

University of Louisville

## ThinkIR: The University of Louisville's Institutional Repository

---

Electronic Theses and Dissertations

---

5-2010

### Technology-driven heart failure management system : changes in functional capacity and quality of life.

Venkata S. Chilakapati  
*University of Louisville*

Follow this and additional works at: <https://ir.library.louisville.edu/etd>

---

#### Recommended Citation

Chilakapati, Venkata S., "Technology-driven heart failure management system : changes in functional capacity and quality of life." (2010). *Electronic Theses and Dissertations*. Paper 242.  
<https://doi.org/10.18297/etd/242>

This Doctoral Dissertation is brought to you for free and open access by ThinkIR: The University of Louisville's Institutional Repository. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of ThinkIR: The University of Louisville's Institutional Repository. This title appears here courtesy of the author, who has retained all other copyrights. For more information, please contact [thinkir@louisville.edu](mailto:thinkir@louisville.edu).

TECHNOLOGY-DRIVEN HEART FAILURE MANAGEMENT SYSTEM:  
CHANGES IN FUNCTIONAL CAPACITY AND QUALITY OF LIFE

By

Venkata S. Chilakapati  
M.D., Andhra Medical College, 1999  
M.P.H., Western Kentucky University, 2001  
Internal Medicine Residency, Michael Reese Hospital, 2007  
Cardiology Fellowship, University of Cincinnati, 2010

A Dissertation  
Submitted to the Faculty of the  
School of Public Health and Information Sciences of the University of Louisville  
in Partial Fulfillment of the Requirements  
for the Degree of

Doctor of Philosophy

Department of Epidemiology and Clinical Investigation Sciences  
University of Louisville  
Louisville, Kentucky

May 2010

Copyright 2010 by Venkata S. Chilakapati

All rights reserved

TECHNOLOGY-DRIVEN HEART FAILURE MANAGEMENT SYSTEM:  
CHANGES IN FUNCTIONAL CAPACITY AND QUALITY OF LIFE

By

Venkata S. Chilakapati  
M.D., Andhra Medical College, 1999  
M.P.H., Western Kentucky University, 2001  
Internal Medicine Residency, Michael Reese Hospital, 2007  
Cardiology Fellowship, University of Louisville, 2010

A Dissertation Approved on

April 14, 2010

by the following Dissertation Committee:

---

Dr. Carlton Hornung, Ph.D., M.P.H. (Chair)

---

Dr. Matthew Williams, M.D.

---

Dr. Paul McKinney, M. D.

---

Dr. Richard Wilson, D.H.Sc., M.P.H.

---

Dr. Sumanth Prabhu, M. D., F.A.C.C., F.A.H.A.

## DEDICATION

This dissertation is dedicated to my beloved professor, mentor, and role model,

Dr. Carlton A. Hornung, Ph.D., M.P.H.,

who has given me invaluable support and encouragement.

## ACKNOWLEDGEMENTS

I would like to thank my dissertation committee chair, Dr. Carton A. Hornung, for his guidance, patience and thorough human understanding. I would also like to thank other committee members, Dr. Matthew Williams, Dr. Paul McKinney, Dr. Richard Wilson, and Dr. Sumanth Prabhu, for their input and assistance over the past few years of my research. I would like to thank my wife, Rajani, for her patience and understanding. Without her cooperation and support I wouldn't have accomplished this in my life. I am guilty of not spending enough time with my two year old daughter, Shivani, while pursuing cardiology fellowship and PhD simultaneously. Her cute words and beautiful smile provided lot of refreshment and ample energy to complete this task. I would like to thank my sister, Lyka, for her support and encouragement. She is always there and made me to stick to my goals and objectives. I would like to thank Dr. Neal Weintraub, Dr. Andrew Burger and Dr. Robin Vandivier for their support and encouragement. I would like to thank Ms. Tammi A. Thomas for her help and administrative support. I would like to thank Mr. Mark Evans, who assisted me for proof reading and formatting.

## ABSTRACT

### TECHNOLOGY-DRIVEN HEART FAILURE MANAGEMENT SYSTEM: CHANGES IN FUNCTIONAL CAPACITY AND QUALITY OF LIFE

Venkata S. Chilakapati

April 14, 2010

**Background:** Heart failure treatment guidelines emphasize the importance of daily weight monitoring. To support this practice, the Health Care Financing Administration Heart Failure Demonstration Project (2000) was designed to evaluate the effect of technology driven monitoring of daily weight and symptoms in elderly HF patients (NYHA class II to IV). This sub-study is conducted to assess the effect of a technology driven HF monitoring system on clinically meaningful change in functional capacity and quality of life (QOL).

**Methods:** This is a randomized, multi-centered, controlled clinical trial in which Medicare beneficiaries with a history of hospitalization within one year were randomized to standard care or standard care + a technology driven heart failure monitoring system. Primary end points were clinically meaningful change in functional capacity [6-minute walk distance (6MWD) or 6-minute work (6MW)] and QOL [Minnesota Living with Heart Failure Questionnaire (MLHF)] after 6-months of follow-up. Regression to the mean effects were estimated and adjusted according to the Edwards-Nunnally (EN) method. Clinically meaningful change is then defined in terms of the standard error of measurement (SEM) criterion.

