGLT2 inhibitor therapy) will be assessed over the first 6 months after randomization. Secondary and exploratory endpoints will be assessed during longer follow-up as prespecified in Table 1 (up to 24 months after randomization). Following completion of follow-up and extraction of registry data for all endpoints, data will be obtained from the Danish Health Data

Authority (Sundhedsdatastyrelsen) and the database will be locked for final analyses. Database lock is anticipated in February 2026**.**

5. Statistical Principles

5.1 Analysis populations

Intention-to-treat (ITT) population: The ITT population will include all randomized patients, analysed according to the group they were randomized to (digital letter strategy vs. usual care), regardless of subsequent response to the invitation, later eligibility assessment or initiation or discontinuation of SGLT2 inhibitor therapy.

'Per-protocol' population: The 'per-protocol' population in this registry-based trial setting, will include patients randomized to the intervention arm who (i) were successfully delivered the digital letter (Digital Post), (ii) attended a clinical evaluation at the participating heart failure clinic within the prespecified assessment window, and (iii) were confirmed eligible for SGLT2 inhibitor initiation at the evaluation, and (iv) subsequently redeemed a presciption of an SGLT2 inhibitor. Patients randomized to the control arm will be included provided they remained resident and observable in registries during the corresponding follow-up window.

Safety population: The safety population will include all randomized patients who initiate an SGLT2 inhibitor during follow-up (defined by first redeemed prescription for an SGLT2 inhibitor in the national prescription registry). Safety outcomes will be assessed using routinely collected registry data and will focus on adverse events of special interest and commonly recognised adverse reactions associated with SGLT2 inhibitors, including (as available in registries) diabetic ketoacidosis, severe hypoglycaemia, genital mycotic infections, urinary tract infections, volume depletion, acute kidney injury, and lower-limb amputation (only hospitalizations). Safety analyses will be descriptive and performed according to randomized group.

5.2 Eligibility

Eligibility is determined using prespecified inclusion and exclusion criteria based on nationwide health registries, as detailed in the study protocol and the design paper. Any randomized patients subsequently found to be ineligible will remain in the ITT population.

5.3 Withdrawal and follow-up

As this is a registry-based trial, complete follow-up for outcomes is expected through linkage to nationwide registries. Participants in the intervention arm may decline the digital invitation and/or any clinical evaluation; such non-participation will be recorded where available and does not constitute withdrawal from follow-up for registry-based endpoints. If a participant requests withdrawal, the level of withdrawal will be tabulated and classified as: (1) withdrawal from the intervention only (no further contact), or (2) withdrawal from intervention and from further follow-up and data use, to the extent permitted by applicable regulations and data access approvals.

The timing and number of (i) withdrawals, (ii) emigration, (iii) deaths prior to completion of follow-up, and (iv) any exclusions from analysis (if applicable) will be presented in a modified CONSORT flow diagram, with numbers and reasons provided at each stage.

5.4 Baseline characteristics

Baseline characteristics will be described for the ITT population overall and by randomized group (digital letter strategy vs. usual care). Characteristics will be obtained from nationwide health registries and will include demographic variables, HF–related characteristics, relevant comorbidities, and baseline medication use. Baseline characteristics will be summarized descriptively in a table (see Table 3). Continuous variables will be presented as mean (standard deviation) or median (interquartile range), and categorical variables as counts and percentages. Formal statistical comparisons of baseline characteristics will not be performed, in accordance with CONSORT recommendations(4).

For the intervention population (digital strategy) who respond to the letter and give consent, additional clinical variables available from routine clinical documentation in their medical journals will also be reported, including LVEF, NYHA class, BMI, alcohol use, smoking status. These variables will be presented descriptively (see Table 4).

