SUPPLEMENTAL MATERIAL

Do self-management interventions work in patients with heart failure work?
An individual patient data meta-analysis

Page	Content
2	**Supplemental Methods:** Statistical analysis plan
6	**Supplemental Table 1:** Effects of self-management interventions on subordinate outcomes in patients with heart failure included in the individual patient data meta-analysis.
7	**Supplemental Table 2:** Effects of self-management interventions on subordinate outcomes in subgroups of patients with heart failure included in the individual patient data meta-analysis.
8	**Supplemental Table 3:** Effects of self-management interventions on main outcomes in subgroups of patients with heart failure included in the individual patient data meta-analysis.
9	**Supplemental Table 4:** Sensitivity analysis on main outcomes by including published main effects of eligible studies without available individual patient data.
10	**Supplemental Table 5:** Sensitivity analysis on main outcomes by excluding trials with enhanced usual care in the comparison group.
11	**Supplemental References**

Supplemental Methods: Statistical analysis plan

This document contains the plan for the statistical analysis for the individual patient data (IPD) meta-analysis in heart failure (HF) patients. Input from the conference calls on March 20th, 2014 and March 31st, 2014 and email contact has been processed in the statistical plan presented in this document.

A schematic overview of the statistical analysis is present in Figure 1. Each step will be explained in more detail in the subsequent paragraphs. For all statistical analyses, the software R for Windows version 3.1.1 (R Development Core Team. Released 2013. Vienna, Austria: R Foundation for Statistical Computing) will be used.

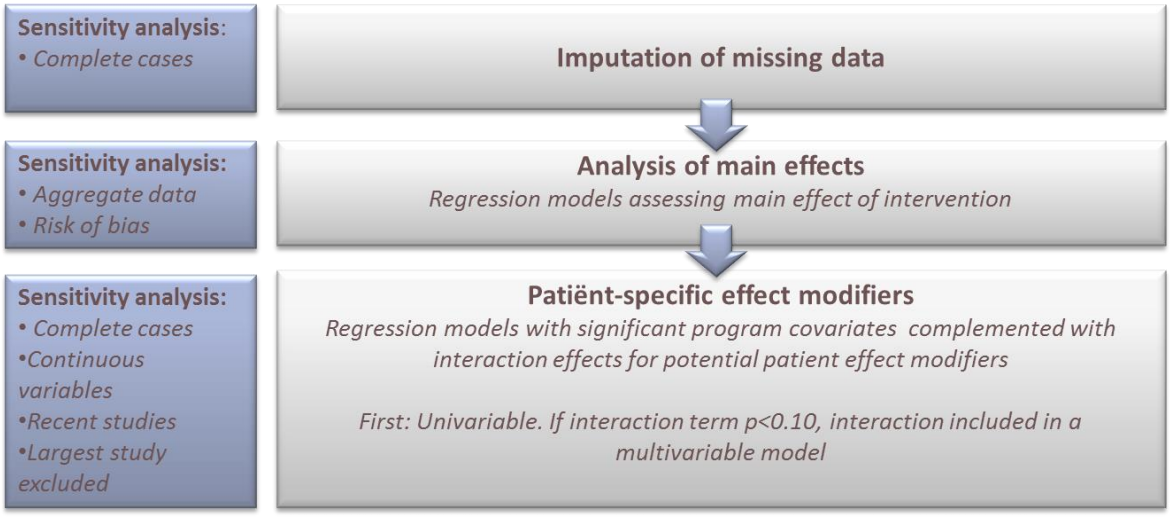

Figure 1: Steps of the statistical analysis of patient-specific determinants of self-management interventions.

1. Imputation of missing data

To address bias due to missing data, we will impute missing data using multiple imputation by chained equations (MICE).[1] The MICE algorithm accounts for the order in which the values of separate variables are predicted through chained equations. To address the uncertainty of just one single imputation, MICE creates multiple imputations, resulting in multiple imputed datasets.

The imputation will be performed according to the following principles:

- Missing values will only be imputed *within* studies: this implies that only the correlation between variables available within one study will be used to estimate the missing values in that particular study
- All available variables (except patient identifiers) will be used to estimate missing values
- Multiple imputation will be used to estimate missing values for patient characteristics and outcomes
- Multiple imputation will be performed 25 times, resulting in 25 imputed datasets
- As a result, all analyses will be carried out 25 times. Results will be pooled using Rubin's rule for the final results.[2]

A complete-case analysis, using only the available patient data, will be performed as a sensitivity analysis to assess the impact of imputing data (see '4. Sensitivity analyses').

