Mannheim, 9 May 2025

FINAL REPORT

Evaluation of a Smartphone App for Heart Failure Patients

for

ProCarement GmbH

ZEW - Leibniz Centre for
European Economic Research
Contact: Prof. Dr. Simon Reif
Health Care Markets and Health Policy
L 7, 1 · 68161 Mannheim · Germany
Tel. +49 621 1235-171
simon.reif@zew.de · www.zew.de

Authors: Simon Reif · Sabrina Schubert · Paul Peters

Leibniz Gemeinschaft

ZEW

SUMMARY

Objective: Heart failure (HF) is a prevalent condition affecting millions of individuals worldwide. Continuous monitoring and targeted behavioral interventions have been shown to improve health status and quality of life for HF patients. Digital therapeutics offer the possibility to make more frequent monitoring and targeted behavioral interventions available for more people. The ProHerz app aims to support patients suffering from HF using easy-to-use monitoring of medical parameters for early detection of disease progression as well as self-help features. This report presents results from a randomized controlled trial (RCT) to assess the impact of the app.

Methods: An RCT with 252 HF patients was conducted where half of patients received access to a digital therapeutic (ProHerz). Clinical indicators as well as patient-reported outcomes were collected at entry and exit examinations. We conduct statistical analyses with and without covariate adjustment and using different imputation strategies for missing values.

Results: We find significant positive effects of the intervention on 6-minute walk test distance (6MWT), self-care behavior (EHFScBS score) and HF-specific health literacy (number of correct answers in AHFKT). The intervention group also showed better progression in NYHA class compared to the control group.

Conclusion: Patients assigned to use the app experienced significant improvements in their condition. The statistical analysis is robust to different sensitivity analyses.

TABLE OF CONTENTS

Summary i

Table of Contents ii

List of Figures iii

List of Tables iv

1 Introduction 1

2 The ProHerz App 2

3 Study Design 4
- 3.1 Inclusion criteria . . . 4
- 3.2 Recruitment and Examinations . . . 4
- 3.3 Randomization . . . 6
- 3.4 Endpoints . . . 7
- 3.5 Imputation of missing values . . . 8
- 3.6 Statistical Analyses . . . 9
- 3.7 Sample description . . . 10

4 Results 11
- 4.1 Full Analysis Set . . . 11
- 4.2 Predictive Mean Matching Imputation . . . 15
- 4.3 Per Protocol Sample . . . 15
- 4.4 Explorative Analyses . . . 18

5 Discussion 22

6 Conclusion 23

References 23

Appendix 30

LIST OF FIGURES

1 User View of the ProHerz app. 4
2 CONSORT Flow Diagram 6
3 Subgroup Analyses 14
3 Change in NYHA Class 18
4 Distribution of Measures per Week 20
5 Evolution of Outcomes over Time 21
A.1 Differences in Outcome Measures by Group 31

LIST OF TABLES

1 Sample Overview 11
2 Main Results - Full Analysis Set 13
3 Responder Analysis EHFScBS - Full Analysis Set 13
4 Main Results - PMM Imputation 15
5 Main Results - Per Protocol Sample 16
6 Health Care Utilization - Per Protocol Sample 18
7 Correlation of Usage Intensity on Outcomes - Full Analysis Set 21
A.1 Results without Covariate Adjustment - Full Analysis Set . . . 32
A.2 Responder Analyses KCCQ - Full Analysis Set 32
A.3 Responder Analyses EHFScBS - Full Analysis Set 33
A.4 Responder Analyses SDMQ - Full Analysis Set 33
A.5 Results Full Analysis Set - Subsample Age < 65 - N = 121 33
A.6 Results Full Analysis Set - Subsample Age ≥ 65 - N = 131 . . . 34
A.7 Results Full Analysis Set - Subsample Men - N = 190 34
A.8 Results Full Analysis Set - Subsample Women - N = 62 35
A.9 Results Full Analysis Set - Subsample NYHA Class 1 | 2 - N = 150 35
A.10 Results Full Analysis Set - Subsample NYHA Class 3 - N = 102 . 36
A.11 Results Full Analysis Set - Subsample Interim Participation - N = 88 36
A.12 Results Full Analysis Set - Subsample No Interim Participation - N = 164 37
A.13 Results without Covariate Adjustment - PMM Imputation . . . 38
A.14 Results without Covariate Adjustment - Per Protocol Sample . 39
A.15 Change in NYHA Class - Per Protocol Sample 40
A.16 Change in NYHA Class - Full Analysis Set 40
A.17 Statistical Summary of Usage Intensity 41
A.18 Descriptive Statistics Split by Interim Assignment 41
A.19 Results with Interim Examination Control - Full Analysis Set . . 42
A.20 Results with Interim Examination Control - PMM 42
A.21 Results with Interim Examination Control - Per Protocol Sample 43

1 INTRODUCTION

Worldwide, heart failure (HF) prevalence is widespread and increases with aging populations (Groenewegen et al., 2020; Roth et al., 2015; Savarese & Lund, 2017). Poor self-management of the condition often leads to hospitalizations and generally increases costs to health systems (Riegel et al., 2009). Widely used approaches to control risk factors are public health promotion, face-to-face health education, and telephone consultations (Lloyd-Jones et al., 2022; Mok et al., 2013; Wood et al., 2008; Zhu et al., 2024). Due to their time-consuming and labor-intensive nature, the idea of using additional tools was first implemented with mobile health (mHealth) tools (such as SMS reminders) and in the last decade also brought digital therapeutics to the market (Cruz-Cobo et al., 2022; Jin et al., 2019). Modern technology can support patients with tools to support lifestyle changes, improve patient knowledge about medical conditions, enhance self-care behavior, guide individual therapy decisions and may hence improve health and avoid costs (Hamine et al., 2015; Robinson et al., 2021).

In this report, we present results from a randomized controlled trial (RCT) to assess the benefits of a digital therapeutic to support patients with HF — the ProHerz app. This RCT builds on the pilot study described in Reif et al. (2022). Key elements of the app — installed on patients' smartphones or tablets — are easy-to-use monitoring of medical parameters for early detection of disease progression as well as self-help features to guide patients.

The research on digital health apps and telemonitoring designed for patients with HF has grown in the past decade. However, these studies are heterogeneous, both in terms of what functionality the digital interventions provide as well as outcome measures used to assess their effectiveness (Coorey et al., 2018; Zhu et al., 2024). Several digital therapeutics for HF patients have already been evaluated in RCTs, yielding mixed results (Cajita et al., 2016). Popular outcomes to assess the effect of a digital therapeutic for HF patients include the New York Heart Association (NYHA) classification for patients with HF, objective measures like oxygen uptake, BMI, or blood pressure, but also subjective measures like the perceived quality of life (QoL). An early systematic review of ten interventions finds that the results for app usage and health outcomes (e.g. NYHA class or number of hospitalizations) are inconsistent at best (Cajita et al., 2016). In a recent meta-analysis of 34 RCTs, Zhu et al. (2024) show that there are fewer adverse events and significant im-

provements in maximum oxygen consumption (VO2 max) in the intervention groups. There are however no significant changes in other clinical markers like BMI or blood pressure and there is high heterogeneity across studies. Recent RCTs generally find an improvement for health status measures like NYHA, BMI, and blood pressure, or physical activity in the treatment group compared to the control group but effects are, in general, small in magnitude (Choi et al., 2023; Gallagher et al., 2023; Ni et al., 2022; Saleh et al., 2023; Zhu et al., 2024). The effectiveness of app-based digital therapeutics has also been analyzed using more subjective outcome measures like QoL. Some RCTs find that using targeted applications significantly increases QoL for patients with HF (Davoudi et al., 2020; Saleh et al., 2023; Varnfield et al., 2014; Widmer et al., 2015) while others do not find a significant improvement (N. Johnston et al., 2016; Victoria-Castro et al., 2024).

Furthermore, self-care and self-management behaviors are increasingly recognized as critical outcomes for HF patients. Effective self-management and self-care can reduce hospitalizations and mitigate disease progression (Lee et al., 2017). Some studies find significant improvements in self-care behavior scores after app use (N. Johnston et al., 2016; Zhu et al., 2024). W. Johnston et al. (2022) highlight the potential for apps integrated with tools like activity trackers to enhance self-care by improving patients' confidence and motivation. However, the overall impact on self-management remains mixed across studies, as some studies find no improvement among app-treated patients (Gallagher et al., 2023; Redfern et al., 2020)

In summary, previous research shows that improvements in health status, QoL and self-management are possible but depend on the specific design of the intervention. This report proceeds as follows: In the next section, we first introduce the digital therapeutic used as the intervention in this study. Next, we explain the study design in detail, followed by the presentation of the results. Finally, we conclude by summarizing the key implications of the study.

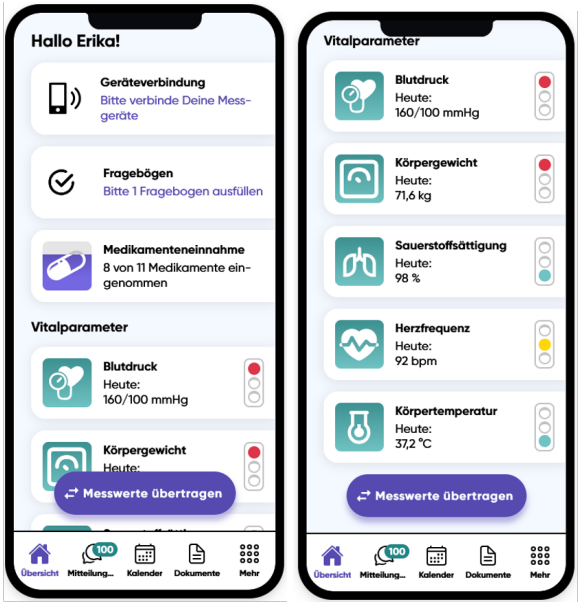
2 THE PROHERZ APP

The ProHerz app is a software application for patients suffering from HF, available on the Android Play Store and Apple App Store since March 2021. Since May 2023, patients in the German Social Health Insurance have been able to use ProHerz under a preliminary listing as an "App on Prescription" in

the DiGA scheme. The app aims to provide positive care effects to patients, in particular improved health status and QoL, as well as improvements in health literacy, self-care behavior, and better involvement of patients in medical decisions. The core function of the app is the provision of individualized recommendations based on regular measures of relevant patient parameters. The digital diagnostic component of the app consists of regular measurement and documentation of relevant vital signs, archiving and providing easy access to medical documents and daily medication reminders, and delivering educational information.

Vital signs collected by the app are blood pressure, heart rate, blood oxygen saturation, body weight, and temperature. These measurements are collected either by medical devices via Bluetooth or by manual patient input. Regular screening of vital signs offers more detailed surveillance than the current standard of care in many countries, including Germany. The measurements — graphically visualized and presented to the patient with additional information in a comprehensible way — allow the algorithm behind the app to identify deviating levels at an early stage and inform patients and, if patients share their data, also physicians. Following the detection of deviating patterns (e.g. signs of decompensations), the digital therapeutic component of the app advises patients to seek medical advice in the early stages of disease progression and incentivizes behavioral changes. The app also regularly monitors patient well-being and keeps track of possible comorbidities. In addition, the app offers a platform to save all medical documents and conveniently share them with health care providers. Finally, the app sends medication reminders and patients can track their medication schedule. Two screenshots from the patient-view of the ProHerz app are shown in Figure 1.

Figure 1: User View of the ProHerz app.



3 STUDY DESIGN

3.1 Inclusion criteria

Patients included in the study needed to be at least 18 years old, have a diagnosed HF (ICD-10 Code I50.-) with NYHA classes I, II, and III, validated as EF ‹ 40% or NT-proBNP › 500. In addition, patients needed to be able to speak and write German, use a smartphone app and give written consent to take part in the study. Patients were excluded if they had cognitive impairment, had severe HF (NYHA class IV), had been hospitalized within eight weeks prior to the entry examination, had recently undergone or were scheduled for a revascularization, or had an alcohol or drug addiction. In addition, patients could not participate if they were enrolled in any other clinical trial, including the pilot study on the ProHerz app (Reif et al., 2022).

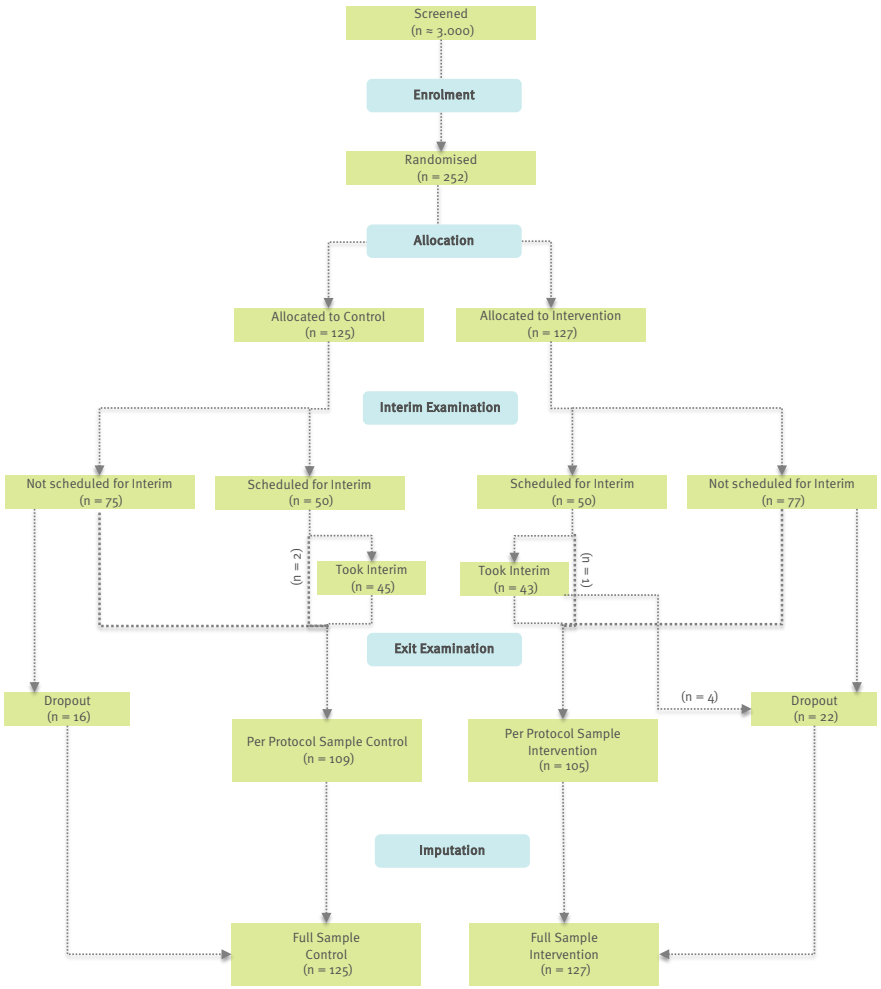
3.2 Recruitment and Examinations

252 patients in nine study sites were included between March 2022 and June 2024.[1] The exit examination was scheduled for all patients six months after

[1] Participating hospitals: Hospital Nürnberg Süd, University Hospital Munich, Albertinen-Krankenhaus Hamburg, HDZ NRW, Klinikum Fulda, SLK-Kliniken Heilbronn, Klinikum Bamberg, University Hospital Essen, University Hospital Greifswald

entry took place between September 2022 and December 2024. For the first 100 patients, an interim examination took place between June and November 2022. During all examinations, patients first had a conversation with a physician where patients gave their written consent to participate in the study. The physician collected patient characteristics (sex, age, BMI, comorbidities from the Elixhauser index, current medication, smoking status, and whether the patient lived on their own), and added them to the study's electronic case report forms (eCRF). Next, patients completed a set of questionnaires on a dedicated tablet and underwent a 6-minute walk test, supervised by a study nurse. There was no requirement that the same physician or nurse performed all examinations for one patient. The CONSORT flow diagram is presented in Figure 2.