5.5 Endpoints

All prespecified endpoints will be reported in accordance with the CONSORT recommendations(4). For each randomized group, descriptive endpoint data will be presented, together with estimated between-group effects and 95% confidence intervals. Statistical testing will follow the prespecified

multiplicity strategy (hierarchical testing for primary and secondary endpoints; exploratory endpoints will be reported without formal multiplicity adjustment unless otherwise specified), with corresponding p-values presented in the results tables (see Table 5-6)

5.6 Statistical analysis

The statistical analyses are prespecified and aligned with the design paper(1). The primary endpoint is initiation of SGLT2 inhibitor therapy, defined as redemption of at least one prescription for an SGLT2 inhibitor within the prespecified follow-up window after randomization (e.g., 6 months), ascertained from the Danish national prescription registry(5). The effect of the digital letter strategy versus usual care on the primary endpoint will be analysed using logistic regression with randomized group as the main explanatory variable. To avoid classifying death as 'non-initiation', the primary logistic regression analysis will be performed among patients who are alive and resident/observable in Danish registries at the end of the uptake window (i.e., not emigrated). Results will be reported as between-group differences in initiation with 95% confidence intervals and two-sided p-values. As a supportive analysis addressing the competing risk of death, cumulative incidence of initiation by the end of the uptake window will be estimated by randomized group with death treated as a competing event and emigration treated as censoring.

The secondary endpoint, time to first HF hospitalization or all-cause mortality, will be analysed in the ITT population using a Cox proportional hazards model. Results will be presented as hazard ratios with 95% confidence intervals and two-sided p-values. Multiplicity will be handled using a hierarchical testing strategy: the primary endpoint will be tested first at a two-sided $\alpha=0.05$. If the primary endpoint is statistically significant, the secondary endpoint will subsequently be tested at a two-sided $\alpha=0.05$. Otherwise, inference for the secondary endpoint will be considered exploratory.

Missing data are expected to be limited due to the registry-based design. If missingness in key baseline covariates required for adjusted or subgroup analyses exceeds 5%, multiple imputation will be applied using an imputation model including randomized group, key baseline characteristics, and outcome information as appropriate.

All primary and secondary analyses will be conducted in the ITT population. Supplementary analyses will be performed using a 'per-protocol' supportive approach, as described in Section 5.8. Exploratory endpoints will be analysed using methods appropriate to endpoint type (e.g., logistic regression for binary outcomes, Cox models for time-to-event outcomes, and regression models for count

outcomes). Exploratory analyses will be reported with effect estimates and 95% confidence intervals and unadjusted p-values.
All analyses will be performed using R (version 4.0 or later) and SAS statistical software.

5.7 Subgroup analyses

Potential heterogeneity of the intervention effect will be explored across predefined baseline subgroups using interaction terms, as prespecified in the design paper(1). Subgroup analyses will be considered exploratory. Results will be presented in forest plots showing subgroup-specific effect estimates and confidence intervals, alongside interaction P-values.

Predefined subgroups of interest are:

- Age (≥75 vs. <75 years)
- Sex (female vs. male)
- Type 2 diabetes (yes vs. no)
- Ischaemic heart disease (yes vs. no)
- Chronic kidney disease / renal function (eGFR<60 ml/min and/or CKD diagnosis where applicable)
- Atrial fibrillation (yes vs. no)
- Vulnerable populations / digital inequality analyses, including: elderly (≥75 years), severe comorbidity burden, non-Western immigrants, lower socioeconomic/educational level, rural residence, and mental illness.

5.8 Supportive analysis

As a supportive analysis, we will evaluate whether the association between SGLT2 inhibitor initiation and the secondary outcome (time to first HF hospitalization or all-cause mortality) differs depending on whether initiation occurred in the context of the digital strategy versus usual care. Initiation will be defined as the first redeemed prescription for an SGLT2 inhibitor and treated as a post-randomization exposure. The secondary outcome will be analysed using Cox regression. Results will be reported as hazard ratios with 95% confidence intervals.

