University of Stavanger

University of Stavanger

STATISTICAL ANALYSIS PLAN for eHealth@H-2-H

NCT05750953

Anne Marie Lunde Husebø, principal investigator
25.08.2026

STATISTICAL ANALYSIS PLAN for eHealth@H-2-H

This Statistical Analysis Plan (SAP) adheres to scientific guidelines of the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH), i.e., ICH E9 "Statistical principles for clinical trials" including the addendum ICH E9(R1) on estimands and sensitivity analysis (ICH official website, 2025).

Administrative information:

Sponsor name	University of Stavanger
Sponsor address	Kjell Arholms gt. 41, 4021 Stavanger, Norway
EU CT number / REC no	REC ID 556114
Trial title	Nurse-assisted Intervention "eHealth@Hospital-2-Home" (Ehealth@H2H")
Trial ID	eHealth@Hospital-2-Home, Norwegian Research Council funding number 301472
Trial registration number	ClinicalTrial.gov ID NCT05750953

SAP and protocol version:

SAP version and date:	This SAP is version 1.0, dated 25 August 2025
Protocol version	This document has been written based on information contained in the study protocol version 9.0, dated 15 August 2025.

SAP revision history:

Protocol version	SAP version	Section number changed	Description and reason for change	Date changed
9.0	1.0	-	Original version	-

SIGNATURE PAGE

PRINCIPAL/COORDINATING INVESTIGATOR:

Name, title: Anne Marie Lunde Husebø, professor
Affiliation: Department of Public Health, Faculty of Health Sciences, University of Stavanger, Norway; Department of Research, Stavanger University Hospital, Norway

28.08.2025

Signature | Date (dd.mm.yyyy)

TRIAL STATISTICIAN:

Name, title: Ingvild Dalen, PhD
Affiliation: Section of Biostatistics, Department of Research, Stavanger University Hospital, Norway; Department of Quality and Health Technology, Faculty of Health Sciences, University of Stavanger, Norway

25.08.2025

Signature | Date (dd.mm.yyyy)

ABBREVIATIONS

AFT	Accelerated failure time
AIC	Akaike information criterion
ANCOVA	Analysis of covariance
BoT	Burden of treatment
CI	Confidence Interval
CONSORT	Consolidated Standards of Reporting Trials
CRC	Colorectal cancer
DAOH	Days alive-and-out of hospital
DOD	Date of death
EQ-5D-5L	European Quality of life 5 Dimensions - 5 Levels questionnaire
EQ-Index	EQ-5D index score
EQ-VAS	EQ visual analogue scale
FAS	Full analysis set
FDA	U.S. Food and Drug Administration
HF	Heart failure
HR	Hazard ratio
HRQoL	Health-related quality of life
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
IPW	Inverse probability weighting
MAR	Missing at random
ML	Maximum likelihood
MNAR	Missing not at random
N	Number of observations
NCD	Non-communicable disease
PETS	Patient Experience with Treatment and Self-management questionnaire
PSHP	Perceived support from healthcare personnel
RCT	Randomized controlled trial
REML	Restricted maximum likelihood
RMST	Restricted mean survival time
RPM	Remote patient monitoring
SAP	Statistical analysis plan
SEMCD	Self-Efficacy for Managing Chronic Disease questionnaire
TAU	Treatment as usual
TR	Time ratio

TABLE OF CONTENTS

1 INTRODUCTION 5
1.1 Background and rationale 5
1.2 Intervention 5
1.3 Trial objectives 5
2 TRIAL METHODS 7
2.1 Trial design 7
2.2 Randomisation and blinding 7
2.3 Statistical framework 7
2.4 Timing of outcome assessments 8
2.5 Statistical interim analyses and stopping guidance 8
2.6 Timing of main analysis 8
3 TRIAL POPULATION 8
3.1 Progression of participants through the trial 8
3.2 Adherence and protocol deviations 9
3.3 Analysis populations 10
4 OUTCOME DEFINITIONS 10
4.1 Primary outcome 10
4.2 Secondary outcomes 10
5 ANALYSIS METHODS 11
5.1 Baseline patient characteristics 11
5.2 Methods for primary outcome 11
5.3 Methods for continuous secondary outcomes 14
5.4 Methods for time to event secondary outcomes 14
5.5 Sample size 16
6 STATISTICAL SOFTWARE 16
7 REFERENCES 17
8 APPENDIX 19
Adapted CONSORT 2010 flow diagram 19

1 Introduction

1.1 Background and rationale

Heart failure (HF) and colon-rectal cancer (CRC) are two non-communicable diseases (NCDs) prone to a high rate of hospital admissions and re-admissions, and complex health care needs. For many patients with HF and CRC, self-management following hospitalization can be a challenge, and they may leave the hospital unprepared for self-managing their disease at home. Symptoms and follow-up treatment using remote patient monitoring (RPM) may serve as a valuable addition to traditional follow-up care to support patients in key aspects of self-management. This study aims to provide NCD patients with a nurse-assisted RPM intervention to support self-management, with the potential to improve patients' self-efficacy to self-manage chronic illness and health-related quality of life, and to reduce burden of treatment and hospital readmissions. The efficacy of this intervention will be assessed in a randomized controlled trial (RCT).