Figure 2: CONSORT Flow Diagram



3.3 Randomization

After the entry examination, patients were randomized to either receive access to the digital therapeutic ProHerz (intervention group) or remain with the standard-of-care (control group). Patients were assigned to either group via stratified randomization following Pocock's algorithm according to gender (male / female), age (< 65 / ≥ 65) and NYHA class (I and II, III) as reported in the eCRF. Study nurses and physicians were not informed about the outcomes of the random assignment and patients were instructed not to disclose their status to clinical staff. Patients in the intervention group received

Bluetooth devices for their regular measurements (blood pressure monitor, pulse oximeter, scale, body thermometer) and detailed instructions on how to use the app and the devices by mail.

3.4 Endpoints

The following **primary endpoints** measures were used in this study:

- Differences in health status are evaluated using the change in distance walked in meters in the 6-minute walk test (**6MWT**). The 6MWT is a widely used measure of physical performance, where larger distances indicate better health status (Demers et al., 2001).
- Differences in QoL are evaluated by a change in the Kansas City Cardiomyopathy Questionnaire (**KCCQ**) index. The KCCQ is a measurement tool validated in German to assess the quality of life for patients with chronic HF (Faller et al., 2005). The questionnaire consists of twelve questions with answers on ordinal scales from either 1 to 5 or 1 to 6. Each question belongs to one of four dimensions. For each dimension a score is constructed where the sum of the ordinal points is transformed to range between 0 and 100, with higher values indicating higher quality of life. In this analysis, we use the summary score over all four dimensions, which is calculated as the average of the four dimension scores (Spertus & Jones, 2015).
- Differences in self-care behavior are evaluated by a change in the 9-item European Heart Failure Self-care Behaviour Scale (**EHFScBS**). The EHFScBS is a health behavior questionnaire validated in German for patients with HF (Köberich et al., 2013). The responses to the nine questions are stored as numeric values in nine variables, ranging from 1 (“I completely agree”) to 5 (“I completely disagree”). These values are then recoded such that 1 corresponds to “I completely disagree” and 5 to “I completely agree.” The sum of the recoded values is subsequently scaled to create an index ranging from 0 (worst possible score) to 100 (best possible score), with higher values indicating better self-care behavior (Vellone et al., 2014).
- Differences in health literacy are evaluated by the number of correct answers on the Atlanta Heart Failure Knowledge Test (**AHFKT**), a validated knowledge test on HF (Reilly et al., 2009). Each correct answer

to one of the thirty questions is coded as 1 and an index is created by summing the number of all correct answers. A higher number of correct answers indicates better health literacy.

In addition, data on the following **secondary endpoints** were collected:

- Differences in general health-related QoL are assessed using the change in the **EQ5D** index between entry and exit examination. The EQ-5D-L questionnaire is often used for health economic evaluations and has been validated for Germany (Ludwig et al., 2018). Its values can range from 0 to 1 where higher values indicate higher quality of life.
- Differences in the involvement in care decisions are evaluated using the Shared Decision Making Questionnaire (**SDMQ**) score (Kriston et al., 2010). The answers to each of the nine questions range from 1 ("completely disagree") to 6 ("completely agree") and the summary score is generated by transforming the sum of these ordinal values to range between 0 and 100, where higher values indicate better participation in the decision-making process.
- Differences in HF-specific health status is assessed using the change in NYHA class. The NYHA classification rates the severity of HF from class I, the least severe condition, to class IV, the most severe condition (Bennett et al., 2002).
- Number of hospital and ambulatory care visits.

3.5 Imputation of missing values

We employ three different strategies to deal with missing values. First, all results are reported for the full analysis set (FAS) containing all patients who had an entry examination. For the FAS, all endpoints that are missing for the exit examination are imputed using a jump to reference approach based on average values in the control group. Specifically, for endpoints that are collected during entry and exit examination, the average change between entry and exit is derived for the control group and this change is added to the values from the entry examination to obtain an imputed endpoint for patients with missing values in the exit examination. The average change in the control group was used to impute the exit examination endpoints for both, the control as well as the intervention group. For endpoints that are collected

only at the exit examination, the average value in the control group was used as the imputed endpoint for all patients.

Second, we use predictive mean matching (PMM) to impute missing values in the exit examination. PMM imputes missing values by first predicting them with a regression model fitted on observed data. Then we identify observed values with predicted values closest to the predicted value of the missing ones. One of these nearest observed values is randomly selected and used to impute the missing value, ensuring realistic imputations that preserve the distribution of the original data (Morris et al., 2014). Third, we limit our analysis to the per protocol sample (PPS) which only includes patients with valid exit examination values.

3.6 Statistical Analyses

The objective of this study is to evaluate potential benefits for HF patients of using the ProHerz app for six months. The effectiveness of the app is measured by changes in health status, QoL, self-care behavior, health literacy, and general involvement of patients in medical decisions. All data preparation and data analysis were conducted using R (Version 4.3.2). The maximum acceptable level of statistical uncertainty is set at an alpha error of 5%, with a Bonferroni correction applied to account for multiple hypothesis testing for four primary outcomes. Accordingly, a p-value of <0.0125 indicates a significant difference. The secondary analyses are exploratory and are not adjusted for multiple comparisons.

For the primary endpoints 6MWT, KCCQ, and EHFScBS and the secondary endpoint EQ5D, data is available from the entry as well as the exit examination. The main estimate here is the difference in outcome Y at the exit examination (Y_i^{exit}) between intervention and control group adjusting for the level of Y at the entry examination (Y_i^{entry}) and a set of covariates X. The linear regression model is described below:

$$Y_i^{exit} = \beta_1 Intervention_i + \beta_2 Y_i^{entry} + \gamma X + \epsilon_i$$

Here, $Intervention_i$ is an indicator variable which is equal to 1 if the patient was assigned to the intervention group and 0 if the patient was assigned to the control group. The main coefficient of interest is β_1, which captures the difference in Y between intervention and control group, adjusted for the initial values of outcome Y and a set of covariates X. To account for the strati-

fication variables in the empirical analysis, X in addition contains three indicator variables for gender (male), age (under 65), and disease severity (NYHA class I and II). These variables were used as input into the stratified randomization. In a sensitivity analyses, we repeat the analyses without covariates where X is just an allones vector to estimate the intercept.The idiosyncratic error term ϵ_i contains unexplained variation in Y_i^{exit}.

For the primary outcome AHFKT and the secondary outcome SDMQ, no baseline values are collected, hence no adjustment for initial values is performed. Therefore, the difference in the outcomes only collected in the exit examinations between the intervention and control groups is estimated with the linear regression model described below:

$$Y_i^{exit} = \beta_1 Intervention_i + \gamma X + \epsilon_i$$

In addition to the point estimates, we report effect size estimates (Cohen's d) for the analyses. The estimations are performed for the three different samples (FAS, PMM, PPS) each with covariate adjustment as the main specification and without covariate adjustment as sensitivity analyses. In addition, we conduct exploratory analyses using different subgroups of the sample and additional outcomes.

3.7 Sample description

The majority of patients come from three sites: Hospital Nuremberg Süd (53%, n=133), University Hospital Essen (22%, n=55), and University Hospital Munich (13%, n=32). The remaining 12% of the sample comes from six additional study sites, each contributing between 2 and 12 patients. An overview of the sample is presented in Table 1. In total, 252 patients took part in the entry examination and 50.4% (n=127) were randomly assigned to the intervention group. 214 patients completed the exit examination (87%, (n=109) of patients in the control group and 83% (n=105) of patients in the intervention group). The average intervention time was 189 days for the intervention as well as the control group. Patients were, on average, 64 years old at the time of the entry examination both for the intervention and the control group. Three-quarters of the patients in the sample are men. There are slightly more men in the intervention group with 77% (n=98), compared to 74% (n=92) in the control group, although the difference is not statistically significant. One-quarter (n=31) of patients report living alone in both

groups. There are slightly more patients in the intervention group who report to be smokers (18%, n=23) compared to the control group (11%, n=12) but the difference is not significantly different from zero. The intervention group is slightly healthier (their Elixhauser score based on the van Walraven algorithm (van Walraven et al., 2009) was 6.8 compared to 7.9 in the control group) and has a marginally lower BMI of around 27.8 kg/m^2 compared to the control group average of 28.3 although both differences are not statistically significant. Both groups have similar HF-specific health conditions where around 15% are categorized as NYHA class I, 45% as NYHA class II and 40% as NYHA class III.

Table 1: Sample Overview

	Control (N=125)		Intervention (N=127)			
	Mean	Std. Dev.	Mean	Std. Dev.	Diff. in Means	p
Study Characeristics						
Follow-up	87.20%		82.68%		-4.52	0.317
Intervention time	188.58	19.06	188.69	18.51	0.11	0.967
Sociodemographics						
Age	64.45	13.03	64.38	13.47	-0.07	0.967
Male	73.60%		77.17%		3.57	0.513
Living alone	25.81%		24.41%		-1.40	0.800
Smoker	11.38%		18.11%		6.73	0.134
Health Status						
Elixhauser score	7.95	6.27	7.35	6.84	-0.60	0.470
BMI	28.26	6.17	27.67	5.21	-0.60	0.406
NYHA Class						
I	14.40%		15.75%		1.35	0.766
II	44.80%		44.09%		-0.71	0.911
III	40.80%		40.16%		-0.64	0.918

4 RESULTS

The results in this section are structured as follows: We first present the main estimations for the FAS and subgroup analyses. Second, we compare these results to estimates from the sample after more complex PMM imputation. Third, we repeat the analysis for the PPS and add additional secondary outcomes that are challenging to impute.

4.1 Full Analysis Set

The main results for the pre-specified estimations using the FAS without covariate adjustment as well as the means for both treatment and control group

at entry and exit are presented in Table A.1 (in the Appendix). A graphical representation of the unadjusted main results is in the boxplots in Figure A.1 (in the Appendix). We present results with covariate adjustment in Table 2.

The 6MWT distance increased from 403m to 434m in the intervention group and slightly decreased from 388m to 386m in the control group. With covariate adjustment in the main model specification, these differences yield a significant +33.1m main effect (CI-98.75: 5.76 ; 60.46) from the regression analysis, corresponding to an effect size of 0.38. There is a small increase in the KCCQ score for both groups (Intervention +3.9; Control +4.7), but the difference is not statistically different from zero in the main model. While both groups exhibited higher EHFScBS scores in the exit examinations, the increase was larger among patients in the intervention group. This difference leads to a significant main effect of +5.2 in the regression analysis (CI-98.75: 0.67 ; 9.72) with a corresponding effect size of 0.37.

The number of correct answers in the AHFKT was significantly higher in the intervention group (18.8) compared to the control group (16.8), yielding a point estimate in the regression analysis of +2.04 (CI-98.75: 1.07 ; 3.01) and a corresponding effect size of 0.67. EQ5D weights are, on average, 0.84 for the control as well as the intervention group both at the entry and exit examinations and there is no significant difference in the main estimation. SDMQ scores are slightly higher at the exit examination for the intervention group (57.9) compared to the control group (54.1), although the regression coefficient of +3.48 is not significantly different from zero (CI-98.75: -5.91 ; 12.87).

We also conduct a responder analysis for EHFScBS using the threshold of 70 which Wagenaar et al. (2017) consider appropriate for discriminating between inadequate and adequate self-care behavior. The results in Table 3 show that the intervention group was more often above the threshold at the exit examination compared to the control group. This difference is only significant at the 10% level due to the loss of power from dichotomizing the dependent variable. We also present responder analyses using logistic regression models in Table A.2 (KCCQ), Table A.3 (EHFScBS) and Table A.4 (SDMQ) in the Appendix. These responder analyses confirm the null result for the full scale analyses of the KCCQ and SDMQ and show Odds-Ratios larger than one for all thresholds of the EHFScBS. For the threshold values 60, 70 and 90 the p-values of these estimates is below 0.1, confirming the positive results from

Table 2: Main Results - Full Analysis Set

	Primary Endpoints				Secondary Endpoints	
	6MWT	KCCQ	EHFScBS	AHFKT	EQ5D	SDMQ
Main Effects						
Point Estimate	33.11**	-0.83	5.20**	2.04***	-0.01	3.48
98.75% CI	[5.76, 60.46]	[-5.52, 3.86]	[0.67, 9.72]	[1.07, 3.01]	[-0.05, 0.03]	[-5.91, 12.87]
p-value	0.003	0.657	0.004	< 0.001	0.578	0.353
Cohen's d	0.38	-0.06	0.37	0.67	-0.07	0.12
95% CI	[0.14, 0.63]	[-0.3, 0.19]	[0.12, 0.61]	[0.42, 0.92]	[-0.32, 0.18]	[-0.13, 0.37]
Covariates						
Male	0.34	-1.13	-0.04	-1.21**	0.00	-0.18
	[-24.56, 25.25]	[-5.4, 3.14]	[-4.15, 4.07]	[-2.09, -0.34]	[-0.04, 0.04]	[-8.71, 8.36]
Age < 65	26.88*	3.77	-0.69	0.79	0.01	10.53**
	[4.05, 49.71]	[-0.01, 7.55]	[-4.33, 2.95]	[0.01, 1.57]	[-0.02, 0.04]	[2.98, 18.08]
NYHA Class 1 or 2	7.36	1.28	-1.86	0.15	0.06***	4.82
	[-16.44, 31.15]	[-2.88, 5.43]	[-5.55, 1.84]	[-0.64, 0.94]	[0.02, 0.09]	[-2.85, 12.5]
Entry Examination	0.90***	0.77***	0.69***		0.75***	
	[0.8, 1]	[0.68, 0.86]	[0.6, 0.78]		[0.67, 0.84]	

Notes: Asterisks indicate p-values following the Bonferroni adjustment for four primary endpoints with * p < 0.025, ** p < 0.0125, *** p < 0.0025. For covariates: 98.75% confidence intervals in brackets. Endpoints: 6MWT distance in meters walked in six minutes; KCCQ Kansas City Cardiomyopathy Questionnaire summary score; EHFScBS: European Heart Failure Self-care Behaviour Scale summary Score; AHFKT number of correct answers in the Atlanta Heart Failure Knowledge Test; EQ5D EQ-5D-L utility weight based on Ludwig et al., 2018; SDMQ summary score from the 9-item Shared Decision Making Questionnaire. N = 252. Missing values are imputed following a jump to reference approach where all observations are assigned the average values from the control group.