2. Analysis of main effects

All data will be analyzed according to the intention-to-treat principle. A so-called one-stage approach will be used, where all patients are analyzed simultaneously in one model while clustering of observations within studies is taken into account.[3]

The present study will analyze the following main outcome measures:

- Composite of time to first disease-related hospital admission or all-cause death;
- Change in health-related quality of life (HRQoL) at 12 months, compared to baseline;
    - A distinction will be made between disease-specific and generic HRQoL to address the different instruments used by original studies
- Time to first disease-related hospital admission;
- Total number of days spent in hospital for HF at 12 months.
- Time to all-cause death;
- Time to first all-cause hospital admission;
- Total number of days spent in hospital for any cause at 12 months.

Additionally, the following subordinate outcomes measures will be analyzed:

- Change in health-related quality of life (HRQoL) at 6 months, compared to baseline;
    - A distinction will be made between disease-specific and generic HRQoL
- Total number of days spent in hospital for HF at 6 months and at 12 months;
- Hospitalized for HF at 6 months;
- All-cause mortality at 6 months and at 12 months;
- Hospitalized for any cause at 6 months and at 12 months;
- Total number of days spent in hospital for any cause at 6 months.

For time-to-event data, effects of self-management will be quantified by estimating hazard ratios (HR) and 95% confidence interval (CI). Cox proportional-hazard models will be used to analyze the data, including a cluster statement to allow inter study variability. For binary outcome data (mortality, all-cause and disease-related hospital admissions), risk ratios (RR) and 95% CI will be estimated using log-binomial mixed effects models. Effects on continuous outcomes (HRQoL) will be quantified by mean differences and 95% CI and will be estimated using linear mixed effects models. Effects on total length of hospital stay will be analyzed with negative binomial mixed effects models to model overdispersion in the data. In the (generalized) linear mixed effects models, random intercepts and random slopes will be included to take clustering within studies into account.

3. Patient-specific effect modifiers

The aforementioned models will be extended to study effect modification by patient characteristics. Effect modification implies that the effect of the intervention on an outcome differs depending on the value of a third variable, the effect modifier. As such, we will be able to identify subgroups of patients in which self-management interventions work best. Interaction terms will be included in the final model resulting from the previous step, which includes the significant program determinants.

We have selected clinically relevant patient characteristics as potential effect modifiers, these are presented in Table 1. Numbers of patients differ per variable due to the fact that some baseline variables have not been collected in one (or more) studies. We would like to categorize the variables to create relevant subgroups for the interpretation of findings. This has been discussed extensively during the conference calls, and the proposed categories are a result of the discussions.

Like the analysis of program characteristics, patient characteristics with $p<0.10$ in the separate analyses will be fitted together in a multivariable model. Effect modifiers will be presented with 95% confidence intervals. Results will be interpreted with great caution to decrease the risk of type I error (i.e. descriptive analysis, consistency with expectations, other findings).

After consulting the investigators during the conference calls we have decided to exclude baseline ***self-efficacy level*** of patients from the analysis. This variable has only been collected in 4 studies (n=1321), each using a different instrument.

Table 1: Patient characteristics to be analyzed as potential effect modifiers.