1.2 Intervention

1.2.1 Brief description of the study intervention

The intervention group will participate in a 6-weeks nurse-assisted intervention including RPM of vital signs (HF: heart rate, blood pressure, body weight; CRC: temperature, body weight) and of self-reported symptoms, health, and well-being (daily and weekly questionnaires), access to educational information about illness and health resources, and communication (text and video) between the patients and a hospital-based Nurse Navigator.

1.2.2 Control settings

The control groups will receive treatment as usual (TAU).

For patients with HF, some patients may have a follow-up appointment at the hospital HF outpatient clinic at 3-8 weeks post-discharge, if clinically appropriate. The CRC patients will have a follow-up appointment with the surgeon at the gastro-surgical outpatient clinic 4-6 weeks post-discharge. If deemed clinically appropriate, they are also referred to being seen by a primary care nurse post-discharge. In both patient populations, the patients are encouraged to make appointments with their general practitioner. Some may also be offered home nursing.

These services will also be available to the participants in the intervention arm.

1.3 Trial objectives

1.3.1 Primary objective

The main objective of this study is to evaluate the efficacy of the eHealth@H2H intervention in improving patients' self-efficacy to self-manage chronic illness (SEMCD) as compared to treatment as usual (TAU). The primary estimand will be the difference in mean change in SEMCD from baseline to post intervention between the study arms, see Table 1 for further details.

Table 1: Definition of the estimand for eHealth@H2H

Treatment	Treatment condition of interest is the eHealth@H2H intervention, whether or not the participants engage in the intervention and whether or not they receive other treatments; alternative treatment condition is TAU.
Population	Consenting HF and CRC patients eligible for the eHealth@H2H intervention.
Endpoint variable of interest	SEMCD shortly after the intervention period.
Handling of intercurrent events	Non-adherence to the allocated treatment will be handled using the treatment policy strategy (tantamount to an intention-to-treat strategy) (as indicated under Treatment). Non-eligibility uncovered post randomization (before study end) will be handled by the principal stratum approach (i.e., participants excluded from analysis). Withdrawal of consent to use data will also be handled by the principal stratum approach, i.e., with exclusion. Serious illnesses and deaths leading to discontinuation of intervention and of follow-up will be handled by the hypothetical approach, i.e., by handling as missing data.
Population-level summary for the endpoint	Difference in mean change from baseline between treatment conditions.

1.3.2 Secondary objectives

The secondary objectives of this study are:

- To assess long-term effects of intervention on SEMCD.
- To investigate the short- and long-term effects of the intervention on health-related quality of life (HRQoL), burden of treatment (BoT), perceived collaboration and social support from healthcare personnel, and on clinical outcomes (readmissions and death).

1.3.3 Other objectives not treated further in this SAP

We aim to explore and describe the patients' and nurse navigators' engagement with the eHealth@H2H intervention.

The intervention's effect on healthcare utilization will be part of a broader health economic evaluation.

The eHealth@H2H trial includes a nested sub study on the follow-up of HF patients, for which the objectives are to investigate the effects of the intervention on self-care behaviour, medication compliance, and 6- and 12-month readmissions in participants with HF. The results of this sub study will be considered exploratory, and the plan for these analyses will be detailed elsewhere.

2 Trial Methods

2.1 Trial design

The eHealth@Hospital-2-Home study is designed as a randomized, controlled, partially blinded, multi-centre, single-country, study. Treatment allocation is 1:1 to either intervention or TAU.

2.2 Randomisation and blinding

Participants will be allocated by stratified block randomisation (1:1) to either the eHealth@H2H intervention or TAU arm using sequentially numbered, opaque, sealed envelopes. Stratification will occur by diagnostic groups (HF and CRC) and study sites. Separate randomisation lists will be produced for all three strata (CRC@site1, HF@site1, HF@site2) by an independent statistician. The statistician performing the data analysis will be blinded to the randomisation allocation.

2.3 Statistical framework

2.3.1 Hypothesis test

This trial is designed to establish the superiority of the eHealth@H2H intervention compared to TAU regarding change in self-efficacy for self-care among people with chronic illnesses.

- The primary null hypothesis is that the intervention is no different than TAU with regards to patients' self-efficacy to self-care reported at the follow-up at six weeks after hospital discharge.
- The primary alternative hypothesis is two-sided, i.e., that the intervention is either better or worse than TAU with regards to patients' self-efficacy for self-care reported at six weeks.

There is only one identified primary analysis in this trial. All other efficacy analyses will be regarded as supportive or exploratory.