Table 3: Responder Analysis EHFScBS - Full Analysis Set

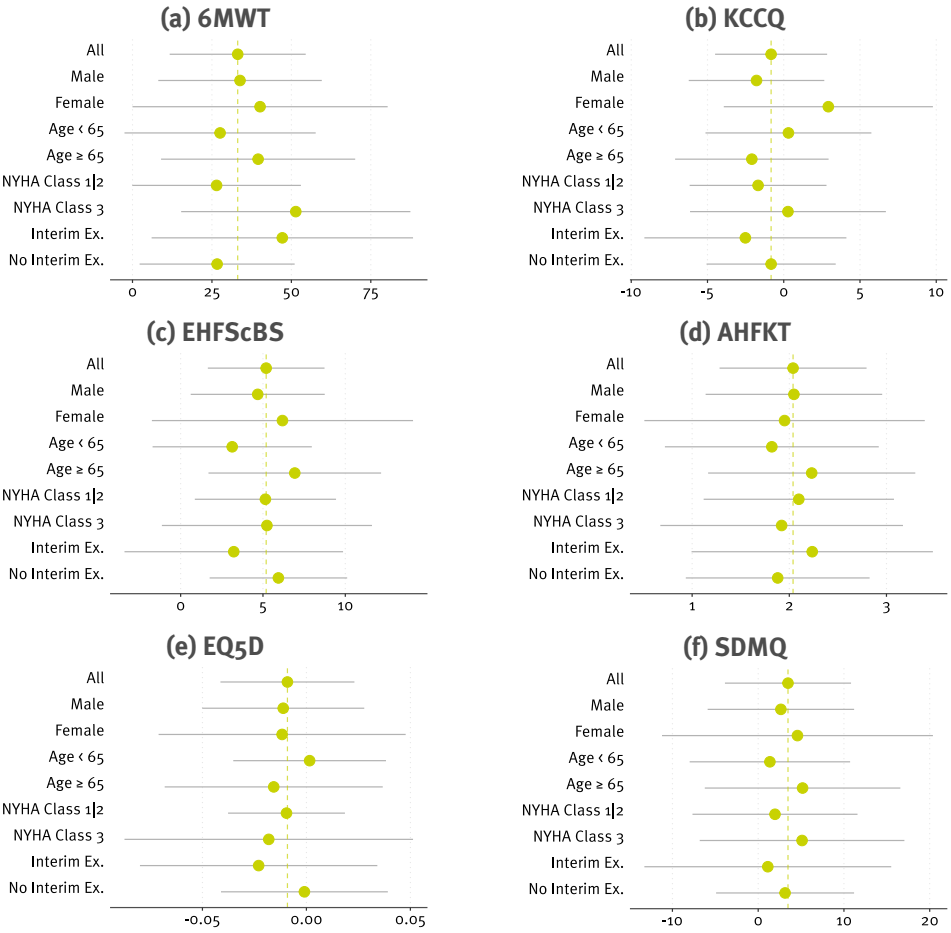
	EHFScBS < 70	EHFScBS ≥ 70
Control	55 (44%)	70 (56%)
Intervention	42 (33%)	85 (67%)

Notes: European Heart Failure Self-care Behaviour Scale summary score ≥ 70 at exit examination by treatment. Chi-Square Test: $x^2 = 2.73$, $p = 0.098$. N = 252. Missing values are imputed following a jump to reference approach where all observations are assigned the average values from the control group.

the full scale analysis as well as the simple responder analysis.

For Figure 3, we split the FAS into subgroups and repeated the main estimation with covariate adjustment. The forest plots contain the main coefficients and the corresponding 95% confidence intervals for each outcome measure across different samples. In each subplot, the first estimate corresponds to the main analysis using the full sample for comparison. Overall, the point estimates are very stable across subsamples. The regression results for the subgroups are also provided in Table A.5, Table A.6, Table A.7, Table A.8, Table A.9, Table A.10, Table A.11 and Table A.12 in the Appendix.

Figure 3: Subgroup Analyses



Notes: Forest plots for main coefficient with covariate adjustment. Dots indicate point estimates from the main regression with the full analysis set for different subsamples. "All" refers to the full sample. Horizontal lines indicate 95% confidence intervals. Endpoints: 6MWT distance in meters walked in six minutes; KCCQ Kansas City Cardiomyopathy Questionnaire summary score; EHFScBS: European Heart Failure Self-care Behaviour Scale summary score; AHFKT number of correct answers in the Atlanta Heart Failure Knowledge Test; EQ5D EQ-5D-L utility weight based on Ludwig et al., 2018; SDMQ summary score from the 9-item Shared Decision Making Questionnaire. N = 252. Missing values are imputed following a jump to reference approach where all observations are assigned the average values from the control group.

4.2 Predictive Mean Matching Imputation

While the jump to reference imputation in the FAS provides a valuable lower bound for the true treatment effect, we use predictive mean matching as an alternative, less conservative way to impute missing values in the exit examination. The results based on this sample presented in Table 4 are qualitatively the same as for the FAS. Quantitatively, the estimated effects are larger for the 6MWT (+48.57m) and the EHFScBS (+5.51) compared to the FAS and the same as in the FAS for AHFKT (+2.04) compared to the FAS. Results without covariate adjustment for the PMM sample are presented in Table A.13 (in the Appendix).

Table 4: Main Results - PMM Imputation

	Primary Endpoints				Secondary Endpoints	
	6MWT	KCCQ	EHFScBS	AHFKT	EQ5D	SDMQ
Main Effects						
Point Estimate	48.57***	1.90	5.51**	2.04***	-0.01	2.50
98.75% CI	[17.98, 79.16]	[-3.71, 7.51]	[0.47, 10.56]	[1.01, 3.08]	[-0.1, 0.07]	[-7.62, 12.62]
p-value	< 0.001	0.395	0.006	< 0.001	0.741	0.535
Cohen's d	0.50	0.11	0.35	0.63	-0.04	0.08
95% CI	[0.26, 0.75]	[-0.14, 0.36]	[0.1, 0.6]	[0.38, 0.88]	[-0.29, 0.21]	[-0.17, 0.33]
Covariates						
Male	-7.57	-2.08	2.03	-1.16*	-0.01	-0.87
	[-35.43, 20.29]	[-7.19, 3.02]	[-2.55, 6.61]	[-2.1, -0.22]	[-0.09, 0.06]	[-10.07, 8.33]
Age < 65	26.51	6.60**	1.45	1.06**	-0.00	12.78***
	[0.97, 52.05]	[2.08, 11.11]	[-2.61, 5.51]	[0.23, 1.89]	[-0.07, 0.06]	[4.64, 20.92]
NYHA Class 1 or 2	5.28	-0.46	-0.82	-0.37	0.05	8.11
	[-21.34, 31.9]	[-5.42, 4.51]	[-4.95, 3.3]	[-1.21, 0.48]	[-0.02, 0.12]	[-0.17, 16.39]
Entry Examination	0.85***	0.58***	0.49***		0.59***	
	[0.74, 0.96]	[0.47, 0.68]	[0.39, 0.58]		[0.42, 0.77]	

Notes: Asterisks indicate p-values following the Bonferroni adjustment for four primary endpoints with * p < 0.025, ** p < 0.0125, *** p < 0.0025. For covariates: 98.75% confidence intervals in brackets. Endpoints: 6MWT distance in meters walked in six minutes; KCCQ Kansas City Cardiomyopathy Questionnaire summary score; EHFScBS: European Heart Failure Self-care Behaviour Scale summary Score; AHFKT number of correct answers in the Atlanta Heart Failure Knowledge Test; EQ5D EQ-5D-L utility weight based on Ludwig et al., 2018; SDMQ summary score from the 9-item Shared Decision Making Questionnaire. N = 252. Missing values are imputed using predictive mean matching.

4.3 Per Protocol Sample

Results from analyzing only patients that had valid responses in the exit estimation (the PPS) are presented in Table 5 are qualitatively the same as for the FAS. Quantitatively, the estimated effect for the 6MWT (+39.94m) is larger compared to the FAS but smaller than the PMM estimate. The coefficient for the EHFScBS (+6.37) as well as for the number of correct AHFKT answers

(+2.48) is larger compared to FAS and PMM. Results without covariate adjustment are presented in Table A.14 (in the Appendix).

Table 5: Main Results - Per Protocol Sample

	Primary Endpoints				Secondary Endpoints	
	6MWT	KCCQ	EHFScBS	AHFKT	EQ5D	SDMQ
Main Effects						
Point Estimate	39.94***	-0.81	6.37**	2.48***	-0.01	3.80
98.75% CI	[7.98, 71.9]	[-6.37, 4.74]	[1.09, 11.65]	[1.35, 3.61]	[-0.06, 0.04]	[-7.43, 15.03]
p-value	0.002	0.712	0.003	< 0.001	0.68	0.395
Cohen's d	0.43	-0.05	0.42	0.77	-0.06	0.12
95% CI	[0.16, 0.7]	[-0.32, 0.22]	[0.15, 0.69]	[0.49, 1.04]	[-0.33, 0.22]	[-0.15, 0.39]
Covariates						
Male	-1.00	-1.69	0.48	-1.50**	0.00	-0.25
	[-29.91, 27.92]	[-6.7, 3.31]	[-4.29, 5.24]	[-2.51, -0.48]	[-0.04, 0.05]	[-10.37, 9.87]
Age < 65	35.61**	4.96	-0.39	0.99	0.01	12.68**
	[8.46, 62.75]	[0.47, 9.45]	[-4.63, 3.86]	[0.08, 1.89]	[-0.03, 0.05]	[3.63, 21.72]
NYHA Class 1 or 2	7.32	1.18	-1.56	0.05	0.06**	5.53
	[-20.16, 34.79]	[-3.69, 6.06]	[-5.89, 2.77]	[-0.88, 0.97]	[0.02, 0.1]	[-3.71, 14.76]
Entry Examination	0.86***	0.71***	0.61***		0.72***	
	[0.73, 0.98]	[0.61, 0.82]	[0.51, 0.72]		[0.62, 0.82]	

Notes: Asterisks indicate p-values following the Bonferroni adjustment for four primary endpoints with * p < 0.025, ** p < 0.0125, *** p < 0.0025. For covariates: 98.75% confidence intervals in brackets. Endpoints: 6MWT distance in meters walked in six minutes; KCCQ Kansas City Cardiomyopathy Questionnaire summary score; EHFScBS: European Heart Failure Self-care Behaviour Scale summary Score; AHFKT number of correct answers in the Atlanta Heart Failure Knowledge Test; EQ5D EQ-5D-L utility weight based on Ludwig et al., 2018; SDMQ summary score from the 9-item Shared Decision Making Questionnaire. N = 214 for 6MWT and N = 211 non-missing observations for the other endpoints where there were non-missing values in the exit examination, out of the original sample of 252.

We can use the PPS also to analyze additional outcomes which are challenging to reasonably impute. In Figure 3, we illustrate how NYHA class assessments evolve between entry and exit examinations for both the intervention and control groups. A summary table is in Table A.15 (in the Appendix). For the 214 patients who attended the exit examination, there was no NYHA assessment for one patient. The NYHA class distribution for the remaining 213 patients is described below.

From the 15 patients who were assigned NYHA class I in the entry examination in the control group, 8 remained in class I, 5 progressed to class II and 2 progressed to class III. From the 18 patients who were assigned NYHA class I in the entry examination in the intervention group, 15 remained in class I, 2 progressed to class II and 1 progressed to class III. This shows that the share of patients progressing to worse NYHA classes is higher in the control group (47%) compared to the intervention group (17%).

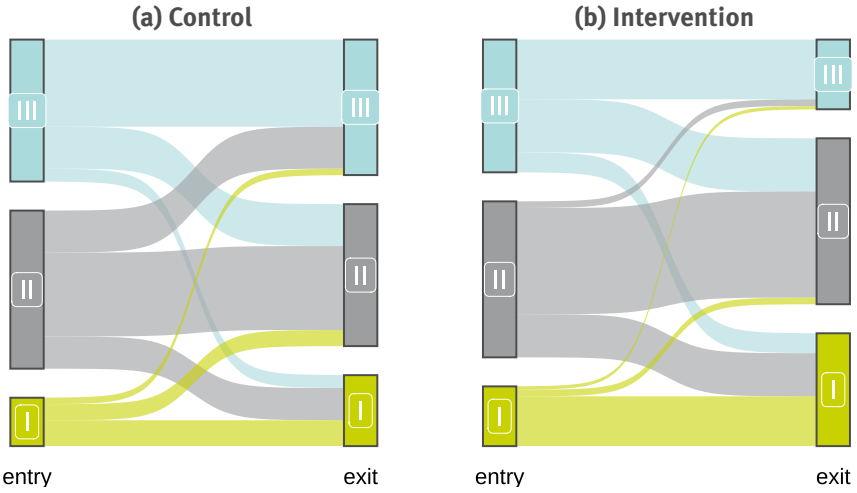
Among patients who were assigned NYHA class II in the entry examination, in the control group 10 improved to class I, 26 remained in class II and 13 progressed to class III. In the intervention group, 13 improved to class I, 32 remained in class II and 2 progressed to class III. This shows that the share of patients progressing to the worse NYHA class is higher in the control group (27%) compared to the intervention group (4%) and also that the share of patients that improved to NYHA class I is higher in the intervention group (27%) compared to the control group (20%).

For patients that were assigned NYHA class III in the entry examination, the control group had 27 patients remain in class III, 13 patients improve to class II and 4 patients improve to class I. The intervention group had 18 patients remain in class III, 16 patients improve to class II and 6 patients improve to class I. These numbers show that the share of patients with improvement was higher in the intervention group compared to the control group both for improvements to class II (40% vs. 30%) as well as improvements to class I (15% vs. 9%).[2]

At the exit examination, patients were also asked to report their healthcare utilization over the preceding four weeks. The results in Table 6 show that patients in the intervention group reported slightly fewer outpatient physician visits (including general practitioners and specialists), with an average of 1.29 visits compared to 1.40 visits in the control group. 8% of patients in the control group and 4% in the intervention group reported a hospital visit in the last four weeks. Among those who did report hospital visits, more were planned rather than emergency visits. Patients were also asked about the number of nights spent in the hospital. However, as only 12 individuals reported hospital visits, the length of stay reflects highly individual circumstances and is not suitable for generalization.

[2] To illustrate NYHA class changes in the FAS, Table A.16 in the Appendix presents results with missing exit values conservatively imputed by assigning patients with missing NYHA examination to one NYHA class worse than the assignment at entry. These results should be interpreted with caution, as patients are only classified into NYHA class IV due to this imputation approach.

Figure 3: Change in NYHA Class



Notes: Change in NYHA Class assignment by study physician between entry and exit examination.