5.9 Health economic analyses

Health economic outcomes will be evaluated descriptively by comparing healthcare utilization and related direct healthcare costs between the digital strategy and usual care groups using routinely

collected registry data. Analyses will include key components such as inpatient admissions (including length of stay), outpatient and primary care contacts, and prescription drug use/costs (including SGLT2 inhibitors). The intervention cost will be estimated pragmatically based on staff time and administrative resources. Results will be reported as mean costs/utilization per patient and between-group differences with uncertainty measures, and interpreted as supportive evidence on the economic impact of the digital implementation strategy.

5.10 Missing data

Due to the registry-based design, missing outcome data are expected to be minimal for the primary endpoint (redeemed prescription for an SGLT2 inhibitor) and for time-to-event outcomes, as these are ascertained through nationwide administrative registries with near-complete coverage. Accordingly, the primary and secondary endpoint analyses will be performed without imputation of outcome data.

For time-to-event endpoints, incomplete follow-up will be handled through censoring at the date of emigration or end of the follow-up period. Deaths will be captured through registries and incorporated in the composite endpoint definition as prespecified.

Missing data may occur for selected baseline covariates and for additional clinical variables available only in the per-protocol population (e.g., LVEF, BMI, blood pressure, laboratory values recorded in routine care). The extent and patterns of missingness will be described, and the number of missing values will be reported for each variable.

If missing data in variables used for adjusted or subgroup analyses are few (≤5%), a complete-case approach will be used. If missingness exceeds 5% for key covariates (or appears imbalanced between randomized groups), sensitivity analyses using multiple imputation under a missing-at-random assumption will be performed. Results from imputed analyses will be compared with complete-case analyses to assess robustness.

5.11 Statistical software

All statistical analyses will be performed using R (R Foundation for Statistical Computing, Vienna, Austria; version 4.0 or later) and/or SAS (SAS Institute Inc., Cary, NC, USA). Data management and analyses will be conducted in a secure computing environment in accordance with applicable data protection and governance requirements.

6. Ethical considerations

EMAIL-HF is a pragmatic, registry-based randomized implementation trial conducted in accordance with applicable ethical principles and Danish regulations. The trial has been evaluated and approved by The Ethical Committee of the Capital Region of Denmark (project ID: F-22035916), the Danish Health Data Authority (project IDs: FSEID-00006340 and FSEID-00006476), and the Capital Region of Denmark (project ID: P-2022-675).

7. References

1. Elmegaard M, Køber L, Ersbøll MK, Lange T, Tuxen CD, Mouridsen MR, et al. Digital implementation strategy to increase SGLT2 inhibitor uptake in heart failure: Study design of EMAIL-HF. ESC Heart Fail. 2025;12(6):3953-65.

2. Schmidt M, Schmidt SA, Sandegaard JL, Ehrenstein V, Pedersen L, Sørensen HT. The Danish National Patient Registry: a review of content, data quality, and research potential. Clin Epidemiol. 2015;7:449-90.

3. Agency for Digital Government D. Digital Post 2025 [Available from: https://lifeindenmark.borger.dk/apps-and-digital-services/Digital-Post.

4. Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. Int J Surg. 2011;9(8):672-7.

5. Pottegård A, Schmidt SAJ, Wallach-Kildemoes H, Sørensen HT, Hallas J, Schmidt M. Data Resource Profile: The Danish National Prescription Registry. Int J Epidemiol. 2017;46(3):798-f.