2.3.2 Confidence intervals and p-values

All efficacy estimates will be presented with two-sided 95% confidence intervals (CI). All calculated p-values will be two-sided and compared to a 5% significance level. If less than 0.05, the corresponding treatment group difference will be denoted as statistically significant. As there is only one primary null hypothesis to be tested in this trial, there will be no adjustments for multiplicity.

2.3.3 Decision rule

This trial is designed to address a single primary outcome. Superiority is claimed if the primary null hypothesis is rejected on the significance level (alpha) of 0.05 (two-sided) and the SEMCD scores at six weeks are higher in the intervention arm than in the control arm.

2.4 Timing of outcome assessments

The target days and visits windows are defined as:

Visit label	Target day	Day window
Baseline	Day 0 (day of discharge)	Target day minus 0-7 days
Post-1	Day 42 (6 weeks post discharge)	Target day plus 0-14 days
Post-2	Day 183 (24 weeks post discharge)	Target day plus 0-30 days

Baseline assessments are performed when the patients are considered ready to be discharged (i.e., post-operatively for CRC patients). However, postponement of discharge may happen due to further testing or deterioration in health. In general, data from outside of day windows will not be excluded from analyses.

2.5 Statistical interim analyses and stopping guidance

There will be no formal interim analyses conducted during the trial. However, HF/CRC patient group data may be analysed separately once data collection for each patient group is complete. These analyses will not influence or delay the finalization of the trial for the other patient group.

2.6 Timing of main analysis

The main analysis is planned when all participants have concluded the intervention period, all data up-to and including Post-2 have been entered and verified, the primary database has been locked, and the SAP has been finalized and made available on ClinicalTrials.gov.

3 Trial Population

3.1 Progression of participants through the trial

A CONSORT flow diagram (see Appendix) will be used to summarise the number of patients who were:

- Screened
- Excluded after screening
- Completed baseline
- Excluded after baseline assessment*
- Randomised

Then separately for each randomisation arm, the patients who were:

- Allocated to the randomisation arm
- Withdrawn or lost to follow-up before Post-1*
- Assessed for Post-1
- Withdrawn or lost to follow-up before Post-2*

- Assessed for Post-2
- Included in the primary analysis
- Excluded from the primary analysis*

*reasons will be provided.

Furthermore, separate CONSORT flow diagrams will be presented for each patient group (HF/CRC).

3.1.1 Follow-up and withdrawal

The trial includes three assessments: baseline, Post-1 and Post-2. The status of eligible and randomised patients at trial end will be tabulated by treatment arm (and additionally per randomisation strata) according to

- Completed all assessments
- Dropout after baseline assessment
- Dropout after post-1 assessment

Each dropout will be categorized as either

- Withdrawal
- Death
- Serious illness
- Lost to follow-up

3.2 Adherence and protocol deviations

3.2.1 Adherence to allocated treatment

In this trial, adherence will refer to the extent to which participants in the intervention group actively engage with, are receptive to, and/or use the materials (i.e., symptom checklists) or recommended resources (i.e., monitoring equipment) of the eHealth@H2H intervention, including completing daily and weekly measurements and symptom checklists. Data on adherence will be collected from the RPM software DignioPrevent. Exploration and classification of adherence is an exploratory objective of the study.

Non-adherence to the allocated treatment will be handled using the treatment policy strategy (tantamount to an intention-to-treat strategy) (as indicated under Treatment) and therefore does not need to be pre-defined.

3.2.2 Protocol deviations

In addition to non-adherence, the following are defined as major protocol deviations (intercurrent events):

- Entering the trial when the eligibility criteria should have prevented trial entry
- Withdrawal of consent to use data

Non-eligibility uncovered post randomization (before study end) will be handled by the principal stratum approach (i.e., participants excluded from analysis), as will any withdrawals of consent to use data.

3.3 Analysis populations

The Full Analysis Set (FAS) will be defined as all eligible patients randomly assigned to either the eHealth@H2H intervention or to TAU.

4 Outcome Definitions

4.1 Primary outcome

4.1.1 Self-efficacy for managing chronic disease (SEMCD)

The primary outcome is the patients' self-reports of self-efficacy to self-manage their HF or CRC disease as measured by the six-item questionnaire Self-Efficacy for Managing Chronic Disease (SEMCD) (Lorig 2001). The items ask participants to rate their confidence in doing activities across several domains common for many chronic diseases (i.e., symptom control, role function, emotional functioning, and communicating with physicians). All items are scored on a 10-point scale (1=not at all confident to 10=confident). The total SEMCD score is the mean score of the six items, provided at least four items have responses.

4.2 Secondary outcomes

4.2.1 Health-related quality of life (HRQoL)

HRQoL will be measured using the validated and reliable EuroQoL 5-Dimension 5-Level (EQ-5D-5L) questionnaire (Herdman 2011). It includes five health dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has a five-level response range from 1) no problem to 5) extreme problems, on the day of completing the questionnaire. The five-dimensional health state is translated into one overall score, the EQ-5D index score (EQ-Index) using a scoring algorithm based on derived value sets anchored in 1 (=full health) and 0 (=dead), with values <0 indicating health states valued less than death. We will use the newly developed Norwegian value set (Garratt 2025) giving possible values between -0.453 and 1.