Table 6: Health Care Utilization - Per Protocol Sample

			Hospital							
	Outpatient Visits		Visits		Nights		Planned Visits		Emergency Visits	
Group	Mean	N	Mean	N	Median	N	Mean	N	Mean	N
Control	1.41	106	0.08	106	1.50	8	1.00	8	0.50	8
Intervention	1.29	105	0.04	105	2.00	4	1.25	4	0.75	4

Notes: This table shows the self-reported utilization of health care services in the four weeks prior to the exit examination. N indicates the number or patients providing information. Emergency hospital stays refer to hospital stays via the emergency room or the ambulance.

4.4 Explorative Analyses

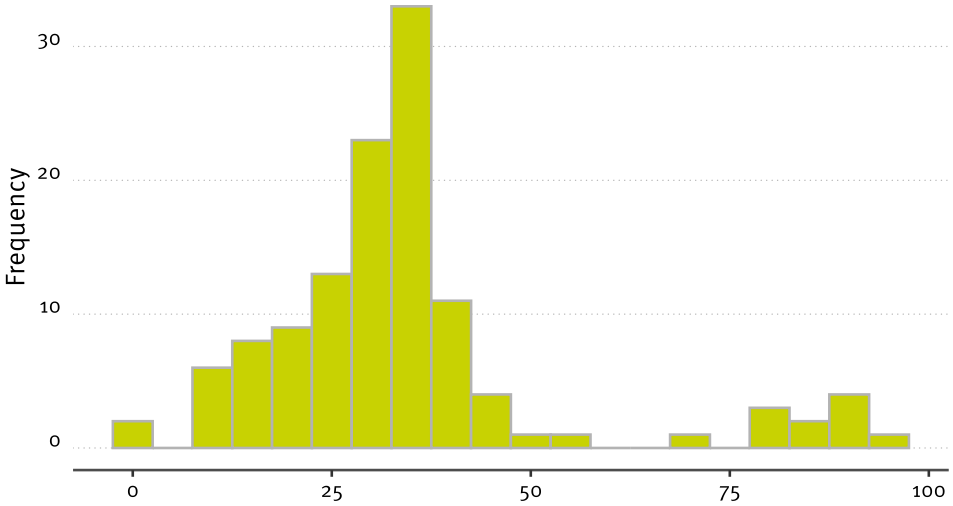
As a last part of the analyses, we explore how intervention intensity and length of the intervention affect outcomes. Figure 4 plots the distribution of average weekly measures taken by the participants during the intervention period. On average, participants used the app to record 34 measures per week (4.8 per day) (a statistical summary is provided in Table A.17 in the Appendix). Two individuals did not use the app to take any measures and the one patient with the highest usage intensity conducted 94 weekly measures (13.4 per day). We use linear regression models to investigate how usage

intensity moderates the change in the endpoints at exit examination in Table 7. One additional average measure per week is associated with a significant 1m longer distance walked in the 6MWT. The correlation between usage intensity and increase in the KCCQ score is close to zero and insignificant. One additional average measure per week is associated with a significant 0.1 point increase in the EHFSCBS score. For AHFKT, EQ5D and SDMQ there is no correlation between usage intensity and change in outcomes at the exit examination.

In addition, we can leverage the interim evaluation that was conducted for the first 100 patients after three months (half of the intervention period) to assess how the effect of the intervention evolves over time. We plot the average values for the outcome measures that were collected on each of the three examinations for those patients where all measures were available in Figure 5. For the distance in the 6MWT, both groups start at a similar level but the increase from entry to midterm is larger for the intervention group. From midterm to exit the distance further increases for the intervention group, albeit at a lower rate. In the control group in contrast the average distance decreases from midterm to exit examination. The KCCQ score evolves very similar for intervention and control group. There is an increase in both time spans but the increase from entry to midterm is stronger than the increase from midterm to exit examination. For the EHFCcBS score, both groups start at a similar level and there is a stronger increase from entry to midterm examination for the intervention group than for the control group. For both groups there is an additional small increase from midterm to exit examination. Average EQ5D weights are slightly higher et entry examination for the control group. There is a marginal increase for the intervention group from entry to midterm examination and a decline back to the original level from midterm to exit examination. While a causal interpretation of these patterns is not possible since the fact that there was an examination might have had an effect on both groups, the patterns for the two endpoints where there is an overall significant treatment effect suggest that using the intervention for longer can increase (6MWT) or sustain (EHFScBS) the treatment effect.

To assess the comparability of patients with and without an interim examination, we compare the patients assigned for an interim evaluation (the first 100 patients that participated in the study) to the remaining patients based on their baseline characteristics at the entry examination (see Table A.18).

Figure 4: Distribution of Measures per Week



Notes: Histogram of app usage intensity. Each bar indicates the frequency of a participant taking on average the number of measures with the app per week. The statistical description of the distribution is presented in Table A.17.

The analysis shows no significant differences between the groups in dropout rates, intervention length, sex distribution, single-household status, smoking status, and BMI based on t-tests for differences in means. However, there is a significant difference in age as well as the Elixhauser comorbidity score at baseline, where patients that are scheduled to participate in the interim examination are older and have better health. Regarding NYHA class, the interim group includes fewer patients in class II and more in class III. For the main endpoint 6MWT, those assigned to the interim examination had a significantly lower distance at entry compared to those that were not. For the other endpoints (KCCQ, EHFScBS, and EQ-5D), there are no significant differences between the groups at baseline.

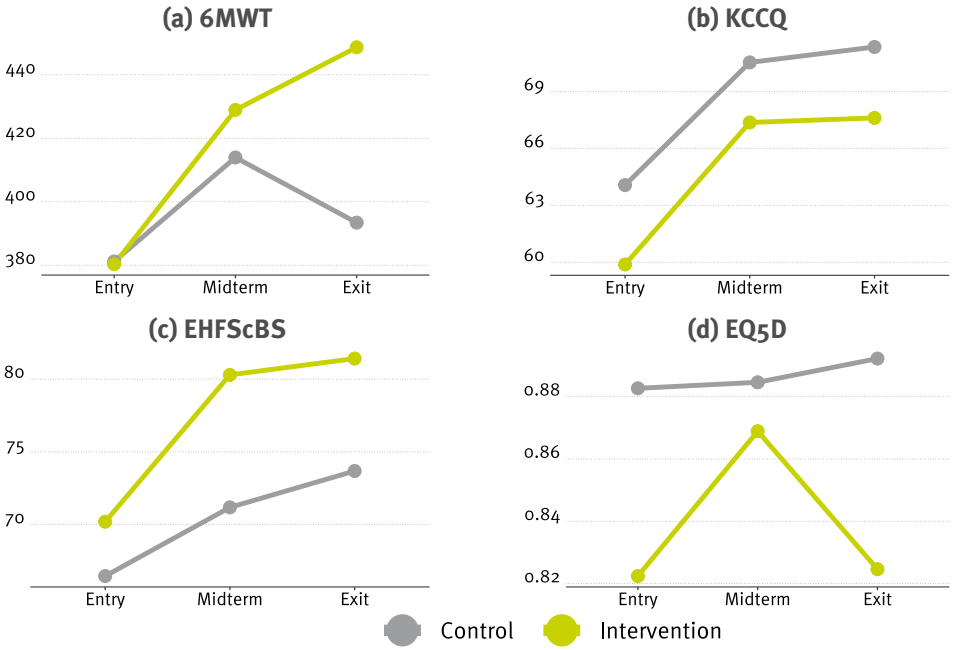
When we repeat the main analyses from Table 2 with an additional control variable that captures whether patients had an interim examination, the estimates for the treatment indicator are quantitatively and qualitatively very similar (see Table A.19 in the Appendix). The coefficient for the interim indicator reveals that patients who underwent the interim examination walked significantly more in the 6MWT, had a higher quality of life by the KCCQ index and answered one more question correctly at the AHFKT in the FAS. Table A.20 and Table A.21 provide also the PMM and the PPS results with the additional interim control.

Table 7: Correlation of Usage Intensity on Outcomes - Full Analysis Set

	Primary Endpoints				Secondary Endpoints	
	6MWT	KCCQ	EHFScBS	AHFKT	EQ5D	SDMQ
Weekly Measures	1.03*	0.05	0.13*	0.02	0.00	-0.04
	(0.44)	(0.08)	(0.06)	(0.01)	(0.00)	(0.15)
Entry Examination Value	0.95***	0.74***	0.67***		0.82***	
	(0.07)	(0.06)	(0.06)		(0.07)	
Intercept	18.00	18.68***	27.17***	18.32***	0.14*	59.49***
	(34.28)	(5.23)	(4.27)	(0.52)	(0.06)	(5.83)

Notes: Each column shows estimates from one linear regression for observations in the intervention group. Main explanatory variable: Average weekly numbers of measures taken using the ProHerz App. Measures is sum of measures taken for heart rate, temperature, blood pressurge and oxygen saturation. Asterisks indicate p-values * p ‹ 0.05, ** p ‹ 0.01, *** p ‹ 0.001. Standard errors in parentheses. Endpoints: 6MWT distance in meters walked in six minutes; KCCQ Kansas City Cardiomyopathy Questionnaire summary score; EHFScBS: European Heart Failure Self-care Behaviour Scale summary Score; AHFKT number of correct answers in the Atlanta Heart Failure Knowledge Test; EQ5D EQ-5D-L utility weight based on Ludwig et al., 2018; SDMQ summary score from the 9-item Shared Decision Making Questionnaire. N = 127. Missing values are imputed following a jump to reference approach where all observations are assigned the average values from the control group.

Figure 5: Evolution of Outcomes over Time



Notes: Evolution of outcomes over time for patients who were included in the interim examination. 6MWT change is difference in distance (in meters walked) in six minutes between entry and exit examination; KCCQ change is difference in Kansas City Cardiomyopathy Questionnaire summary score between entry and exit examination; EHFScBS change is difference in European Heart Failure Self-care Behaviour Scale summary score between entry and exit examination; EQ5D change is difference in EQ-5D-L utility weight based on Ludwig et al., 2018 between entry and exit examination.

5 DISCUSSION

This study is the first randomized controlled evaluation of a smartphone intervention to support HF patients in Germany with a sample size comparable to the largest international studies in the field (only 4 of the 34 reviewed RCTs in Zhu et al. (2024) have larger sample sizes). The results show that the intervention had significant positive effects on patients' health (better NYHA class, longer 6MWT distance) as well as on self-care behavior and health literacy. These results confirm the conclusion from similar interventions that smartphone interventions can indeed have positive health effects.

In addition to the statistical significance of the difference in outcomes, their effect size needs to be discussed. For the 6MWT, previous research suggests that a 14m to 36m increase is the minimal clinically important difference (MCID) among patients with HF (Bohannon & Crouch, 2017; Täger et al., 2014). In this study, the estimates for the increase in the 6MWT range between 33m (FAS) and 49m (PMM). The corresponding effect sizes range between Cohen's d of 0.38 (FAS) and 0.50 (PMM). For all samples the 95% confidence interval for the effect sizes excludes zero. Hence, the treatment effect is above the MCID and the standardized effect size is small to medium, so we can conclude that the intervention had a meaningful effect on patients' health. This interpretation is confirmed by the better NYHA class progression in the intervention group compared to the control group.

For the EHFScBs, the intervention led to a significant increase between 5.2 (FAS) and 6.37 (PPS). The corresponding effect sizes range between Cohen's d of 0.35 (PMM) and 0.42 (PPS). These effect size estimates are above the irrelevance threshold of 0.2 and the corresponding confidence intervals are above zero for all samples. Supported by the higher share of patients at exit above the threshold of 70 suggested by Wagenaar et al. (2017), we can conclude that intervention had a meaningful effect on patients' self-care behavior.

Regarding health literacy, participants in the intervention group gave an additional 2.0 (FAS) to 2.5 (PMM) correct answers in the AHFKT at exit. This higher number corresponds to an effect size of Cohen's d between 0.63 (PMM) and 0.77 (PPS) indicating a medium size effect with all confidence intervals for the effect size above zero. Although the literature has not yet defined MCID values for the AHFKT, the effect size estimates are a strong signal that the intervention improved health literacy.

Similar to other studies, there was no significant effect on quality of life, both measured with the disease-specific KCCQ and with the general EQ5D. In addition, self-care behavior increased, although not significantly. It is up future research to analyze whether smartphone interventions in general struggle to improve these subjective scores or whether the results are due to the sample size and intervention period choices.

6 CONCLUSION

In this study, we present findings from a randomized controlled trial evaluating the effectiveness of the ProHerz app, a digital therapeutic designed for patients with HF. This RCT builds on insights from the preceding pilot study, incorporating a larger sample size, a longer intervention period, and a control group for comparison.

Patients in the intervention group experienced significant and clinically meaningful improvements in their health status measured in the distance walked in 6 minutes as well as better progression in NYHA class. In addition, self-care behaviors and HF-specific health literacy were significantly improved.

By including a control group and extending the duration of the intervention, this RCT provides robust evidence of ProHerz's causal impact on patient relevant outcomes for HF patients. The observed effect sizes are consistent with the outcomes of prior research on digital therapeutics.