Determinant	Data in database	Proposed categories for analysis	Statistics in database
Sex (n=5624)	1. Male 2. Female	1. Male 2. Female	57.2% 42.8%
Age (n=5624)	Years	 1. <65 years 2. 65-80 years 3. >80 years	Mean(SD)=69.7(12.4) 30.5% 50.7% 18.9%
Disease severity (n=3562)	% LVEF	 1. ≤35% LVEF (REF) 2. >35% LVEF (based on ESC Guidelines 2012)	Mean(SD)=39.2(18.2) 52.0% 48.0%
Symptom severity (n=5328)	NYHA class (I-IV)	1. NYHA I & II 2. NYHA III 3. NYHA IV	46.1% 36.9% 17.0%
Comorbidity index (n=5079)	# of clusters of comorbid conditions	1. No comorbid conditions 2. Comorbid conditions in 1 cluster 3. Comorbid conditions in ≥2 clusters	*Categories still to be calculated for each individual study*
Depression (n=2998)	Score on instrument	1. No/mild depression 2. Moderate/severe depression (based on validated cut-offs of each instrument)	*Cut-offs still to be calculated for each individual study*
Level of education (n=4216)	1. Primary or below 2. Secondary 3. Higher	1. Primary or below 2. Secondary 3. Higher	40.7% 39.1% 20.1%
Years since diagnosis (n=2310)	Months/Years/Categories	 1. <1 year diagnosed 2. 1 -<2 years diagnosed 3. ≥2 years diagnosed	Median(IQR)=1.6(0.1-5.4) 44.1% 12.5% 43.4%
Living status (n=2883)	1. Living alone 2. Not living alone	1. Living alone 2. Not living alone	25.8% 74.2%
Body Mass Index (n=3201)	BMI score	 1. BMI <25 (underweight/normal) 2. BMI 25 - 29.99 (overweight) 3. BMI ≥30 (obese)	Mean(SD)=28.0(6.6) 35.3% 35.0% 29.7%
Smoking status (n=2376)	1. Current smoker 2. Former smoker 3. Never smoker	1. Current smoker 2. Current non-smoker	18.9% 81.1%

BMI indicates Body Mass Index;; LVEF, Left Ventricular Ejection Fraction; NYHA, New York Heart Association; and REF, Reduced Ejection Fraction.

Explanation of scoring comorbidity index

We would like to study the effect of comorbid conditions on effectiveness of self-management, since we expect patients with a higher comorbid burden to benefit less from self-management interventions. Yet, comorbidity has been collected very differently across the different studies. If we simply score the number of comorbidities as all comorbidities collected in a study, patients in studies collecting more diagnoses have a higher risk of having a higher comorbidity score (which biases the results).

We propose a recoding of comorbid diagnoses collected in each study into the following clusters:

1. Cardiovascular conditions
2. Endocrine conditions (incl. diabetes)
3. Neurological/psychiatric conditions
4. Respiratory conditions
5. Renal/hepatic/gastrointestinal conditions
6. Cancer
7. Musculoskeletal conditions

Patients will be scored on presence of a comorbid condition within each cluster. Clusters of comorbid conditions are based on the Cumulative Illness Rating Scale.[4]

We aim to score the comorbid burden of patients by categorizing patients in:

- No comorbid conditions
- Comorbid conditions in 1 cluster
- Comorbid conditions in ≥2 clusters

Data will be analyzed more in-depth in a descriptive manner to cautiously interpret any findings with regard to this comorbidity index.

4. Sensitivity analyses

Several sensitivity analyses will be performed to assess the robustness of the findings:

1. *Inclusion of aggregate data of studies for which IPD are unavailable:*
   To assess if IPD included are representative of all studies invited for this project, an aggregate data meta-analysis will be performed to assess the impact of missing studies on the main effect for each endpoint.
2. *Inclusion of only studies with a low risk of bias:*
   To assess whether methodological quality of studies has an impact on findings, the studies scoring a 'high risk of bias' on attrition bias on tool from the Cochrane Collaboration[5] will be left out of the analysis to assess the impact on the main effect for each endpoint.
3. *Inclusion of only complete cases:*
   To assess the effect of imputing missing data, all analyses will be repeated with a dataset containing only the patients for whom data are available. This will be performed for the analyses for main effects as well as patient-specific effect modifiers.
4. *Inclusion of continuous patient characteristics instead of categorized scores:*
   To assess the loss of information by categorizing continuous patient characteristics for the subgroup analysis, a sensitivity analysis is performed using the continuous data instead of categorized data for those patient characteristics. This applies to the effect modification of the variables age, % LVEF, years since diagnosis, and BMI.
5. *Inclusion of only newer studies (recruitment since 2000):*
   To assess of observed effects are robust over time, the sensitivity analysis will be repeated by only including more recently conducted studies (recruitment since 2000).
6. *Excluding the largest trial:**
   To assess if subgroup effects are attributable to a specific study (particularly the largest trial) or whether they can be generalized across studies, the subgroup analysis will be repeated without the largest trial.

**N.B: this sensitivity analysis was extended post hoc, by excluding each study one-by-one and repeating the subgroup analysis without that study (i.e., a leave-one-out analysis) to assess the impact of each study.*

Supplemental Table 1: Effects of self-management interventions on subordinate outcomes in patients with heart failure included in the individual patient data meta-analysis.