The second part of the EQ-5D-5L questionnaire consists of a visual analogue scale (EQ-VAS), where patients report on their overall health ("your health today") using a scale between 0 ("worst imaginable health") and 100 ("best imaginable health").

Missing values for any of the EQ-5D-5L health state dimensions will not be imputed, meaning the EQ-Index will be missing too.

4.2.2 Days alive-and-out of hospital (DAOH)

Days alive and out of hospital (Cohn 2009) will be collected from hospital records and subtracting the number of days spent away from home due to hospitalisation from the date of discharge up to date of death/six months (Post-2), i.e. possible range will be 0-180 days.

4.2.3 Time to first readmission or death

Time of first readmission within six months will be collected for hospital records, as will date of death (DOD) for those who died within six months from the date of discharge from the primary hospitalisation. The first of these dates will define the end point for this outcome.

4.2.4 Burden of treatment (BoT)

Burden of Treatment comprise patient self-management work and how it may affect wellness and will be measured using parts of the Norwegian translation of the Patient Experiences with Treatment and Self-Management (PETS) (Eton 2017, Husebø 2018, Eton 2020). Four dimensions will be used to focus on treatment burden related to areas of non-adherence in chronic illness: medical information (seven items on five-point scales from "very easy" to "very difficult", plus the option "not applicable"), medications (seven items), medical appointments (six items), and monitoring health (two items); for the three latter dimensions all items are measured on five-point scales from "not at all" to "very much" a problem. All items refer to a recall period of the past four weeks. Each dimension will be presented as a mean score rescaled to 0-100 (high score indicating high burden) provided >50% of items for the respective dimension have valid responses (Eton 2020).

4.2.5 Perceived support from healthcare personnel (PSHP)

Perceived support from healthcare personnel will be measured using 12 items on constructive support (Karlsen 2004, Oftedal 2011). All items are measured on a five-point Likert scale, ranging from 'agree strongly' to 'disagree strongly'. Sum score with range 0-40 with high scores indicating high levels of perceived support. Missing items will be imputed using the mean of valid responses provided >50% (i.e., at least seven) items have valid responses.

4.2.6 Collaboration with healthcare personnel

Collaboration with healthcare personnel will be measured using the validated three-item questionnaire CollaboRATE (Elwyn 2013, Forcino 2018). The three items represent three core shared decision-making activities: explanation of the health issue, preference elicitation and preference integration (Drivenes 2019). Each item is scored on a 10-point anchor scale from 0 ("No effort was made") to 9 ("Every effort was made"). Only instances where all three items have been completed will be given a total score, which will be the mean of the three items applying the "Means score" scoring approach.

5 Analysis Methods

5.1 Baseline patient characteristics

The patient demographics and baseline characteristics to be summarised include study site, diagnosis group, age in years, gender, education, work status, habitus status, comorbidity, and frailty. These will be summarised for all FAS data by randomised treatment arm and overall using descriptive statistics (N, mean, standard deviation, median, 25/75 percentiles, minimum, and maximum) for continuous variables, and number and percentages of patients for categorical variables. There will be no statistical analysis of differences. Any clinically important imbalance between the treatment arms will be noted.

5.2 Methods for primary outcome

5.2.1 Descriptive statistics

Descriptive statistics will include mean and standard deviation for each follow-up time point by treatment arm. Descriptive statistics will be based on non-imputed FAS data, thus the number of evaluable outcome measurements at each time point will also be presented.

5.2.2 Primary inferential analysis

The primary assessment of the intervention effect will be performed using a constrained mixed linear model (Coffman 2016, Twisk 2018) in which both baseline and post-intervention SEMCD scores will be included in the dependent variable, and with fixed effects of measurement occasion as a categorical variable (Post-1 or Post-2 vs. baseline) and of the interaction between study arm (intervention or control) and measurement occasion (i.e. no main effect of study arm), in essence constraining the model to allow for no systematic differences between the study arms at baseline, i.e. any differences seen at baseline will be attributed to random effects. The model allows for potentially differing treatment effects at Post-1 and Post-2. The stratifying variables used in the randomisation (site and diagnosis) will also be included in the model as fixed effects (as an interaction term, in effect producing three categories: CRC@site1, HF@site1, HF@site2).

To account for dependencies between repeated measurements, the model will incorporate subject-specific random intercepts and random effects of measurement occasion. This mixed-model approach ensures that effect estimates are unbiased given the assumption of missing at random (MAR) missingness and dropout (see, e.g., Ibrahim & Molenberghs 2009, or Allison 2012).

Important predictors of SEMCD or of dropout (see Section 5.2.4), will be included in the model to increase precision in the effect estimates (Senn 2022, FDA 2023, Kahan 2014) and to increase the likelihood that the MAR assumption holds, thereby minimizing potential bias from dropouts (Groenwold 2012).