REFERENCES

Bennett, J. A., Riegel, B., Bittner, V., & Nichols, J. (2002). Validity and reliability of the nyha classes for measuring research outcomes in patients with cardiac disease. *Heart & Lung*, *31*(4), 262–270. https://doi.org/10.1067/mhl.2002.124554

Bohannon, R. W., & Crouch, R. (2017). Minimal clinically important difference for change in 6-minute walk test distance of adults with pathology: A systematic review. *Journal of Evaluation in Clinical Practice*, *23*(2), 377–381. https://doi.org/10.1111/jep.12629

Cajita, M. I., Gleason, K. T., & Han, H.-R. (2016). A Systematic Review of mHealth-Based Heart Failure Interventions. *Journal of Cardiovascular Nursing*, *31*(3), E10–E22. https://doi.org/10.1097/JCN.0000000000000305

Choi, E.-Y., Park, J.-S., Min, D., Ahn, S., & Ahn, J.-A. (2023). Heart Failure-Smart Life: A randomized controlled trial of a mobile app for self-management in patients with heart failure. *BMC Cardiovascular Disorders*, *23*(1), 7. https://doi.org/10.1186/s12872-023-03039-8

Coorey, G. M., Neubeck, L., Mulley, J., & Redfern, J. (2018). Effectiveness, acceptability and usefulness of mobile applications for cardiovascular disease self-management: Systematic review with meta-synthesis of quantitative and qualitative data. *European Journal of Preventive Cardiology*, *25*(5), 505–521. https://doi.org/10.1177/2047487317750913

Cruz-Cobo, C., Bernal-Jiménez, M. Á., Vázquez-García, R., & Santi-Cano, M. J. (2022). Effectiveness of mHealth Interventions in the Control of Lifestyle and Cardiovascular Risk Factors in Patients After a Coronary Event: Systematic Review and Meta-analysis. *JMIR mHealth and uHealth*, *10*(12), e39593. https://doi.org/10.2196/39593

Davoudi, M., Ghezeljeh, T. N., & Aghouee, F. V. (2020). Effect of a Smartphone-Based App on the Quality of Life of Patients With Heart Failure: Randomized Controlled Trial. *JMIR Nursing*, *3*(1), e20747. https://doi.org/10.2196/20747

Demers, C., McKelvie, R. S., Negassa, A., & Yusuf, S. (2001). Reliability, validity, and responsiveness of the six-minute walk test in patients with heart failure. *American Heart Journal*, *142*(4), 698–703. https://doi.org/10.1067/mhj.2001.118468

Faller, H., Steinbüchel, T., Schowalter, M., Spertus, J. A., Störk, S., & Angermann, C. E. (2005). Der Kansas City Cardiomyopathy Questionnaire (KCCQ) - ein neues krankheitsspezifisches Messinstrument zur Erfassung der Lebensqualität bei chronischer Herzinsuffizienz: Psychometrische Prüfung der deutschen Version. *PPmP - Psychotherapie · Psychosomatik · Medizinische Psychologie*, *55*(3/4), 200–208. https://doi.org/10.1055/s-2004-834597

Gallagher, R., Chow, C. K., Parker, H., Neubeck, L., Celermajer, D. S., Redfern, J., Tofler, G., Buckley, T., Schumacher, T., Hyun, K., Boroumand, F., & Figtree, G. (2023). The effect of a game-based mobile app 'MyHeartMate' to promote lifestyle change in coronary disease patients: A randomized controlled trial. *European Heart Journal - Digital Health*, *4*(1), 33–42. https://doi.org/10.1093/ehjdh/ztac069

Groenewegen, A., Rutten, F. H., Mosterd, A., & Hoes, A. W. (2020). Epidemiology of heart failure. *European Journal of Heart Failure*, *22*(8), 1342–1356. https://doi.org/10.1002/ejhf.1858

Hamine, S., Gerth-Guyette, E., Faulx, D., Green, B. B., & Ginsburg, A. S. (2015). Impact of mHealth Chronic Disease Management on Treatment Adherence and Patient Outcomes: A Systematic Review. *Journal of Medical Internet Research*, *17*(2), e52. https://doi.org/10.2196/jmir.3951

Jin, K., Khonsari, S., Gallagher, R., Gallagher, P., Clark, A. M., Freedman, B., Briffa, T., Bauman, A., Redfern, J., & Neubeck, L. (2019). Telehealth interventions for the secondary prevention of coronary heart disease: A systematic review and meta-analysis. *European Journal of Cardiovascular Nursing*, *18*(4), 260–271. https://doi.org/10.1177/1474515119826510

Johnston, N., Bodegard, J., Jerström, S., Åkesson, J., Brorsson, H., Alfredsson, J., Albertsson, P. A., Karlsson, J.-E., & Varenhorst, C. (2016). Effects of interactive patient smartphone support app on drug adherence and lifestyle changes in myocardial infarction patients: A randomized study. *American Heart Journal*, *178*, 85–94. https://doi.org/10.1016/j.ahj.2016.05.005

Johnston, W., Keogh, A., Dickson, J., Leslie, S. J., Megyesi, P., Connolly, R., Burke, D., & Caulfield, B. (2022). Human-Centered Design of a Digital Health Tool to Promote Effective Self-care in Patients With Heart Failure: Mixed Methods Study. *JMIR formative research*, *6*(5), e34257. https://doi.org/10.2196/34257

Köberich, S., Glattacker, M., Jaarsma, T., Lohrmann, C., & Dassen, T. (2013). Validity and reliability of the German version of the 9-item European Heart Failure Self-care Behaviour Scale. *European Journal of Cardiovascular Nursing*, *12*(2), 150–158. https://doi.org/10.1177/1474515112438639

Kriston, L., Scholl, I., Hölzel, L., Simon, D., Loh, A., & Härter, M. (2010). The 9-item Shared Decision Making Questionnaire (SDM-Q-9). Development and psychometric properties in a primary care sample. *Patient Education and Counseling*, *80*(1), 94–99. https://doi.org/10.1016/j.pec.2009.09.034

Lee, H., Boo, S., Yu, J., Suh, S.-R., Chun, K. J., & Kim, J. H. (2017). Physical Functioning, Physical Activity, Exercise Self-Efficacy, and Quality of Life Among Individuals With Chronic Heart Failure in Korea: A Cross-

Sectional Descriptive Study. *Journal of Nursing Research*, *25*(2), 131–139. https://doi.org/10.1097/JNR.0000000000000150

Lloyd-Jones, D. M., Allen, N. B., Anderson, C. A., Black, T., Brewer, L. C., Foraker, R. E., Grandner, M. A., Lavretsky, H., Perak, A. M., Sharma, G., Rosamond, W., & on behalf of the American Heart Association. (2022). Life's Essential 8: Updating and Enhancing the American Heart Association's Construct of Cardiovascular Health: A Presidential Advisory From the American Heart Association. *Circulation*, *146*(5). https://doi.org/10.1161/CIR.0000000000001078

Ludwig, K., Graf von der Schulenburg, J.-M., & Greiner, W. (2018). German Value Set for the EQ-5D-5L. *PharmacoEconomics*, *36*(6), 663–674. https://doi.org/10.1007/s40273-018-0615-8

Mok, V. K. F., Sit, J. W. H., Tsang, A. S. M., Chair, S. Y., Cheng, T. L., & Chiang, C.-s. (2013). A Controlled Trial of a Nurse Follow-up Dietary Intervention on Maintaining a Heart-Healthy Dietary Pattern Among Patients After Myocardial Infarction. *Journal of Cardiovascular Nursing*, *28*(3), 256–266. https://doi.org/10.1097/JCN.0b013e31824a37b7

Morris, T. P., White, I. R., & Royston, P. (2014). Tuning multiple imputation by predictive mean matching and local residual draws. *BMC Medical Research Methodology*, *14*(1), 75. https://doi.org/10.1186/1471-2288-14-75

Ni, Z., Wu, B., Yang, Q., Yan, L. L., Liu, C., & Shaw, R. J. (2022). An mHealth Intervention to Improve Medication Adherence and Health Outcomes Among Patients With Coronary Heart Disease: Randomized Controlled Trial. *Journal of Medical Internet Research*, *24*(3), e27202. https://doi.org/10.2196/27202

Redfern, J., Coorey, G., Mulley, J., Scaria, A., Neubeck, L., Hafiz, N., Pitt, C., Weir, K., Forbes, J., Parker, S., Bampi, F., Coenen, A., Enright, G., Wong, A., Nguyen, T., Harris, M., Zwar, N., Chow, C. K., Rodgers, A., … Peiris, D. (2020). A digital health intervention for cardiovascular disease management in primary care (CONNECT) randomized controlled trial. *npj Digital Medicine*, *3*(1), 117. https://doi.org/10.1038/s41746-020-00325-z

Reif, S., Schubert, S., Stiefel, J., Husri, F., Fischlein, T., Pauschinger, M., & Klucken, J. (2022). Supporting patients with heart failure with digital

therapeutics—A pilot study in Germany. *DIGITAL HEALTH*, *8*, 205520762211438. https://doi.org/10.1177/20552076221143899

Reilly, C. M., Higgins, M., Smith, A., Gary, R. A., Robinson, J., Clark, P. C., McCarty, F., & Dunbar, S. B. (2009). Development, Psychometric Testing, and Revision of the Atlanta Heart Failure Knowledge Test. *Journal of Cardiovascular Nursing*, *24*(6), 500–509. https://doi.org/10.1097/JCN.0b013e3181affob0

Riegel, B., Moser, D. K., Anker, S. D., Appel, L. J., Dunbar, S. B., Grady, K. L., Gurvitz, M. Z., Havranek, E. P., Lee, C. S., Lindenfeld, J., Peterson, P. N., Pressler, S. J., Schocken, D. D., & Whellan, D. J. (2009). State of the Science: Promoting Self-Care in Persons With Heart Failure: A Scientific Statement From the American Heart Association. *Circulation*, *120*(12), 1141–1163. https://doi.org/10.1161/CIRCULATIONAHA.109.192628

Robinson, A., Husband, A. K., Slight, R. D., & Slight, S. P. (2021). Digital technology to support lifestyle and health behaviour changes in surgical patients: Systematic review. *BJS Open*, *5*(2), zraa009. https://doi.org/10.1093/bjsopen/zraa009

Roth, G. A., Forouzanfar, M. H., Moran, A. E., Barber, R., Nguyen, G., Feigin, V. L., Naghavi, M., Mensah, G. A., & Murray, C. J. (2015). Demographic and Epidemiologic Drivers of Global Cardiovascular Mortality. *New England Journal of Medicine*, *372*(14), 1333–1341. https://doi.org/10.1056/NEJMoa1406656

Saleh, Z. T., Elshatarat, R. A., Elhefnawy, K. A., Helmi Elneblawi, N., Abu Raddaha, A. H., Al-Za'areer, M. S., Mofdy Almarwani, A., Alzahrani, N. S., Aqel, A. A., Shawashi, T. O., & Tayeh, M. (2023). Effect of a Home-Based Mobile Health App Intervention on Physical Activity Levels in Patients With Heart Failure: A Randomized Controlled Trial. *Journal of Cardiovascular Nursing*, *38*(2), 128–139. https://doi.org/10.1097/JCN.0000000000000911

Savarese, G., & Lund, L. H. (2017). Global Public Health Burden of Heart Failure. *Cardiac Failure Review*, *3*(1), 7–11. https://doi.org/10.15420/cfr.2016:25:2

Spertus, J. A., & Jones, P. G. (2015). Development and Validation of a Short Version of the Kansas City Cardiomyopathy Questionnaire. *Circula-*

tion: Cardiovascular Quality and Outcomes, *8*(5), 469–476. https://doi.org/10.1161/CIRCOUTCOMES.115.001958

Täger, T., Hanholz, W., Cebola, R., Fröhlich, H., Franke, J., Doesch, A., Katus, H. A., Wians, F. H., & Frankenstein, L. (2014). Minimal important difference for 6-minute walk test distances among patients with chronic heart failure. *International Journal of Cardiology*, *176*(1), 94–98. https://doi.org/10.1016/j.ijcard.2014.06.035

van Walraven, C., Austin, P. C., Jennings, A., Quan, H., & Forster, A. J. (2009). A Modification of the Elixhauser Comorbidity Measures into a Point System for Hospital Death Using Administrative Data. *Medical Care*, *47*(6), 626–633.

Varnfield, M., Karunanithi, M., Lee, C.-K., Honeyman, E., Arnold, D., Ding, H., Smith, C., & Walters, D. L. (2014). Smartphone-based home care model improved use of cardiac rehabilitation in postmyocardial infarction patients: Results from a randomised controlled trial. *Heart*, *100*(22), 1770–1779. https://doi.org/10.1136/heartjnl-2014-305783

Vellone, E., Jaarsma, T., Strömberg, A., Fida, R., Årestedt, K., Rocco, G., Cocchieri, A., & Alvaro, R. (2014). The European Heart Failure Self-care Behaviour Scale: New insights into factorial structure, reliability, precision and scoring procedure. *Patient Education and Counseling*, *94*(1), 97–102. https://doi.org/10.1016/j.pec.2013.09.014

Victoria-Castro, A. M., Martin, M. L., Yamamoto, Y., Melchinger, H., Weinstein, J., Nguyen, A., Lee, K. A., Gerber, B., Calderon, F., Subair, L., Lee, V., Williams, A., Shaw, M., Arora, T., Garcez, A., Desai, N. R., Ahmad, T., & Wilson, F. P. (2024). Impact of Digital Health Technology on Quality of Life in Patients With Heart Failure. *JACC: Heart Failure*, *12*(2), 336–348. https://doi.org/10.1016/j.jchf.2023.09.022

Wagenaar, K., Broekhuizen, B., Rutten, F., Strömberg, A., Van Stel, H., Hoes, A. W., & Jaarsma, T. (2017). Interpretability of the European Heart Failure Self-care Behaviour scale. *Patient Preference and Adherence*, *Volume 11*, 1841–1849. https://doi.org/10.2147/PPA.S144915

Widmer, R. J., Allison, T. G., Lerman, L. O., & Lerman, A. (2015). Digital Health Intervention as an Adjunct to Cardiac Rehabilitation Reduces Cardiovascular Risk Factors and Rehospitalizations. *Journal of Cardiovascular Translational Research*, *8*(5), 283–292. https://doi.org/10.1007/s12265-015-9629-1

Wood, D., Kotseva, K., Connolly, S., Jennings, C., Mead, A., Jones, J., Holden, A., De Bacquer, D., Collier, T., De Backer, G., & Faergeman, O. (2008). Nurse-coordinated multidisciplinary, family-based cardiovascular disease prevention programme (EUROACTION) for patients with coronary heart disease and asymptomatic individuals at high risk of cardiovascular disease: A paired, cluster-randomised controlled trial. *The Lancet*, *371*(9629), 1999–2012. https://doi.org/10.1016/S0140-6736(08)60868-5

Zhu, Y., Zhao, Y., & Wu, Y. (2024). Effectiveness of mobile health applications on clinical outcomes and health behaviors in patients with coronary heart disease: A systematic review and meta-analysis. *International Journal of Nursing Sciences*, *11*(2), 258–275. https://doi.org/10.1016/j.ijnss.2024.03.012

APPENDIX

Ethical approval

The ethics committee of the Bavarian State Medical Association (BLÄK) approved this study (REC number: 21073).

Pre-registration

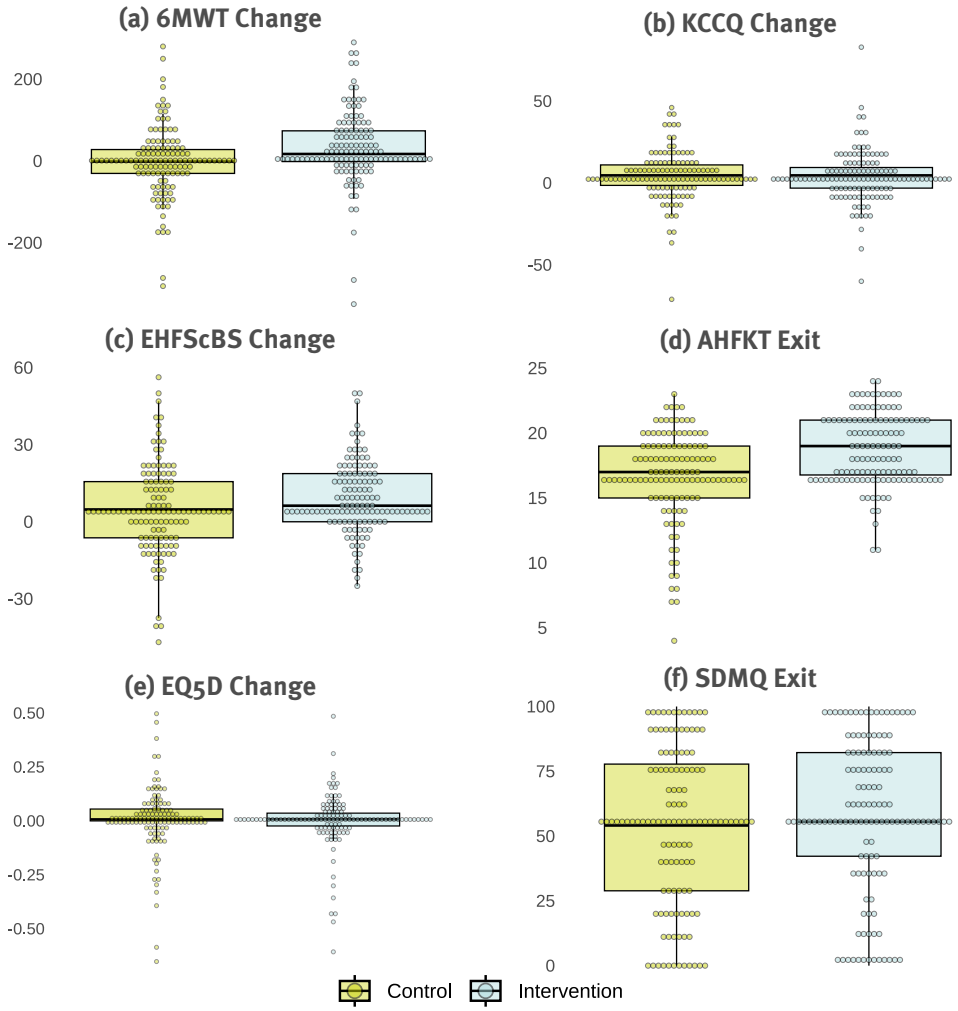
This study was registered at the German Clinical Trials Register under ID DRKS00027949.