Outcome	N studies	n patients	Effect measure	Treatment effect (95% CI)
Heart failure-related outcomes				
HF-related QoL – 6 months	10	3419	SMD	0.13 (0.00-0.26)
HF-related hospitalization – 6 months	12	3742	RR	0.81 (0.66-0.99)
HF-related hospitalization – 12 months	11	3503	RR	0.82 (0.64-1.05)
Total days HF-related hospital stay – 6 months	8	1734	RLOS	0.67 (0.46-0.99)
General outcomes				
Generic QoL – PCS – 6 months	3	888	MD	1.13 (-2.25-4.52)
Generic QoL – MCS – 6 months	3	888	MD	1.89 (-2.90-6.68)
Mortality – 6 months	17	4999	RR	0.83 (0.66-1.05)
Mortality – 12 months	14	4204	RR	0.86 (0.72-1.03)
All-cause hospitalization – 6 months	14	4329	RR	0.92 (0.83-1.01)
All-cause hospitalization – 12 months	13	4266	RR	0.95 (0.87-1.04)
Total days all-cause hospital stay – 6 months	10	2820	RLOS	0.96 (0.74-1.25)

CI indicates confidence interval; HF, heart failure; MCS, mental component scale Short Form Health Survey; MD, mean difference; PCS, physical component scale Short Form Health Survey; QoL, quality of life; RLOS, relative length of stay; RR, risk ratio; and SMD, standardized mean difference.

Supplemental Table 2: Effects of self-management interventions on subordinate outcomes in subgroups of patients with heart failure included in the individual patient data meta-analysis.

Outcome	n patients	Effect measure	UNIVARIABLE ANALYSIS		MULTIVARIABLE ANALYSIS	
			Treatment effect (95% CI)	p-value for interaction	Treatment effect (95% CI)	p-value for interaction
Heart failure-related outcomes						
Subgroup						
HF-related QoL – 6 months						
No subgroup effects.						
HF-related hospitalization – 6 months						
No subgroup effects.						
HF-related hospitalization – 12 months						
NYHA I-II	1770	OR	0.81 (0.56-1.18)	0.06	*	
NYHA III	1323		0.90 (0.62-1.28)			
NYHA IV	410		0.44 (0.26-0.75)			
Total days HF-related hospital stay – 6 months						
<65 years	339	RLOS	0.24 (0.10-0.61)	0.06	*	
65-80 years	985		0.82 (0.45-1.48)			
>80 years	410		0.91 (0.39-2.12)			
General outcomes						
Subgroup						
Generic QoL – PCS – 6 months						
No subgroup effects.						
Generic QoL – MCS – 6 months						
No subgroup effects.						
Mortality – 6 months						
<65 years	1538	RR	1.32 (0.85-2.06)	0.02	1.32 (0.70-2.48)	0.07
65-80 years	2537		0.80 (0.61-1.06)		0.82 (0.47-1.43)	
>80 years	934		0.63 (0.45-0.89)		0.64 (0.36-1.16)	
No comorbidities	835	RR	0.87 (0.56-1.35)	0.02	1.32 (0.70-2.48)	0.02
Comorbidities in 1 cluster	1632		0.59 (0.44-0.79)		0.86 (0.50-1.48)	
Comorbidities in >1 cluster	1885		0.99 (0.77-1.27)		1.56 (0.92-2.66)	
Mortality – 12 months						
No subgroup effects.						
All-cause hospitalization – 6 months						
No subgroup effects.						
All-cause hospitalization – 12 months						
Not living alone	1555	RR	0.88 (0.78-0.99)	0.08	*	
Living alone	571		1.05 (0.87-1.27)			
Total days all-cause hospital stay – 6 months						
No subgroup effects.						

CI indicates confidence interval; HF, heart failure; MCS, mental component scale Short Form Health Survey; NYHA, New York Heart Association; OR, odds ratio; PCS, physical component scale Short Form Health Survey; QoL, quality of life; RLOS, relative length of stay; and RR, risk ratio.

Results of the subgroup analyses are only presented if a potential effect modifier showed an effect with $p<0.10$ in the univariable analysis.