8. Supplemental

Figure 1. CONSORT diagram for EMAIL-HF

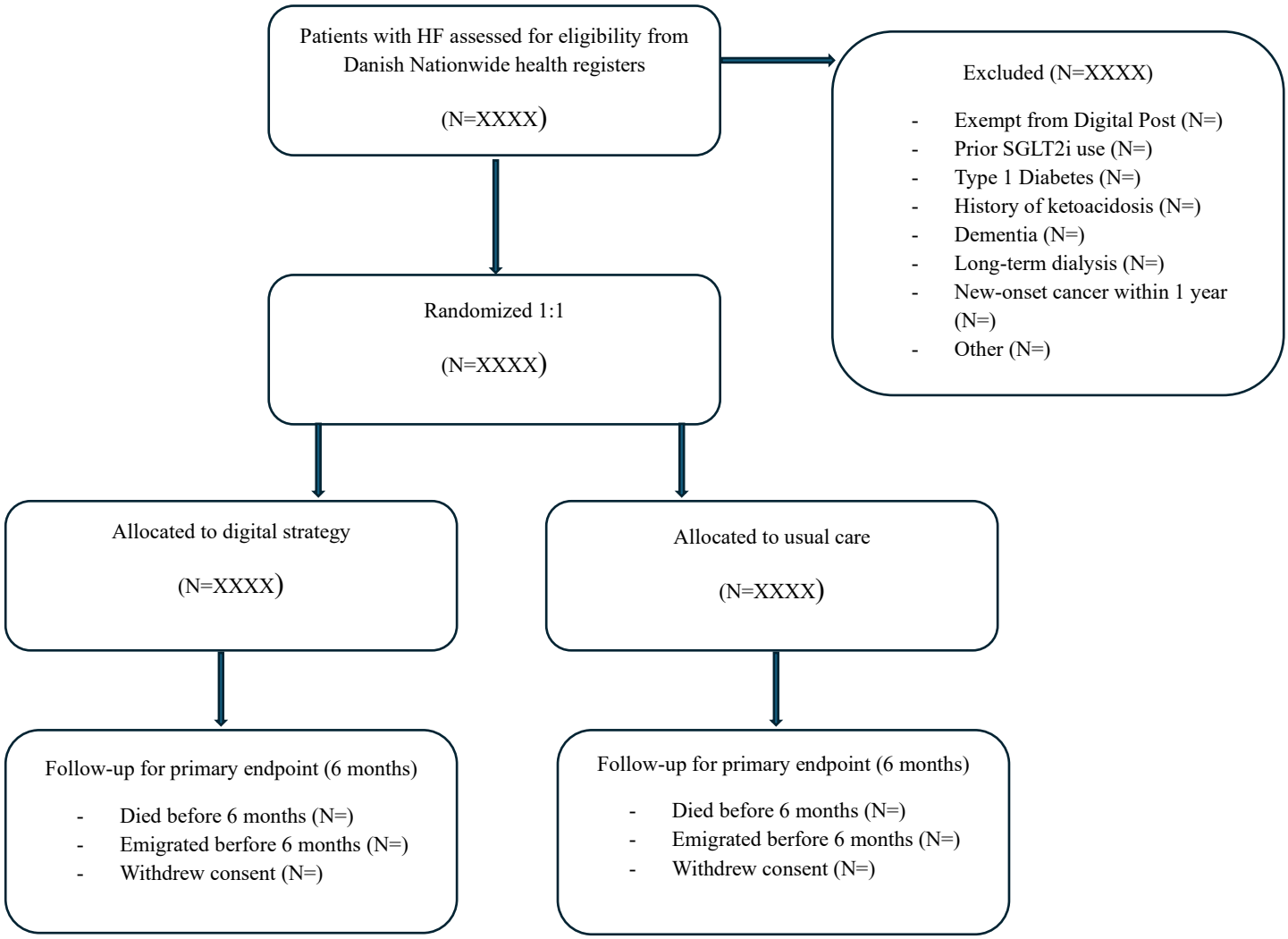

Table 1. Pre-specified definitions of the endpoints and subgroups

	Description	**ICD-10/ATC-codes**	**Type**	**Timeframe**
Primary outcome	Proportion of patients initiating therapy with dapagliflozin or empagliflozin	ATC: A10BK01, A10BK03	≥1 redeemed prescription	From randomization up to six months
Secondary outcome	Time to first occurrence of a composite heart failure endpoint consisting of all-cause death or heart failure hospitalization.	ICD-code: I50*	Type A diagnosis, unplanned urgent visit or in-patient overnight stay	From randomization up to 24 months
Other outcome measures	Number of all-cause deaths and heart failure hospitalizations	ICD-code: I50*	Type A diagnosis, unplanned urgent visit or in-patient overnight stay	From randomization up to 24 months
	Time to first occurrence of a 3-point expanded composite heart failure endpoint consisting of all-cause death, heart failure hospitalization, or kidney failure, with examination of the components of this composite	Heart failure: ICD-code: I50*. Kidney failure: DZ992*, N185, BJFD2*, or decline in eGFR ≥50%	Type A diagnosis, unplanned urgent visit or in-patient overnight stay	From randomization up to 24 months
	Time to first occurrence of a 4-point expanded composite heart failure endpoint consisting of all-cause death, heart failure hospitalisation, non-fatal myocardial infarction, non-fatal stroke, with examination of the components of this composite.	Heart failure: ICD-code: I50*. Myocardial infarction: ICD-code: I21 Stroke: ICD-code: I63-I64	Type A diagnosis, unplanned urgent visit or in-patient overnight stay	From randomization up to 24 months
	Time to initiation of SGLT2 inhibitor therapy	ATC: A10BK01, A10BK03	≥1 redeemed prescription	From randomization up to 24 months

	Adherence to SGLT2 inhibitor therapy	ATC: A10BK01, A10BK03	Proportion of days covered ≥80%	From initiation of therapy up to 12 months
Subgroups	Elderly >75 years			
	Patients with lower educational level		Defined according to the International Standard Classification of Education (ISCED)	
	Patients with lower socioeconomic level		Defined according to the Organization for Economic Co-operation and Development (OECD) modified scale	
	Immigrants from non-Western countries			