Acknowledgments

We would like to thank Georg Zweyer for implementing the e-CRF tool. Kristina Kellersmann provided excellent research assistance.

We are grateful for the work of the study nurses and physicians who conducted the data collection. Albertinen Krankenhaus: Prof. Dr. med. Herbert Nägele (Lead), Eva Cramer-Frommann. Herz Lunge Klinik Bamberg: Dr. med. Johannes Haas (Lead). Klinikum Fulda: Prof. Dr. med. Volker Schächinger (Lead), Dr. med. Margit Niethammer, MBA. Universitätsklinikum Greifswald: Prof. Dr. med. Marcus Dörr (Lead), PD Bishwas Chamling. SLK-Kliniken Heilbronn: Prof. Dr. med. Marcus Hennersdorf (Lead), Dr. med. Tilman Schmidt, Dr. med. Saskia Stegmülle. LMU Universitätsklinikum: Priv.-Doz. Dr. med. Ulrich Grabmaier (Lead), Dr. med. Wolf-Stephan Rudi. Universitätsklinikum Ruhr-Universität Bochum: Dr. med. Daniel Dumitrescu (Lead), Dr. med. Denise Guckel, Dr. med. Andrea Graw. Klinikum Nürnberg: Prof. Dr. med. Matthias Pauschinger (Lead), Prof. Dr. med. Theodor Fischlein, Fahed Husri. Universitätsklinikum Essen: Prof. Dr. Tienush Rassaf (Lead), Univ.-Prof. Dr. med. A. A. Mahabadi, Dr. med. Lars Michel, Dr. med. Florian Schindhelm.

Figure A.1: Differences in Outcome Measures by Group



Notes: Boxplots for main outcomes. 6MWT change is difference in distance (in meters walked) in six minutes between entry and exit examination; KCCQ change is difference in Kansas City Cardiomyopathy Questionnaire summary score between entry and exit examination; EHFScBS change is difference in European Heart Failure Self-care Behaviour Scale summary score between entry and exit examination; AHFKT Exit is number of correct answers in the Atlanta Heart Failure Knowledge Test in the exit examination; EQ5D change is difference in EQ-5D-L utility weight based on Ludwig et al., 2018 between entry and exit examination; SDMQ Exit is summary score from the 9-item Shared Decision Making Questionnaire in exit examination. Total N = 252. Missing values are imputed following a jump to reference approach where all observations are assigned the average values from the control group.

Table A.1: Results without Covariate Adjustment - Full Analysis Set

	Primary Endpoints				Secondary Endpoints	
	6MWT	KCCQ	EHFScBS	AHFKT	EQ5D	SDMQ
Main Effect						
Point Estimate	33.28**	-0.75	5.15**	2.02***	-0.01	3.84
98.75% CI	[5.79, 60.77]	[-5.45, 3.96]	[0.65, 9.66]	[1.04, 3.00]	[-0.05, 0.03]	[-5.69, 13.37]
p-value	0.003	0.690	0.004	‹ 0.001	0.613	0.312
Cohen's d	0.38	-0.05	0.28	0.65	-0.07	0.13
95% CI	[0.13, 0.62]	[-0.30, 0.20]	[0.03, 0.53]	[0.40, 0.91]	[-0.31, 0.18]	[-0.12, 0.37]
Differences						
Diff. Int.	31.13	3.87	9.17		0	
	(88.96)	(15.85)	(13.59)		(0.13)	
Diff. Cont.	-1.39	4.67	4.78		0.01	
	(84.11)	(15.23)	(17.22)		(0.15)	
Mean Values						
Entry Int.	403.02	63.16	67.54		0.84	
	(120.51)	(22.06)	(19.73)		(0.18)	
Exit Int.	434.15	67.03	76.71	18.8	0.84	57.93
	(143.38)	(22.3)	(18.51)	(2.7)	(0.2)	(29.89)
Entry Cont.	387.57	62.88	65.08		0.84	
	(124.32)	(23.05)	(21.35)		(0.21)	
Exit Cont.	386.19	67.55	69.85	16.77	0.84	54.09
	(146.55)	(24.08)	(21.57)	(3.46)	(0.21)	(30.23)

Notes: Each column shows estimates from one linear regression. Asterisks indicate p-values following the Bonferroni adjustment for four primary endpoints with * p ‹ 0.025, ** p ‹ 0.0125, *** p ‹ 0.0025. Standard deviations in parenthesis. Endpoints: 6MWT distance in meters walked in six minutes; KCCQ Kansas City Cardiomyopathy Questionnaire summary score; EHFScBS: European Heart Failure Self-care Behaviour Scale summary score; AHFKT number of correct answers in the Atlanta Heart Failure Knowledge Test; EQ5D EQ-5D-L utility weight based on Ludwig et al., 2018; SDMQ summary score from the 9-item Shared Decision Making Questionnaire. N = 252. Missing values are imputed following a jump to reference approach where all observations are assigned the average values from the control group.

Table A.2: Responder Analyses KCCQ - Full Analysis Set

	KCCQ ≥ 60	KCCQ ≥ 70	KCCQ ≥ 80	KCCQ ≥ 90
Intervention	0.9888	0.6972	0.8408	1.0868
	[0.3891, 2.5006]	[0.2902, 1.6433]	[0.3559, 1.9778]	[0.3627, 3.2710]
	(0.9756)	(0.2953)	(0.6111)	(0.8485)

Notes: Odds ratios from logistic regression models for Kansas City Cardiomyopathy Questionnaire summary score above threshold. Control variables: 98.75% confidence interval in brackets. p-values in parenthesis. Asterisks indicate p-values with * p ‹ 0.025, ** p ‹ 0.0125, *** p ‹ 0.0025. N = 252. Missing values are imputed following a jump to reference approach where all observations are assigned the average values from the control group.

Table A.3: Responder Analyses EHFScBS - Full Analysis Set

	EHFScBS ≥ 60	EHFScBS ≥ 70	EHFScBS ≥ 80	EHFScBS ≥ 90
Intervention	2.9928** [1.1967, 8.0114] (0.0037)	1.6902 [0.7664, 3.8048] (0.0998)	1.3054 [0.6126, 2.8028] (0.3791)	3.4060*** [1.3639, 9.1454] (0.0012)

Notes: Odds ratios from logistic regression models for European Heart Failure Self-care Behaviour Scale summary score above threshold. Control variables: 98.75% confidence interval in brackets. p-values in parenthesis. Asterisks indicate p-values with * p ‹ 0.025, ** p ‹ 0.0125, *** p ‹ 0.0025. N = 252. Missing values are imputed following a jump to reference approach where all observations are assigned the average values from the control group.

Table A.4: Responder Analyses SDMQ - Full Analysis Set

	SDMQ ≥ 60	SDMQ ≥ 70	SDMQ ≥ 80	SDMQ ≥ 90
Intervention	1.3412 [0.7061, 2.5616] (0.2540)	1.1952 [0.6137, 2.3381] (0.5042)	1.2543 [0.6090, 2.6068] (0.4341)	1.0287 [0.4335, 2.4503] (0.9343)

Notes: Odds ratios from logistic regression models for 9-item Shared Decision Making Questionnaire summary score above threshold. Control variables: 98.75% confidence interval in brackets. p-values in parenthesis. Asterisks indicate p-values with * p ‹ 0.025, ** p ‹ 0.0125, *** p ‹ 0.0025. N = 252. Missing values are imputed following a jump to reference approach where all observations are assigned the average values from the control group.

Table A.5: Results Full Analysis Set - Subsample Age ‹ 65 - N = 121

	Primary Endpoints				Secondary Endpoints	
	6MWT	KCCQ	EHFScBS	AHFKT	EQ5D	SDMQ
Main Effects						
Point Estimate	27.56	0.31	3.13	1.82***	0.00	1.37
98.75% CI	[-10.98, 66.11]	[-6.66, 7.28]	[-3.06, 9.33]	[0.41, 3.23]	[-0.05, 0.05]	[-10.62, 13.35]
p-value	0.072	0.911	0.202	0.001	0.933	0.773
Cohen's d	0.33	0.02	0.23	0.60	0.02	0.05
95% CI	[-0.03, 0.69]	[-0.34, 0.38]	[-0.13, 0.6]	[0.24, 0.96]	[-0.35, 0.38]	[-0.31, 0.41]
Covariates						
Male	0.37	-0.57	-1.77	-1.28	0.02	-2.55
	[-35.16, 35.89]	[-7.08, 5.94]	[-7.48, 3.94]	[-2.58, 0.03]	[-0.02, 0.06]	[-13.63, 8.52]
Age ‹ 65						
NYHA Class 1 or 2	18.62	-2.26	-1.03	-0.01	0.02	3.08
	[-16.46, 53.71]	[-8.73, 4.22]	[-6.34, 4.28]	[-1.23, 1.2]	[-0.02, 0.06]	[-7.22, 13.39]
Entry Examination	0.79***	0.73***	0.72***		0.83***	
	[0.66, 0.92]	[0.6, 0.87]	[0.6, 0.84]		[0.71, 0.96]	

Notes: Asterisks indicate p-values following the Bonferroni adjustment for four primary endpoints with * p ‹ 0.025, ** p ‹ 0.0125, *** p ‹ 0.0025. For covariates: 98.75% confidence intervals in brackets. Endpoints: 6MWT distance in meters walked in six minutes; KCCQ Kansas City Cardiomyopathy Questionnaire summary score; EHFScBS: European Heart Failure Self-care Behaviour Scale summary Score; AHFKT number of correct answers in the Atlanta Heart Failure Knowledge Test; EQ5D EQ-5D-L utility weight based on Ludwig et al., 2018; SDMQ summary score from the 9-item Shared Decision Making Questionnaire. Missing values are imputed following a jump to reference approach where all observations are assigned the average values from the control group.

Table A.6: Results Full Analysis Set - Subsample Age ≥ 65 - N = 131

	Primary Endpoints				Secondary Endpoints	
	6MWT	KCCQ	EHFScBS	AHFKT	EQ5D	SDMQ
Main Effects						
Point Estimate	39.54**	-2.09	6.93**	2.23***	-0.02	5.18
98.75% CI	[0.42, 78.65]	[-8.54, 4.35]	[0.22, 13.64]	[0.87, 3.6]	[-0.08, 0.05]	[-9.42, 19.78]
p-value	0.012	0.412	0.01	‹ 0.001	0.554	0.371
Cohen's d	0.45	-0.14	0.46	0.73	-0.10	0.16
95% CI	[0.1, 0.8]	[-0.49, 0.2]	[0.11, 0.81]	[0.38, 1.07]	[-0.45, 0.24]	[-0.19, 0.51]
Covariates						
Male	-3.94	-1.16	1.51	-1.19	-0.01	1.66
	[-39.1, 31.22]	[-6.9, 4.58]	[-4.52, 7.55]	[-2.4, 0.03]	[-0.07, 0.05]	[-11.34, 14.65]
Age ‹ 65						
NYHA Class 1 or 2	-4.01	3.98	-2.58	0.27	0.08***	6.07
	[-36.61, 28.59]	[-1.45, 9.41]	[-7.8, 2.64]	[-0.79, 1.33]	[0.03, 0.14]	[-5.29, 17.43]
Entry Examination	1.03***	0.80***	0.65***		0.72***	
	[0.88, 1.18]	[0.68, 0.92]	[0.52, 0.79]		[0.6, 0.84]	

Notes: Asterisks indicate p-values following the Bonferroni adjustment for four primary endpoints with * p ‹ 0.025, ** p ‹ 0.0125, *** p ‹ 0.0025. For covariates: 98.75% confidence intervals in brackets. Endpoints: 6MWT distance in meters walked in six minutes; KCCQ Kansas City Cardiomyopathy Questionnaire summary score; EHFScBS: European Heart Failure Self-care Behaviour Scale summary Score; AHFKT number of correct answers in the Atlanta Heart Failure Knowledge Test; EQ5D EQ-5D-L utility weight based on Ludwig et al., 2018; SDMQ summary score from the 9-item Shared Decision Making Questionnaire. Missing values are imputed following a jump to reference approach where all observations are assigned the average values from the control group.

Table A.7: Results Full Analysis Set - Subsample Men - N = 190

	Primary Endpoints				Secondary Endpoints	
	6MWT	KCCQ	EHFScBS	AHFKT	EQ5D	SDMQ
Main Effects						
Point Estimate	33.84**	-1.78	4.68*	2.05***	-0.01	2.65
98.75% CI	[0.95, 66.73]	[-7.47, 3.9]	[-0.52, 9.89]	[0.89, 3.21]	[-0.06, 0.04]	[-8.27, 13.56]
p-value	0.01	0.429	0.024	‹ 0.001	0.573	0.542
Cohen's d	0.38	-0.12	0.33	0.65	-0.08	0.09
95% CI	[0.09, 0.67]	[-0.4, 0.17]	[0.04, 0.62]	[0.36, 0.93]	[-0.37, 0.21]	[-0.2, 0.38]
Covariates						
Male						
Age ‹ 65	28.71	3.89	-1.64	0.76	0.02	9.20
	[1.59, 55.83]	[-0.68, 8.45]	[-5.83, 2.55]	[-0.17, 1.69]	[-0.02, 0.06]	[0.46, 17.94]
NYHA Class 1 or 2	-2.06	0.94	-1.64	0.20	0.07***	6.05
	[-30.86, 26.74]	[-4.24, 6.12]	[-5.9, 2.62]	[-0.75, 1.15]	[0.03, 0.11]	[-2.88, 14.98]
Entry Examination	0.90***	0.78***	0.67***		0.76***	
	[0.79, 1.02]	[0.67, 0.89]	[0.57, 0.77]		[0.65, 0.87]	

Notes: Asterisks indicate p-values following the Bonferroni adjustment for four primary endpoints with * p ‹ 0.025, ** p ‹ 0.0125, *** p ‹ 0.0025. For covariates: 98.75% confidence intervals in brackets. Endpoints: 6MWT distance in meters walked in six minutes; KCCQ Kansas City Cardiomyopathy Questionnaire summary score; EHFScBS: European Heart Failure Self-care Behaviour Scale summary Score; AHFKT number of correct answers in the Atlanta Heart Failure Knowledge Test; EQ5D EQ-5D-L utility weight based on Ludwig et al., 2018; SDMQ summary score from the 9-item Shared Decision Making Questionnaire. Missing values are imputed following a jump to reference approach where all observations are assigned the average values from the control group.