Stata function mixed will be used for the analysis, with function call:

mixed SEMCD i.time studyarm#time site#diagnosis [baseline covariates] || patID: i.time, reml

All analyses will be conducted on the FAS data, and the model will be fitted using restricted maximum likelihood (REML). The main comparison will be at the six weeks (Post-1) timepoint.

5.2.3 Effect estimates

The primary effect estimator will be assessed by the average marginal difference in SEMCD between treatment arms at Post-1 evaluated from the model in Section 5.2.2 and will be presented with a 95% CI and with the p-value from the two-sided contrast test. Superiority of intervention will be claimed if the hypothesis of no difference at Post-1 is rejected on an alpha level of 0.05 and if also the (adjusted) mean SEMCD in the intervention arm is higher than in the TAU arm at Post-1 (see also Section 2.3).

Standardized effect size will be calculated by dividing the model-based difference at Post-1 by the pooled standard deviation of SEMCD at baseline.

5.2.4 Selection of baseline variables for adjustment

Baseline variables will be selected for adjustment in a two-stepped approach:

1) Based on improving the prediction of post SEMCD scores after accounting for baseline SEMCD and the stratifying variables (i.e., disregarding treatment arm). Hence, the model in Stata syntax (fitted using ordinary maximum likelihood (ML) estimation):

mixed SEMCD@Post i.time site#diagnosis SEMCD@BL [other baseline covariates] || patID: i.time

A variable will be included if its addition reduces the Akaike Information Criterion (AIC) by at least 2 points, indicating a meaningful contribution to predictive power while avoiding overfitting.

2) Based on improving the prediction of dropout after accounting for baseline SEMCD, stratifying variables, and the variables identified under Step 1 (i.e., still disregarding treatment arm).

Dropout due to death or serious illness is more relevant regarding potential missing not at random (MNAR) missingness, and minimizing bias due to potential MNAR missingness is our main concern here. Therefore, dropout will be categorized as 1) death or dropout due to serious health issues before Post-1, 2) death or dropout due to serious health issues before Post-2, 3) drop-out due to other reasons (e.g. disinterest) at any time, 4) complete follow-up. This categorical variable will serve as dependent variable in an ordinal regression. In Stata syntax:

ologit dropout_category site#diagnosis SEMCD@BL [baseline variables identified in Step 1] [other baseline covariates]

Additional baseline variables will be included if AIC is reduced by at least 2 points, indicating that their inclusion improves the prediction of dropout and thus enhances the likelihood of the MAR assumption.

Candidates for inclusion in both Steps 1 and 2 are: age, gender, education, work status, habitus status, comorbidity, and frailty, plus baseline scores of the questionnaire secondary outcomes.

In both steps, variables will be added one by one, and the variable that results in the largest reduction in AIC is selected. The models will then be refitted, adjusting for the selected variable, and the process will continue by adding the next variable that leads to the largest AIC reduction until no further variables produce a reduction of at least 2 points in AIC.

5.2.5 Assumption checks and alternative analyses

In case of skewness > 1 in absolute value in the observed SEMCD, the scores will be log-transformed.

In case of convergence problems for the primary model, the model will be simplified by 1) removing random effects of measurement occasion, 2) iteratively removing covariates that were included last in the process in Section 5.2.4.

Residuals and random effects will be assessed for normality using density plots. Heteroskedasticity will be checked by comparing residual variances between groups.

5.2.6 Missing data

Missing observations of baseline variables will be handled by mean imputation (Groenwold 2012, White 2012, White 2005) to not reduce the sample size.

It is assumed that our mixed modelling approach with inclusion of predictors of dropouts will deal with most of the potential bias due to dropout. Effects of potential remaining MNAR dropout will be explored in sensitivity analyses (see Section 5.2.7).

5.2.7 Sensitivity analyses

The following sensitivity analyses will be performed and reported:

1) Combine an ANCOVA mixed modelling (i.e., with baseline SEMCD included as a predictor) with Inverse Probability Weighting (IPW) to deal with dropouts.
2) Handle irregular measurement times by including as a continuous variable and assessing the average marginal difference at time=42 days.
3) Apply alternative regression models, e.g., tobit regression for handling any substantial floor or ceiling effects for the outcome.
4) Allow for potentially heterogenous residual variance.
5) Explore plausible MNAR missingness scenarios.
6) Outcome assessments outside of day windows will be scrutinized for relevance, in particular baseline assessments made more than seven days before discharge, and analyses excluding data obviously not representing state at discharge will be presented.

5.2.8 Subgroup analyses

Possible differences in treatment effect between diagnoses (HF vs. CRC) will be assessed. This will be done by including diagnosis*time and diagnosis*studyarm*time interaction terms as fixed effects in the mixed model used for assessing the overall effect of the intervention (see Section 5.2.2). Estimated treatment effects for each subgroup at Post-1 will be presented, along with the Wald test of the effect of diagnosis*studyarm*{time=Post-1}.