	Patients with type 2 diabetes	ATC-code: A10B ICD-code: DE11		
	Male and female patients (biological sex)			
	Patients with cardiovascular disease (including atrial fibrillation, cerebrovascular disease, peripheral artery disease, ischemic heart disease)	ICD-10 code: I20-I25, I50, I42, I110, I130, I132, J819, I70, I74, I48, I60-I64, DI1	Type A/B diagnosis	
	Patients with mental illnesses	ICD-10 code:	Type A/B diagnosis	
	Patients with chronic kidney disease	ICD-10 code: N02-N08, N11, N12, N14, N18, N19, N26, E102, E112, E132, E142, I120, N168, M300,	Type A/B diagnosis	

		M313, M319, Q612, Q613, Q615, Q619, T858, T859, Z992, or eGFR< 60ml/min		
	Patients living in rural/urban areas		According to the European Commission's definition of the degree of urbanization (DEGURBA)	
	Patients with low versus high comorbidity burden		Low comorbidity burden= two or less comorbidities. High comorbidity burden=more than two comorbidities.	
	Patients with and without polypharmacy		Polypharmacy= five or more co-medications	

Table 2. Inclusion and exclusion definitions

Inclusion criteria	Definition	Exclusion criteria	Definition
Registered diagnosis of heart failure within the last 10 years, irrespective of LVEF	ICD-10 code: I50* (excluding I501A, I501B, I501C)	Exemption from the Danish mandatory governmental digital letter system ('Digital Post')	-
Aged 20 years or older	-	Redeemed prescription of a SGLT2 inhibitor after 2015	ATC code: A10BK*, A10BD15, A10BD16, A10BD19, A10BD20, A10BD21, A10BD25, A10BD27
Living in the Capital Region of Denmark or municipality of Roskilde	Municipality code: 101, 147, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 173, 175, 183, 185, 187, 190, 201, 210, 217, 219, 223, 230, 240, 250, 260, 265, 270, 400, 411	History of ketoacidosis	ICD-10 code: E141, E131, E111
		Type 1 diabetes mellitus	ICD-10 code: E10*
		Dementia	ICD-10 code: F0*
		Chronic kidney disease in long-term dialysis	ICD-10 code: DZ992*, N185, BJFD2*
		New-onset cancer within the last year (except prostate and non-melanoma skin cancer)	ICD-10 code: C00*-C43*, C45*-C60*, C62*-C99*
		Residing in a nursing home	-

** Includes all ICD-10 codes in the same category.*
ATC: Anatomical Therapeutic Chemical Classification system.
ICD-10: International Classification of Diseases, 10th revision.