Table A.8: Results Full Analysis Set - Subsample Women - N = 62

	Primary Endpoints				Secondary Endpoints	
	6MWT	KCCQ	EHFScBS	AHFKT	EQ5D	SDMQ
Main Effects						
Point Estimate	40.15	2.93	6.18	1.95**	-0.01	4.58
98.75% CI	[-11.57, 91.88]	[-5.91, 11.76]	[-4.03, 16.4]	[0.09, 3.81]	[-0.09, 0.06]	[-15.76, 24.92]
p-value	0.05	0.396	0.124	0.009	0.696	0.564
Cohen's d	0.53	0.23	0.42	0.72	-0.10	0.15
95% CI	[0, 1.07]	[-0.31, 0.76]	[-0.12, 0.95]	[0.19, 1.25]	[-0.64, 0.43]	[-0.38, 0.68]
Covariates						
Male						
Age ‹ 65	19.09	2.44	1.54	0.91	-0.00	14.34
	[-24.11, 62.28]	[-4.4, 9.28]	[-6.37, 9.45]	[-0.53, 2.36]	[-0.06, 0.06]	[-1.49, 30.17]
NYHA Class 1 or 2	36.99	3.18	-2.11	-0.01	0.02	1.63
	[-6.53, 80.51]	[-4.03, 10.39]	[-10.24, 6.02]	[-1.5, 1.48]	[-0.04, 0.08]	[-14.7, 17.96]
Entry Examination	0.92***	0.72***	0.74***		0.74***	
	[0.72, 1.12]	[0.56, 0.88]	[0.54, 0.93]		[0.61, 0.86]	

Notes: Asterisks indicate p-values following the Bonferroni adjustment for four primary endpoints with * p ‹ 0.025, ** p ‹ 0.0125, *** p ‹ 0.0025. For covariates: 98.75% confidence intervals in brackets. Endpoints: 6MWT distance in meters walked in six minutes; KCCQ Kansas City Cardiomyopathy Questionnaire summary score; EHFScBS: European Heart Failure Self-care Behaviour Scale summary Score; AHFKT number of correct answers in the Atlanta Heart Failure Knowledge Test; EQ5D EQ-5D-L utility weight based on Ludwig et al., 2018; SDMQ summary score from the 9-item Shared Decision Making Questionnaire. Missing values are imputed following a jump to reference approach where all observations are assigned the average values from the control group.

Table A.9: Results Full Analysis Set - Subsample NYHA Class 1 | 2 - N = 150

	Primary Endpoints				Secondary Endpoints	
	6MWT	KCCQ	EHFScBS	AHFKT	EQ5D	SDMQ
Main Effects						
Point Estimate	26.46	-1.68	5.15*	2.10***	-0.01	1.96
98.75% CI	[-7.49, 60.42]	[-7.42, 4.06]	[-0.34, 10.64]	[0.84, 3.35]	[-0.05, 0.03]	[-10.35, 14.27]
p-value	0.051	0.46	0.019	‹ 0.001	0.502	0.688
Cohen's d	0.33	-0.12	0.39	0.70	-0.11	0.07
95% CI	[0, 0.65]	[-0.45, 0.2]	[0.07, 0.72]	[0.37, 1.03]	[-0.44, 0.22]	[-0.26, 0.39]
Covariates						
Male	-10.37	-0.84	-0.06	-1.16	0.03	1.52
	[-41.19, 20.46]	[-6.11, 4.44]	[-5, 4.88]	[-2.29, -0.02]	[-0.01, 0.06]	[-9.64, 12.69]
Age ‹ 65	35.49*	1.00	-0.10	0.68	-0.01	9.36
	[7.72, 63.25]	[-3.49, 5.5]	[-4.37, 4.17]	[-0.3, 1.66]	[-0.04, 0.02]	[-0.26, 18.99]
NYHA Class 1 or 2						
Entry Examination	0.80***	0.74***	0.70***		0.74***	
	[0.67, 0.93]	[0.63, 0.85]	[0.6, 0.8]		[0.66, 0.81]	

Notes: Asterisks indicate p-values following the Bonferroni adjustment for four primary endpoints with * p ‹ 0.025, ** p ‹ 0.0125, *** p ‹ 0.0025. For covariates: 98.75% confidence intervals in brackets. Endpoints: 6MWT distance in meters walked in six minutes; KCCQ Kansas City Cardiomyopathy Questionnaire summary score; EHFScBS: European Heart Failure Self-care Behaviour Scale summary Score; AHFKT number of correct answers in the Atlanta Heart Failure Knowledge Test; EQ5D EQ-5D-L utility weight based on Ludwig et al., 2018; SDMQ summary score from the 9-item Shared Decision Making Questionnaire. Missing values are imputed following a jump to reference approach where all observations are assigned the average values from the control group.

Table A.10: Results Full Analysis Set - Subsample NYHA Class 3 - N = 102

	Primary Endpoints				Secondary Endpoints	
	6MWT	KCCQ	EHFScBS	AHFKT	EQ5D	SDMQ
Main Effects						
Point Estimate	51.39**	0.28	5.24	1.92**	-0.02	5.12
98.75% CI	[5.13, 97.64]	[-7.96, 8.51]	[-2.94, 13.42]	[0.32, 3.52]	[-0.11, 0.07]	[-10.18, 20.42]
p-value	0.006	0.932	0.106	0.003	0.606	0.397
Cohen's d	0.56	0.02	0.33	0.61	-0.10	0.17
95% CI	[0.17, 0.96]	[-0.38, 0.41]	[-0.07, 0.72]	[0.21, 1.01]	[-0.5, 0.29]	[-0.23, 0.57]
Covariates						
Male	23.30	-0.57	0.02	-1.33	-0.03	-2.07
	[-18.6, 65.19]	[-8.08, 6.93]	[-7.4, 7.45]	[-2.78, 0.12]	[-0.11, 0.05]	[-15.92, 11.78]
Age ‹ 65	15.53	8.13*	-1.64	0.98	0.05	12.27
	[-23.32, 54.38]	[1.4, 14.86]	[-8.33, 5.05]	[-0.33, 2.28]	[-0.03, 0.12]	[-0.22, 24.76]
NYHA Class 1 or 2						
Entry Examination	1.04***	0.81***	0.68***		0.76***	
	[0.88, 1.19]	[0.66, 0.96]	[0.51, 0.84]		[0.58, 0.94]	

Notes: Asterisks indicate p-values following the Bonferroni adjustment for four primary endpoints with * p ‹ 0.025, ** p ‹ 0.0125, *** p ‹ 0.0025. For covariates: 98.75% confidence intervals in brackets. Endpoints: 6MWT distance in meters walked in six minutes; KCCQ Kansas City Cardiomyopathy Questionnaire summary score; EHFScBS: European Heart Failure Self-care Behaviour Scale summary Score; AHFKT number of correct answers in the Atlanta Heart Failure Knowledge Test; EQ5D EQ-5D-L utility weight based on Ludwig et al., 2018; SDMQ summary score from the 9-item Shared Decision Making Questionnaire. Missing values are imputed following a jump to reference approach where all observations are assigned the average values from the control group.

Table A.11: Results Full Analysis Set - Subsample Interim Participation - N = 88

	Primary Endpoints				Secondary Endpoints	
	6MWT	KCCQ	EHFScBS	AHFKT	EQ5D	SDMQ
Main Effects						
Point Estimate	47.15	-2.52	3.23	2.24***	-0.02	1.12
98.75% CI	[-5.66, 99.97]	[-11.02, 5.99]	[-5.29, 11.75]	[0.64, 3.83]	[-0.1, 0.05]	[-17.36, 19.6]
p-value	0.025	0.452	0.336	‹ 0.001	0.426	0.877
Cohen's d	0.49	-0.16	0.21	0.77	-0.17	0.03
95% CI	[0.06, 0.92]	[-0.59, 0.27]	[-0.22, 0.64]	[0.34, 1.2]	[-0.61, 0.26]	[-0.39, 0.46]
Covariates						
Male	-9.71	-1.80	-0.29	-1.86**	0.01	-0.33
	[-54.57, 35.16]	[-8.99, 5.39]	[-7.49, 6.92]	[-3.21, -0.52]	[-0.05, 0.07]	[-15.97, 15.31]
Age ‹ 65	25.73	11.58***	5.12	1.29	0.04	30.27***
	[-21.48, 72.94]	[4.39, 18.78]	[-2.11, 12.35]	[-0.06, 2.65]	[-0.03, 0.1]	[14.59, 45.96]
NYHA Class 1 or 2	25.74	-0.81	-3.57	-0.77	0.03	-10.99
	[-19.76, 71.25]	[-8.28, 6.67]	[-10.71, 3.57]	[-2.1, 0.56]	[-0.03, 0.09]	[-26.44, 4.47]
Entry Examination	0.93***	0.61***	0.69***		0.79***	
	[0.73, 1.13]	[0.45, 0.77]	[0.52, 0.86]		[0.6, 0.97]	

Notes: Asterisks indicate p-values following the Bonferroni adjustment for four primary endpoints with * p ‹ 0.025, ** p ‹ 0.0125, *** p ‹ 0.0025. For covariates: 98.75% confidence intervals in brackets. Endpoints: 6MWT distance in meters walked in six minutes; KCCQ Kansas City Cardiomyopathy Questionnaire summary score; EHFScBS: European Heart Failure Self-care Behaviour Scale summary Score; AHFKT number of correct answers in the Atlanta Heart Failure Knowledge Test; EQ5D EQ-5D-L utility weight based on Ludwig et al., 2018; SDMQ summary score from the 9-item Shared Decision Making Questionnaire. Missing values are imputed following a jump to reference approach where all observations are assigned the average values from the control group.

Table A.12: Results Full Analysis Set - Subsample No Interim Participation - N = 164

	Primary Endpoints				Secondary Endpoints	
	6MWT	KCCQ	EHFScBS	AHFKT	EQ5D	SDMQ
Main Effects						
Point Estimate	26.64	-0.83	5.94**	1.88***	-0.00	3.14
98.75% CI	[-4.58, 57.85]	[-6.24, 4.59]	[0.59, 11.28]	[0.67, 3.09]	[-0.05, 0.05]	[-7.15, 13.43]
p-value	0.033	0.7	0.006	‹ 0.001	0.963	0.442
Cohen's d	0.34	-0.06	0.44	0.61	-0.01	0.12
95% CI	[0.03, 0.65]	[-0.37, 0.25]	[0.13, 0.75]	[0.31, 0.92]	[-0.32, 0.3]	[-0.19, 0.43]
Covariates						
Male	14.18	-0.44	0.02	-0.62	0.00	-3.56
	[-15.51, 43.87]	[-5.61, 4.74]	[-5.07, 5.12]	[-1.78, 0.53]	[-0.05, 0.05]	[-13.38, 6.27]
Age ‹ 65	26.35	0.35	-3.17	0.73	0.00	1.94
	[0.97, 51.74]	[-3.95, 4.64]	[-7.38, 1.05]	[-0.23, 1.68]	[-0.04, 0.05]	[-6.17, 10.05]
NYHA Class 1 or 2	10.46	3.21	-0.56	0.99	0.08***	8.89
	[-17.92, 38.84]	[-1.89, 8.31]	[-5.09, 3.98]	[-0.03, 2.02]	[0.03, 0.12]	[0.18, 17.61]
Entry Examination	0.90***	0.84***	0.69***		0.73***	
	[0.79, 1.01]	[0.74, 0.95]	[0.59, 0.79]		[0.63, 0.83]	

Notes: Asterisks indicate p-values following the Bonferroni adjustment for four primary endpoints with * p ‹ 0.025, ** p ‹ 0.0125, *** p ‹ 0.0025. For covariates: 98.75% confidence intervals in brackets. Endpoints: 6MWT distance in meters walked in six minutes; KCCQ Kansas City Cardiomyopathy Questionnaire summary score; EHFScBS: European Heart Failure Self-care Behaviour Scale summary Score; AHFKT number of correct answers in the Atlanta Heart Failure Knowledge Test; EQ5D EQ-5D-L utility weight based on Ludwig et al., 2018; SDMQ summary score from the 9-item Shared Decision Making Questionnaire. Missing values are imputed following a jump to reference approach where all observations are assigned the average values from the control group.

Table A.13: Results without Covariate Adjustment - PMM Imputation

	Primary Endpoints				Secondary Endpoints	
	6MWT	KCCQ	EHFScBS	AHFKT	EQ5D	SDMQ
Main Effect						
Point Estimate	48.53***	2.03	5.62**	2.03***	-0.01	2.93
98.75% CI	[17.85, 79.20]	[-3.64, 7.70]	[0.60, 10.65]	[0.98, 3.08]	[-0.10, 0.07]	[-7.45, 13.32]
p-value	‹ 0.001	0.368	0.005	‹ 0.001	0.732	0.478
Cohen's d	0.48	0.10	0.23	0.61	-0.05	0.09
95% CI	[0.23, 0.73]	[-0.15, 0.34]	[-0.02, 0.48]	[0.36, 0.87]	[-0.29, 0.20]	[-0.16, 0.34]
Differences						
Diff. Int.	40.81	6.53	12.13		-0.09	
	(99.35)	(20.57)	(16.89)		(0.28)	
Diff. Cont.	-6.02	4.61	7.77		-0.07	
	(95.17)	(19.57)	(20.9)		(0.27)	
Mean Values						
Entry Int.	403.02	63.16	67.54		0.84	
	(120.51)	(22.06)	(19.73)		(0.18)	
Exit Int.	443.83	69.69	79.68	19.46	0.75	58.92
	(140.37)	(21.9)	(16.23)	(2.7)	(0.29)	(33.53)
Entry Cont.	387.57	62.88	65.08		0.84	
	(124.32)	(23.05)	(21.35)		(0.21)	
Exit Cont.	381.55	67.5	72.85	17.43	0.76	55.98
	(150.55)	(22.67)	(20.89)	(3.84)	(0.29)	(31.96)

Notes: Asterisks indicate p-values following the Bonferroni adjustment for four primary endpoints with * p ‹ 0.025, ** p ‹ 0.0125, *** p ‹ 0.0025. Standard deviations in parentheses. Endpoints: 6MWT distance in meters walked in six minutes; KCCQ Kansas City Cardiomyopathy Questionnaire summary score; EHFScBS: European Heart Failure Self-care Behaviour Scale summary Score; AHFKT number of correct answers in the Atlanta Heart Failure Knowledge Test; EQ5D EQ-5D-L utility weight based on Ludwig et al., 2018; SDMQ summary score from the 9-item Shared Decision Making Questionnaire. N = 252. Missing values are imputed using predictive mean matching.