5.3 Methods for continuous secondary outcomes

Effect of the intervention on SEMCD at Post-2 will be assessed in the model described in Section 5.2.2.

Other continuous secondary outcomes (HRQoL, DAOH, BoT, PSHP, collaboration with healthcare personnel) will be analysed in the same manner as SEMCD, but with potentially other adjustment variables (derived as for SEMCD).

5.4 Methods for time to event secondary outcomes

5.4.1 Descriptive statistics

For the secondary outcome Time to first readmission or death (within six months), the number and percentage of patients who were hospitalised and the number and percentage of patients who died within six months will be presented by treatment arm. Survival probabilities for the composite outcome (i.e. time to either death or first readmission) will be presented using Kaplan-Meier curves stratified by treatment arm, and the median and interquartile range of time to event will be derived by Kaplan-Meier estimates. Stata syntax:

stset time, failure(status)
sts graph, by(treatment) risktable
stsum, by(treatment)

5.4.2 Primary inferential analysis

Differences between treatment arms in time to first readmission or death will be assessed with Cox regression using the FAS data. The stratifying variables site and diagnosis will be included in the

model as covariates. Important predictors of time to event (see 5.4.3), will be included in the model to increase precision in the effect estimates (Senn 2022, FDA 2023, Kahan 2014). We will use the Efron method to handles ties. Stata syntax for the final model:

stcox treatment site#diagnosis [other baseline covariates], efron

5.4.3 Selection of baseline adjustment variables

As for the primary outcome, baseline variables will be selected for adjustment in a data-driven but pre-specified manner:

A variable will be included if its addition reduces by at least 2 points the AIC for the prediction of time to event after accounting for the stratifying variables (i.e., disregarding treatment arm). Hence, the model in Stata syntax:

stcox treatment site#diagnosis [other baseline covariates], efron
estat ic

Variables will be added one by one, and the variable that results in the largest reduction in AIC is selected. The models will then be refitted, adjusting for the selected variable, and the process will continue by adding the next variable that leads to the largest AIC reduction until no further variables produce a reduction of at least 2 points in AIC. Candidates for adjustment variables are the same as for the primary outcome.

5.4.4 Effect estimates

The difference between treatment and control will be presented as an adjusted hazard ratio (HR) with a 95% CI and a p-value from the two-sided likelihood ratio (LR) test of no effect (i.e. test of HR=1).

5.4.5 Assumption checks and alternative analyses

The proportionality of the hazards related to treatment and control will be tested using covariate-specific Schoenfeld residuals test, by visual inspection of log-log survival plots, and by statistical testing of time-varying effect of treatment arm via interactions between treatment arm and time or log(time). In case of substantial violation (any of the tests' $p<0.05$ *and* evident non-parallelism in log-log plot), we will use accelerated failure time (AFT) models with distributions optimized via AIC (Weibull, log-normal, or log-logistic). The selected model will be validated using graphical checks (of deviance residuals and comparing observed vs. predicted hazards). Stata syntax:

estat phtest, detail
stphplot, by(treatment)
stcox treatment [other covariates], tvc(treatment) texp(_t)
stcox treatment [other covariates], tvc(treatment) texp(ln(_t))

streg treatment [other covariates], dist(weibull) time

From the selected AFT model, we will report the estimated time ratio (TR) with 95% CI.

5.4.6 Missing data

Loss to follow-up will be handled as non-informative censoring. We do, however, anticipate that we will be able to extract information about time to event or censoring for all of the FAS sample. Missing baseline data will be handled with mean imputation as for the primary (continuous) outcome.

5.4.7 Sensitivity analyses

Compare unadjusted results from the chosen model (Cox or AFT) with those from restricted mean survival time (RMST) analysis, which is a distribution-free alternative.

If any censoring, apply worst-case scenario by reassign censored treatment-arm patients as events and control-arm as censored.

5.4.8 Subgroup analyses

Possible differences in treatment effect between diagnoses (HF vs. CRC) will be assessed by adding an interaction term between diagnosis and treatment to the model.

5.5 Sample size

The sample size goal was set to 240 patients (HF=180, CRC=60) equally divided between the intervention and control group. Assuming a 20% dropout, a sample size of 240 will provide a power of 80% to detect, within a Type I error probability of α=0.05, a medium-sized standardised difference in the primary outcome of 0.4 at the first time point (Post-1) in the diagnosis groups combined.

6 Statistical Software

All statistical analyses will be performed in Stata/SE (StataCorp. 2023. *Stata Statistical Software: Release 18*. College Station, TX: StataCorp LLC).

7 References

Allison, PD. Handling missing data by maximum likelihood. Paper 312-2012 presented at the SAS Global Forum 2012.

Coffman CJ, Edelman D, Woolson RF. To condition or not condition? Analysing 'change' in longitudinal randomized controlled trials. *BMJ Open* 2016; **6**:e013096. DOI: 10.1136/bmjopen-2016-013096.