Table 3. Baseline characteristics

	Digital strategy (n=XXXX)	**Usual care (n=XXXX)**	**All patients (n=XXXX)**
Age, years — mean (SD)			
Female sex — n (%)			
Male sex — n (%)			
Urban residence – n (%)			
Rural residence – n (%)			
Caucasian – n (%)			
Immigrant – n (%)			
Medical history and comorbidity			
Heart failure duration, months – mean (SD)			
History of HF hospitalization – n (%)			
Ischaemic heart disease — n (%)			
PCI – n (%)			
CABG – n (%)			
Atrial fibrillation/flutter — n (%)			
Prior stroke/TIA — n (%)			
Type 2 diabetes — n (%)			
Chronic kidney disease — n (%)			
COPD/asthma — n (%)			
Sleep apnea – n (%)			
Hypertension — n (%)			
Cancer — n (%)			

Anaemia – n (%)			
Mental illness — n (%)			
Laboratory measures			
eGFR, mL/min/1.73m² — median [Q1–Q3]			
NT-proBNP — median [Q1–Q3]			
Baseline pharmacotherapy			
Diuretic – n (%)			
ACEi— n (%)			
ARB – n (%)			
ACEi or ARB – n (%)			
Beta-blocker — n (%)			
Mineralocorticoid receptor antagonist — n (%)			
Cardiac glycoside – n (%)			
Ivabradine – n (%)			
Statin/lipid-lowering therapy — n (%)			
Antiplatelet therapy — n (%)			
Oral anticoagulant therapy — n (%)			
Glucose-lowering therapy (any) — n (%)			

Table 4. Supplemental baseline characteristics

	Digital strategy – responders only (n=XXXX)
NYHA class – n (%)	
I	
II	
III	
IV	
LVEF – n (%)	
>=50%	
41-49%	
<=40%	
History of LVEF <40	
BMI, kg/m^2 – median [Q1–Q3]	
Smoking – n (%)	
Never	
Former smoker	
Light-moderate-intermittent smoker	
Heavy smoker (>14 cigarettes/day)	
Alcohol misuse – n (%)	

Table 5. Primary outcome (ITT): Initiation of SGLT2i

	Digital strategy (n=)	**Usual care (n=)**	**Effect estimate**	**95% CI**	**P-value**
Initiation of SGLT2i at 6 months, n/N (%) †	xx/xx (xx.x)	xx/xx (xx.x)	Risk difference (percentage points)*	(xx.x to xx.x)	x.xxx
Supportive analysis					
Cumulative incidence of initiation by 6 months (death as competing risk), %	xx.x	xx.x	Absolute difference at 6 months	(xx.x to xx.x)	x.xxx

Footnote:

† Primary analysis excludes patients who die or emigrate before 6 months (i.e., includes those alive and resident/observable at 6 months).

* Risk difference estimated from model-based predicted probabilities (marginal standardization).

Table 6. Secondary outcome (ITT): Time to first heart failure hospitalization or all-cause mortality

	Digital strategy (n=)	**Usual care (n=)**	**Effect estimate**	**95% CI**	**P-value**
Time to first HF hospitalization or all-cause mortality (events, n)	xxxx	xxxx	Hazard ratio (cox)	(xx.x to xx.x)	x.xxx

Footnote:

Cox proportional hazards model with randomized group as main explanatory variable; participants are censored at emigration or administrative end of follow-up.