Table A.14: Results without Covariate Adjustment - Per Protocol Sample

	Primary Endpoints				Secondary Endpoints	
	6MWT	KCCQ	EHFScBS	AHFKT	EQ5D	SDMQ
Main Effect						
Point Estimate	40.90***	-0.64	6.35***	2.45***	-0.01	4.64
98.75% CI	[8.69, 73.11]	[-6.20, 4.93]	[1.13, 11.58]	[1.30, 3.59]	[-0.06, 0.04]	[-6.75, 16.04]
p-value	0.002	0.774	0.002	< 0.001	0.742	0.306
Cohen's d	0.42	-0.06	0.31	0.74	-0.07	0.14
95% CI	[0.15, 0.69]	[-0.33, 0.21]	[0.04, 0.59]	[0.46, 1.02]	[-0.34, 0.20]	[-0.13, 0.41]
Differences						
Diff. Int.	37.95	3.7	10.09		0	
	(96.52)	(17.44)	(14.79)		(0.14)	
Diff. Cont.	-1.39	4.67	4.78		0.01	
	(90.13)	(16.55)	(18.72)		(0.16)	
Mean Values						
Entry Int.	403.02	63.16	67.54		0.84	
	(120.51)	(22.06)	(19.73)		(0.18)	
Exit Int.	449.37	67.6	78.07	19.22	0.85	58.73
	(139.49)	(22.58)	(17.27)	(2.79)	(0.19)	(32.85)
Entry Cont.	387.57	62.88	65.08		0.84	
	(124.32)	(23.05)	(21.35)		(0.21)	
Exit Cont.	389.3	67.28	70.05	16.77	0.84	54.09
	(145.13)	(23.19)	(21.42)	(3.76)	(0.22)	(32.85)

Notes: Asterisks indicate p-values following the Bonferroni adjustment for four primary endpoints with * p < 0.025, ** p < 0.0125, *** p < 0.0025. Standard deviations in parentheses. Endpoints: 6MWT distance in meters walked in six minutes; KCCQ Kansas City Cardiomyopathy Questionnaire summary score; EHFScBS: European Heart Failure Self-care Behaviour Scale summary Score; AHFKT number of correct answers in the Atlanta Heart Failure Knowledge Test; EQ5D EQ-5D-L utility weight based on Ludwig et al., 2018; SDMQ summary score from the 9-item Shared Decision Making Questionnaire. N = 214 for 6MWT and N = 211 non-missing observations for the other endpoints where there were non-missing values in the exit examination, out of the original sample of 252.

Table A.15: Change in NYHA Class - Per Protocol Sample

	Entry NYHA I			Entry NYHA II			Entry NYHA III		
Exit	NYHA I	NYHA II	NYHA III	NYHA I	NYHA II	NYHA III	NYHA I	NYHA II	NYHA III
Control	8 (53.3%)	5 (33.3%)	2 (13.3%)	10 (20.4%)	26 (53.1%)	13 (26.5%)	4 (9.1%)	13 (29.5%)	27 (61.4%)
Intervention	15 (83.3%)	2 (11.1%)	1 (5.6%)	13 (27.7%)	32 (68.1%)	2 (4.3%)	6 (15.0%)	16 (40.0%)	18 (45.0%)

Notes: Patient counts by NYHA class at entry and exit examination. Percentages in parentheses refer to the share of patients within each entry NYHA class.

Table A.16: Change in NYHA Class - Full Analysis Set

	Entry NYHA I				Entry NYHA II				Entry NYHA III			
Exit	NYHA I	NYHA II	NYHA III	NYHA IV	NYHA I	NYHA II	NYHA III	NYHA IV	NYHA I	NYHA II	NYHA III	NYHA IV
Control	8 (44.4%)	8 (44.4%)	2 (11.1%)	0 (0.0%)	10 (17.9%)	26 (46.4%)	20 (35.7%)	0 (0.0%)	4 (7.8%)	13 (25.5%)	27 (52.9%)	7 (13.7%)
Intervention	15 (75.0%)	4 (20.0%)	1 (5.0%)	0 (0.0%)	13 (23.2%)	32 (57.1%)	11 (19.6%)	0 (0.0%)	6 (11.8%)	16 (31.4%)	18 (35.3%)	11 (21.6%)

Notes: Patient counts by NYHA class at entry and exit examination. Percentages in parentheses refer to the share of patients within each entry NYHA class. Missing exit values were conservatively imputed by assigning patients to one NYHA class higher than at entry.

Table A.17: Statistical Summary of Usage Intensity

	Value
Mean	34
SD	18
Median	32
Minimum	0
Maximum	94
25Q	25
75Q	35

Table A.18: Descriptive Statistics Split by Interim Assignment

	Assigned to Interim (N=100)		No Interim assigned (N=152)			
	Mean	Std. Dev.	Mean	Std. Dev.	Diff. in Means	p
Study Characeristics						
Follow-up	87.00%		83.55%		-3.45	0.447
Intervention time	186.59	8.73	190.03	23.18	3.45	0.129
Sociodemographics						
Age	66.36	10.90	63.13	14.45	-3.23	0.045
Male	69.00%		79.61%		10.61	0.064
Living alone	26.00%		24.50%		-1.50	0.791
Smoker	15.00%		14.67%		-0.33	0.942
Health Status						
Elixhauser score	3.72	4.81	10.24	6.27	6.52	‹0.001
BMI	28.53	5.97	27.59	5.51	-0.93	0.212
NYHA Class						
I	13.00%		16.45%		3.45	0.447
II	31.00%		53.29%		22.29	‹0.001
III	56.00%		30.26%		-25.74	‹0.001
Endpoints						
6MWT	373.84	124.47	409.51	119.35	35.67	0.025
KCCQ	62.49	22.72	63.38	22.44	0.89	0.760
EHFScBS	67.53	18.88	65.52	21.60	-2.01	0.436
EQ5D	0.85	0.17	0.83	0.21	-0.02	0.446

Notes: The table reports group specific means and standard deviations and p-values comparing patients assigned to an interim examination with those who were not. Endpoints: 6MWT distance in meters walked in six minutes; KCCQ Kansas City Cardiomyopathy Questionnaire summary score; EHFScBS: European Heart Failure Self-care Behaviour Scale summary Score; AHFKT number of correct answers in the Atlanta Heart Failure Knowledge Test; EQ5D EQ-5D-L utility weight based on Ludwig et al., 2018; SDMQ summary score from the 9-item Shared Decision Making Questionnaire.

Table A.19: Results with Interim Examination Control - Full Analysis Set

	Primary Endpoints				Secondary Endpoints	
	6MWT	KCCQ	EHFScBS	AHFKT	EQ5D	SDMQ
Main Effects						
Point Estimate	33.71***	-0.74	5.25**	2.05***	-0.01	3.37
98.75% CI	[6.88, 60.54]	[-5.38, 3.89]	[0.74, 9.76]	[1.09, 3.01]	[-0.05, 0.03]	[-5.99, 12.73]
p-value	0.002	0.686	0.004	< 0.001	0.595	0.365
Cohen's d	0.40	-0.05	0.37	0.68	-0.07	0.11
95% CI	[0.15, 0.65]	[-0.3, 0.2]	[0.12, 0.62]	[0.43, 0.93]	[-0.32, 0.18]	[-0.13, 0.36]
Covariates						
Male	4.67	-0.49	0.32	-1.11*	0.01	-0.92
	[-19.9, 29.24]	[-4.73, 3.75]	[-3.8, 4.44]	[-1.98, -0.23]	[-0.03, 0.04]	[-9.48, 7.63]
Age < 65	27.41*	3.86	-0.66	0.80	0.01	10.45**
	[5.01, 49.81]	[0.13, 7.6]	[-4.28, 2.97]	[0.03, 1.57]	[-0.02, 0.04]	[2.93, 17.98]
NYHA Class 1 or 2	16.32	2.74	-1.12	0.38	0.06***	3.29
	[-7.64, 40.28]	[-1.5, 6.98]	[-4.91, 2.67]	[-0.43, 1.18]	[0.03, 0.1]	[-4.59, 11.16]
Entry Examination Value	0.90***	0.76***	0.69***		0.74***	
	[0.8, 1]	[0.67, 0.85]	[0.6, 0.77]		[0.66, 0.83]	
Interim Indicator	37.81***	5.41**	3.16	0.94*	0.03	-6.56
	[15.03, 60.59]	[1.45, 9.37]	[-0.67, 7]	[0.13, 1.76]	[-0.01, 0.06]	[-14.52, 1.4]

Notes: Asterisks indicate p-values following the Bonferroni adjustment for four primary endpoints with * p < 0.025, ** p < 0.0125, *** p < 0.0025. For covariates: 98.75% confidence intervals in brackets. Endpoints: 6MWT distance in meters walked in six minutes; KCCQ Kansas City Cardiomyopathy Questionnaire summary score; EHFScBS: European Heart Failure Self-care Behaviour Scale summary Score; AHFKT number of correct answers in the Atlanta Heart Failure Knowledge Test; EQ5D EQ-5D-L utility weight based on Ludwig et al., 2018; SDMQ summary score from the 9-item Shared Decision Making Questionnaire. Interim Indicator is 1 if the patient participated in the interim examination.

Table A.20: Results with Interim Examination Control - PMM

	Primary Endpoints				Secondary Endpoints	
	6MWT	KCCQ	EHFScBS	AHFKT	EQ5D	SDMQ
Main Effects						
Point Estimate	49.26***	1.96	5.54**	2.04***	-0.01	2.38
98.75% CI	[19.3, 79.23]	[-3.63, 7.55]	[0.5, 10.59]	[1.01, 3.08]	[-0.09, 0.07]	[-7.7, 12.46]
p-value	< 0.001	0.378	0.006	< 0.001	0.778	0.553
Cohen's d	0.52	0.11	0.35	0.63	-0.04	0.07
95% CI	[0.27, 0.77]	[-0.14, 0.36]	[0.1, 0.6]	[0.38, 0.87]	[-0.28, 0.21]	[-0.17, 0.32]
Covariates						
Male	-2.57	-1.60	2.25	-1.16*	-0.00	-1.76
	[-30.01, 24.88]	[-6.72, 3.51]	[-2.36, 6.86]	[-2.11, -0.21]	[-0.08, 0.08]	[-10.97, 7.45]
Age < 65	27.13	6.66**	1.47	1.06*	-0.00	12.68***
	[2.11, 52.14]	[2.16, 11.17]	[-2.59, 5.53]	[0.23, 1.89]	[-0.07, 0.07]	[4.58, 20.79]
NYHA Class 1 or 2	15.64	0.64	-0.37	-0.37	0.08*	6.27
	[-11.11, 42.4]	[-4.47, 5.75]	[-4.61, 3.88]	[-1.24, 0.5]	[0.01, 0.15]	[-2.21, 14.75]
Entry Examination Value	0.85***	0.57***	0.48***		0.56***	
	[0.74, 0.96]	[0.46, 0.68]	[0.39, 0.58]		[0.39, 0.73]	
Interim Indicator	43.71***	4.05	1.96	-0.01	0.12***	-7.85
	[18.27, 69.15]	[-0.73, 8.82]	[-2.33, 6.25]	[-0.89, 0.87]	[0.05, 0.19]	[-16.42, 0.72]

Notes: Asterisks indicate p-values following the Bonferroni adjustment for four primary endpoints with * p < 0.025, ** p < 0.0125, *** p < 0.0025. For covariates: 98.75% confidence intervals in brackets. Endpoints: 6MWT distance in meters walked in six minutes; KCCQ Kansas City Cardiomyopathy Questionnaire summary score; EHFScBS: European Heart Failure Self-care Behaviour Scale summary Score; AHFKT number of correct answers in the Atlanta Heart Failure Knowledge Test; EQ5D EQ-5D-L utility weight based on Ludwig et al., 2018; SDMQ summary score from the 9-item Shared Decision Making Questionnaire. Interim Indicator is 1 if the patient participated in the interim examination.

Table A.21: Results with Interim Examination Control - Per Protocol Sample

	Primary Endpoints				Secondary Endpoints	
	6MWT	KCCQ	EHFScBS	AHFKT	EQ5D	SDMQ
Main Effects						
Point Estimate	40.91***	-0.56	6.56***	2.51***	-0.01	3.48
98.75% CI	[9.48, 72.33]	[-6.04, 4.92]	[1.31, 11.81]	[1.39, 3.64]	[-0.06, 0.04]	[-7.72, 14.67]
p-value	0.001	0.797	0.002	< 0.001	0.739	0.435
Cohen's d	0.45	-0.04	0.44	0.78	-0.05	0.11
95% CI	[0.18, 0.72]	[-0.31, 0.24]	[0.16, 0.71]	[0.51, 1.05]	[-0.32, 0.23]	[-0.16, 0.38]
Covariates						
Male	5.13	-0.76	1.14	-1.37**	0.01	-1.45
	[-23.59, 33.86]	[-5.75, 4.22]	[-3.65, 5.93]	[-2.39, -0.35]	[-0.04, 0.05]	[-11.63, 8.72]
Age < 65	34.44**	4.87	-0.49	0.97	0.01	12.85**
	[7.75, 61.14]	[0.45, 9.3]	[-4.71, 3.73]	[0.06, 1.87]	[-0.03, 0.05]	[3.84, 21.86]
NYHA Class 1 or 2	18.60	3.28	-0.25	0.31	0.07***	3.03
	[-9.48, 46.68]	[-1.79, 8.35]	[-4.78, 4.28]	[-0.66, 1.28]	[0.03, 0.12]	[-6.63, 12.68]
Entry Examination Value	0.86***	0.70***	0.61***		0.71***	
	[0.74, 0.98]	[0.59, 0.81]	[0.5, 0.71]		[0.61, 0.81]	
Interim Indicator	38.77**	6.09**	4.17	0.84	0.04	-8.00
	[12.28, 65.26]	[1.44, 10.73]	[-0.27, 8.62]	[-0.11, 1.79]	[0, 0.08]	[-17.45, 1.45]

Notes: Asterisks indicate p-values following the Bonferroni adjustment for four primary endpoints with * p < 0.025, ** p < 0.0125, *** p < 0.0025. For covariates: 98.75% confidence intervals in brackets. Endpoints: 6MWT distance in meters walked in six minutes; KCCQ Kansas City Cardiomyopathy Questionnaire summary score; EHFScBS: European Heart Failure Self-care Behaviour Scale summary Score; AHFKT number of correct answers in the Atlanta Heart Failure Knowledge Test; EQ5D EQ-5D-L utility weight based on Ludwig et al., 2018; SDMQ summary score from the 9-item Shared Decision Making Questionnaire. Interim Indicator is 1 if the patient participated in the interim examination.