Cohn J, Cleland JGF, Lubsen J, Borer JS, Steg PG, Perelman M, Zannad F. Unconventional end points in cardiovascular clinical trials: should we be moving away from morbidity and mortality? *Journal of Cardiac Failure* 2009; **15**(3):199-205. DOI: 10.1016/j.cardfail.2008.10.029.

Drivenes K, Haaland VØ, Mesel T, Tanum L. Practitioners' positive attitudes promote shared decision-making in mental health care. *Journal of Evaluation in Clinical Practice* 2019; **25**(6):1041-1049. DOI: 10.1111/jep.13275.

Elwyn G, Barr PJ, Grande SW, Thompson R, Walsh T, Ozanne EM. Developing CollaboRATE: a fast and frugal patient-reported measure of shared decision making in clinical encounters. *Patient Education and Counseling* 2013; **93**(1):102-107. DOI: 10.1016/j.pec.2013.05.009.

Eton DT, Yost KJ, Lai JS, Ridgeway JL, Egginton JS, Rosedahl JK, Linzer M, Boehm DH, Thakur A, Poplau S, Odell L, Montori VM, May CR, Anderson RT. Development and validation of the Patient Experience with Treatment and Self-management (PETS): a patient-reported measure of treatment burden. *Quality of Life Research* 2017; **26**(2):489-503. DOI: 10.1007/s11136-016-1397-0.

Eton DT, Lee MK, St. Sauver JL, Anderson RT. Known-groups validity and responsiveness to change of the Patient Experience with Treatment and Self-management (PETS vs. 2.0): a patient-reported measure of treatment burden. *Quality of Life Research* 2020; **29**:3143–3154. DOI: 10.1007/s11136-020-02546-x.

U.S. Department of Health and Human Services, Food and Drug Administration. Adjusting for covariates in randomized clinical trials for drugs and biological products. Guidance for industry, 2023. https://www.fda.gov/media/148910/download, , accessed August 12, 2025.

Forcino RC, Barr PJ, O'Malley AJ, Arend R, Castaldo MG, Ozanne EM, Percac-Lima S, Stults CD, Tai-Seale M, Thompson R, Elwyn G. Using CollaboRATE, a brief patient-reported measure of shared decision making: results from three clinical settings in the United States. *Health Expectations* 2018; **21**(1):82-89. DOI: 10.1111/hex.12588.

Garratt AM, Stavem K, Shaw JW, Rand K. EQ-5D-5L value set for Norway: a hybrid model using cTTO and DCE data. *Quality of Life Research* 2025; **34**:417-427. DOI: 10.1007/s11136-024-03837-3.

Groenwold RHH, Donders AR, Roes KC, Harrell FE Jr, Moons KG. Dealing with missing outcome data in randomized trials and observational studies. *American Journal of Epidemiology* 2012; **175**(3):210-217. DOI: 10.1093/aje/kwr302.

Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, Bonsel G, Badia X. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research* 2011; **20**(10):1727-1736. DOI: 10.1007/s11136-011-9903-x.

Husebø AML, Morken IM, Eriksen KS, Nordfonn OK. The patient experience with treatment and self-management (PETS) questionnaire: translation and cultural adaption of the Norwegian version. *BMC Medical Research Methodology* 2018; **18**(1):147. DOI: 10.1186/s12874-018-0612-9.

Ibrahim JG, Molenberghs G. Missing data methods in longitudinal studies: a review. *TEST* 2009; **18**:1-43. DOI: 10.1007/s11749-009-0138-x.

ICH official website. https://www.ich.org/page/efficacy-guidelines, accessed August 12, 2025.

Kahan BC, Jairath V, Doré CJ, Morris TP. The risks and rewards of covariate adjustment in randomized trials: an assessment of 12 outcomes from 8 studies. *Trials* 2014; **15**:139. DOI: 10.1186/1745-6215-15-139.

Karlsen B, Idsoe T, Hanestad B, Murberg T, Bru E. Perceptions of support, diabetes-related coping and psychological well-being in adults with type 1 and type 2 diabetes. *Psychology, Health & Medicine* 2004; **9**(1):53-70. DOI: 10.1080/13548500310001637751.

Lorig KR, Sobel DS, Ritter PL, Laurent D, Hobbs M. Effect of a self-management program on patients with chronic disease. *Effective Clinical Practice* 2001; **4**(6):256-62.

Oftedal B, Bru E, Karlsen B. Social support as a motivator of self-management among adults with type 2 diabetes. *Journal of Nursing and Healthcare of Chronic Illness* 2011; **3**(1):12-22. DOI: 10.1111/j.1752-9824.2010.01074.x.

Groenwold RH, Donders AR, Roes KC, Harrell FE Jr, Moons KG. Dealing with missing outcome data in randomized trials and observational studies. *American Journal of Epidemiology* 2012; **175**(3):210-217. DOI: 10.1093/aje/kwr302.