*To adjust for other relevant effect modifiers, multivariable analysis was only performed if there were two or more potential effect modifiers in the univariable analysis.

Supplemental Table 3: Effects of self-management interventions on main outcomes in subgroups of patients with heart failure included in the individual patient data meta-analysis.

Outcome Subgroup	n patients	Effect measure	UNIVARIABLE ANALYSIS		MULTIVARIABLE ANALYSIS	
			Treatment effect (95% CI)	p-value for interaction	Treatment effect (95% CI)	p-value for interaction
Heart failure-related outcomes *Subgroup*						
HF-related hospitalization/ mortality – time to event						
No subgroup effects.						
HF-related QoL – 12 months						
No subgroup effects.						
HF-related hospitalization – time to event						
NYHA I-II	1579	HR	0.87 (0.70-1.08)	0.06	*	
NYHA III	1399		0.83 (0.68-1.03)			
NYHA IV	483		0.53 (0.37-0.77)			
Total days HF-related hospital stay – 12 months						
<65 years	139	RLOS	0.09 (0.02-0.38)	**0.03**	*	
65-80 years	521		0.95 (0.46-1.94)			
>80 years	232		0.96 (0.31-2.97)			
General outcomes *Subgroup*						
Generic QoL - PCS – 12 months						
No subgroup effects.						
Generic QoL - MCS – 12 months						
Current non-smokers	796	MD	-0.19 (-3.34-2.97)	0.09	*	
Current smokers	113		4.91 (-1.07-10.89)			
Mortality – time to event						
No/mild depression	1619	HR	0.86 (0.69-1.06)	**0.01**	*	
Moderate/severe depression	814		1.39 (1.06-1.83)			
All-cause hospitalization – time to event						
<65 years	1188	HR	1.09 (0.88-1.36)	0.07	0.93 (0.73-1.18)	0.35
65-80 years	1928		0.92 (0.75-1.15)		0.82 (0.69-0.98)	
>80 years	717		0.79 (0.60-1.04)		0.73 (0.57-0.95)	
Primary education	1283	HR	0.82 (0.71-0.96)	**0.02**	0.93 (0.73-1.18)	0.07
Secondary education	1110		0.98 (0.82-1.17)		1.09 (0.86-1.38)	
Higher education	653		1.26 (0.99-1.60)		1.33 (1.01-1.76)	
<1 year diagnosed	822	HR	1.13 (0.91-1.41)	0.08	†	
1-2 years diagnosed	168		1.61 (1.00-2.58)			
>2 years diagnosed	549		0.91 (0.72-1.14)			
Total days all-cause hospital stay – 12 months						
No subgroup effects.						

CI indicates confidence interval; HF, heart failure; HR, hazard ratio; MCS, mental component scale Short Form Health Survey; MD, mean difference; NYHA, New York Heart Association; PCS, physical component scale Short Form Health Survey; QoL, quality of life; and RLOS, relative length of stay.

*To adjust for other relevant effect modifiers, multivariable analysis was only performed if two or more potential effect modifiers in the univariable analysis were $p<0.10$.

†Years diagnosed not included as covariate in multivariable analysis since only N=1 study contained data on all covariates.

Supplemental Table 4: Sensitivity analysis on main outcomes by including published main effects of eligible studies without available individual patient data.

		Primary analysis (individual patient data only)			Pooled analysis of individual patient data and published effects		
	Effect	**Studies**	**Patients**	**Effect size (95% CI)**	**Studies**	**Patients**	**Effect size (95% CI)**
Heart failure-related outcomes							
HF-related hospitalization/ mortality - time to event	HR	10	3461	0.80 (0.71-0.89)	Published data could not be pooled		
HF-related QoL 12 months	SMD	11	3356	0.15 (0.00-0.30)	17	4370	0.14 (0.03-0.26)
HF-related hospitalization time to event	HR	10	3461	0.80 (0.69-0.92)	12	4327	0.79 (0.69-0.90)
Total days HF-related hospital stay - 12 months	RLOS	5	892	0.86 (0.44-1.67)	Published data could not be pooled		
General outcomes							
Generic QoL - PCS 12 months	MD	8	1739	0.95 (-1.15-3.05)	Published data could not be pooled		
Generic QoL - MCS 12 months	MD	8	1739	0.27 (-2.53-3.08)	Published data could not be pooled		
Mortality time to event	HR	14	4312	0.91 (0.79-1.04)	17	5326	0.89 (0.78-1.01)
All-cause hospitalization time to event	HR	12	3833	0.93 (0.85-1.03)	14	4699	0.93 (0.85-1.00)
Total days all-cause hospital stay - 12 months	RLOS	9	2304	0.97 (0.77-1.23)	Published data could not be pooled		

CI indicates confidence interval; MCS, mental component scale Short Form Health Survey; MD, mean difference; PCS, physical component scale Short Form Health Survey; QoL, quality of life; RLOS, relative length of stay; RR, risk ratio; and SMD, standardized mean difference.