Senn S. Empirical studies of balance do not justify a requirement for 1,000 patients per trial. *Journal of Clinical Epidemiology* 2022; **148**:184-188. DOI: 10.1016/j.jclinepi.2022.02.010.

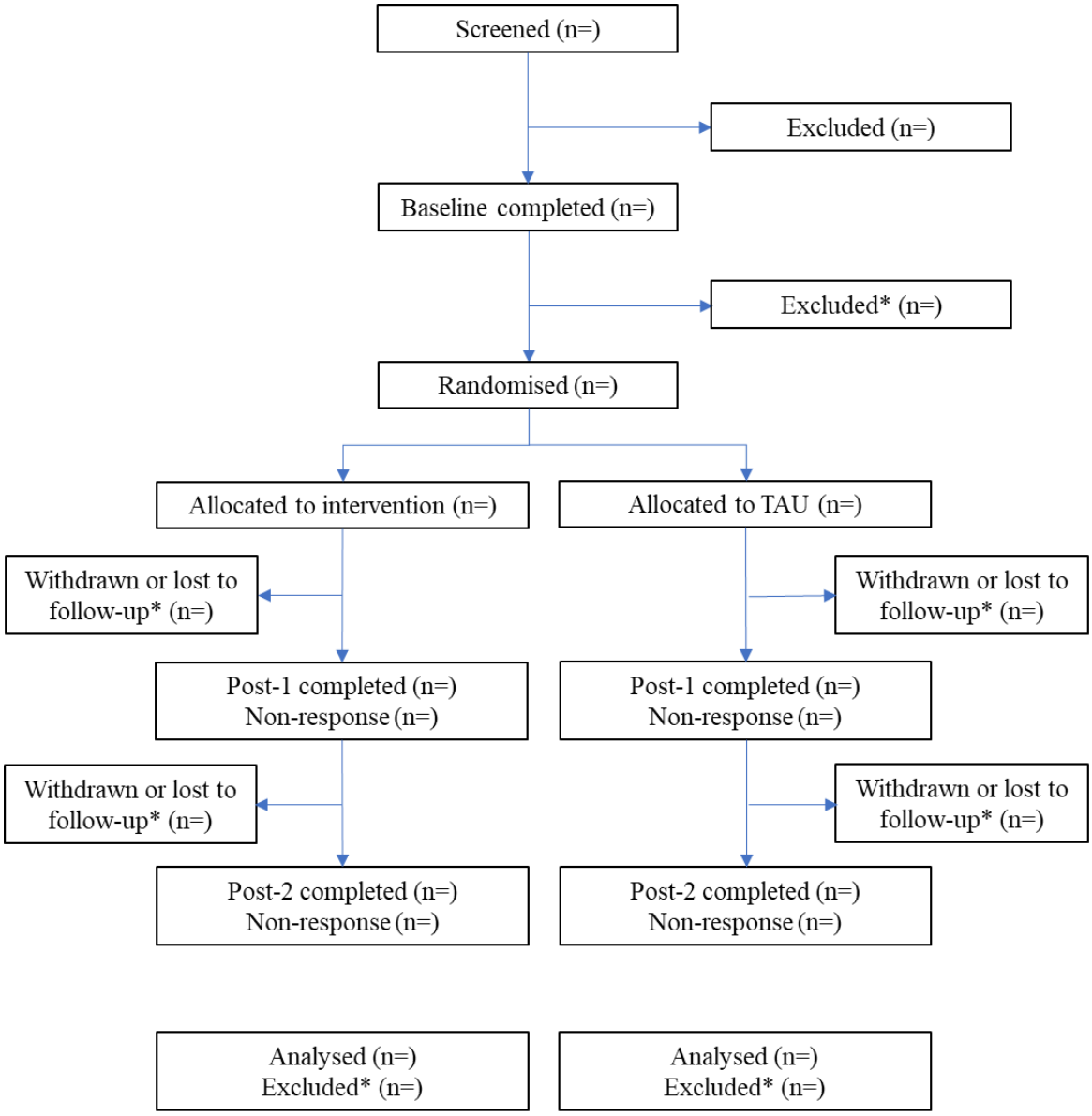
Twisk J, Bosman L, Hoekstra T, Rijnhart J, Welten M, Heymans M. Different ways to estimate treatment effects in randomised controlled trials. *Contemporary Clinical Trials Communications* 2018; **10**:80-85. DOI: 10.1016/j.conctc.2018.03.008.

White IR, Carpenter J, Horton NJ. Including all individuals is not enough: lessons for intention-to-treat analysis. *Clinical Trials* 2012; **9**(4):396-407. DOI: 10.1177/1740774512450098.

White IR, Thompson SG. Adjusting for partially missing baseline measurements in randomized trials. *Statistics in Medicine* 2005; **24**(7):993-1007. DOI: 10.1002/sim.1981.

8 APPENDIX

Adapted CONSORT 2010 flow diagram



* Reasons will be provided