Supplemental Table 5: Sensitivity analysis on main outcomes by excluding trials with enhanced usual care in the comparison group.

		Primary analysis (all studies included)				Analysis without DeWalt, 2012[6] & Heisler, 2013[7]			
	Effect	**Stu-dies**	**Pa-tient**	**Effect size (95% CI)**	I^2	**Stu-dies**	**Pa-tient**	**Effect size (95% CI)**	I^2
Heart failure-related outcomes									
HF-related hospitalization/ mortality - time to event	HR	10	3461	0.80 (0.71-0.89)	51.6%	9	2856	0.78 (0.69-0.88)	53.2%
HF-related QoL 12 months	SMD	11	3356	0.15 (0.00-0.30)	43.6%	10	2751	0.16 (-0.02-0.34)	48.7%
HF-related hospitalization time to event	HR	10	3461	0.80 (0.69-0.92)	60.8%	9	2856	0.76 (0.66-0.89)	59.7%
Total days HF-related hospital stay - 12 months	RLOS	5	892	0.86 (0.44-1.67)	0.0%	No outcomes reported by DeWalt/Heisler			
General outcomes									
Generic QoL - PCS 12 months	MD	8	1739	0.95 (-1.15-3.05)	0.0%	No outcomes reported by DeWalt/Heisler			
Generic QoL - MCS 12 months	MD	8	1739	0.27 (-2.53-3.08)	0.0%	No outcomes reported by DeWalt/Heisler			
Mortality time to event	HR	14	4312	0.91 (0.79-1.04)	43.8%	12	3441	0.87 (0.76-1.00)	24.4%
All-cause hospitalization time to event	HR	12	3833	0.93 (0.85-1.03)	53.1%	10	2962	0.89 (0.80-0.99)	49.6%
Total days all-cause hospital stay - 12 months	RLOS	9	2304	0.97 (0.77-1.23)	82.2%	7	1443	0.90 (0.68-1.20)	86.3%

CI indicates confidence interval; MCS, mental component scale Short Form Health Survey; MD, mean difference; PCS, physical component scale Short Form Health Survey; QoL, quality of life; RLOS, relative length of stay; RR, risk ratio; and SMD, standardized mean difference.

Supplemental References

1. Groothuis-Oudshoorn K, van Buuren S. MICE: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software.* 2011;45.
2. Rubin DB. *Multiple imputation for non-response in surveys*. New York: John Wiley & Sons; 1987.
3. Simmonds MC, Higgins JP, Stewart LA, Tierney JF, Clarke MJ, Thompson SG. Meta-analysis of individual patient data from randomized trials: a review of methods used in practice. *Clin Trials.* 2005;2:209-217.
4. Linn BS, Linn MW, Gurel L. Cumulative illness rating scale. *J Am Geriatr Soc.* 1968;16:622-626.
5. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions, version 5.1.0. The Cochrane Collaboration; 2011.
6. DeWalt DA, Schillinger D, Ruo B, Bibbins-Domingo K, Baker DW, Holmes GM, Weinberger M, abasco-O'Connell A, Broucksou K, Hawk V, Grady KL, Erman B, Sueta CA, Chang PP, Cene CW, Wu JR, Jones CD, Pignone M. Multisite randomized trial of a single-session versus multisession literacy-sensitive self-care intervention for patients with heart failure. *Circulation.* 2012;125:2854-2862.
7. Heisler M, Halasyamani L, Cowen ME, Davis MD, Resnicow K, Strawderman RL, Choi H, Mase R, Piette JD. A Randomized Controlled Effectiveness Trial of Reciprocal Peer Support in Heart Failure. *Circ Heart Fail.* 2013;6:246-253.