**Results:** Two hundred eighty four patients from three centers in USA were randomized. Mean age was 74 +/- 9 yrs, 58% were male, 82% were Caucasians, and New York Heart Association class II (31%), III (59%), and IV (10%). The change in functional capacity in terms of 6WD was 42 m, by 6MW was 3668 kg/m of work, and change in QOL in terms of MLHF total score by 6.0, physical dimension score by 4.0, and emotional dimension score by 2.0 from baseline. Body mass index, Left ventricular Ejection Fraction, beta-blocker usage and intra-cardiac device (ICD) implantation were the best clinically relevant predictors of change in functional capacity. Technology driven heart failure management did not result in a clinically meaningful incremental benefit in functional capacity or in QOL.

**Conclusions:** There is no significant clinically meaningful benefit in functional capacity or in QOL from technology driven HF monitoring system in NYHA class II-IV elderly HF patients. Further, monitoring HF patients increased outpatient care resource utilization and costs, and was associated with a significantly poorer QOL.



## TABLE OF CONTENTS

	PAGE
ACKNOWLEDGEMENTS .....	iv
ABSTRACT .....	v
LIST OF TABLES .....	xiv
LIST OF FIGURES .....	xviii
INTRODUCTION .....	1
Heart Failure .....	1
Disease Impact .....	1
Heart Failure and Clinical Status Changes .....	3
American Heart Association Prognostic Classification .....	4
Heart Failure and Functional Capacity .....	5
Heart Failure and Quality Of Life .....	6
Heart Failure and Clinical Management .....	7
Heart Failure and Telemonitoring Systems .....	9
Telemonitoring Acceptability and Patient Satisfaction .....	10
Observational Studies of Telemonitoring and Hospitalizations and	
Readmission Rates .....	11
Randomized Studies of Telemonitoring and Hospitalizations and Mortality .....	12
Technology-Driven Home Monitoring of Patients With Heart Failure .....	13
Rationale for the Study .....	13
Significance of the Study .....	15

LITERATURE REVIEW .....	17
Functional Capacity .....	17
Measuring Functional Capacity .....	18
6-Minute Walk Test .....	18
6-Minute Work.....	21
Quality of Life.....	22
Definition .....	22
Measures of Quality of Life .....	23
Minnesota Living With Heart Failure Questionnaire .....	24
Clinically Meaningful Difference .....	25
Definition .....	25
Measuring Clinically Meaningful Difference .....	26
Distribution-Based Approach .....	26
Anchor-Based Approach.....	27
Integration of Distribution- and Anchor-Based Methods .....	28
Clinically Meaningful Change in 6-Minute Walk Distance .....	30
Clinically Meaningful Change in MLHF Score.....	31
METHODOLOGY .....	34
Hypotheses .....	34
Primary Hypothesis 1 .....	34
Primary Hypothesis 2.....	34
Secondary Hypothesis 1 .....	35
Secondary Hypothesis 2.....	35

Study Design .....	35
Study Setting .....	36
Eligibility Criteria .....	36
Subject Identification and Enrollment .....	37
Randomization .....	39
Intervention .....	39
Data Collection .....	43
Study Endpoints .....	44
Data Entry .....	44
Data Quality Monitoring .....	44
Data Management .....	45
Sample Size and Power Calculations .....	46
Data Analysis .....	47
Handling of Missing Data .....	47
Adjustment for Confounders .....	49
Generalizability of Study Sample .....	49
Preliminary Analyses .....	50
Analysis of Endpoints .....	51
Adjustment for Regression to the Mean .....	51
Standard Error of Measurement .....	52
Assessment of the Effect of Technology-Driven Heart Failure Management System .....	53

RESULTS .....	55
Sample Description .....	55
Estimation of Skewness and Kurtosis for Baseline Target Variable	
Distribution .....	61
Comparison of Baseline Characteristics .....	63
Description of Target Variables After 6 Months .....	63
Correlations Among Changes in Target Variables After 6 Months.....	66
Regression to the Mean.....	67
Assessment of Regression to Mean .....	67
Regression to Mean for Change in 6-Minute Walk Distance Over 6	
Months .....	67
Regression to Mean for Change in 6-Minute Walk After 6 Months .....	69
Regression to Mean for Change in MLHF Score After 6 Months.....	70
Adjustment of Change in Target Variables From Baseline to 6 Months	
to RTM.....	70
Determination of Clinically Meaningful Change in Target Variables From	
Baseline to 6 Months.....	72
Clinically Meaningful Change in 6MWD for Standard Care Vs. Technology-	
Driven Disease Management .....	74
Clinically Meaningful Change in 6MW From Baseline to 6 Months for	
Standard Care Vs. Technology-Driven Disease Management.....	77
Clinically Meaningful Change in QOL From Baseline to 6 Months for	
Standard Care Vs. Technology-Driven Disease Management.....	79

Clinically Meaningful Change in Physical Domain Score From Baseline to 6 Months for Standard Care Vs. Technology-Driven Disease Management.....	81
Clinically Meaningful Change in MLHF Emotional Domain Score From Baseline to 6 Months for Standard Care Vs. Technology-Driven Disease Management.....	83
Clinically Meaningful Change in 6MWD to Predict Clinically Meaningful Similar Directional Change in MLHF Score .....	85
Clinically Meaningful Change in 6MWD to Predict Clinically Meaningful Similar Directional Change in MLHF Physical Domain Score.....	85
Clinically Meaningful Change in 6MWD to Predict Clinically Meaningful Similar Directional Change in MLHF Emotional Domain Score.....	86
Clinically Meaningful Change in 6MW to Predict Clinically Meaningful Similar Directional Change in MLHF Score .....	88
Clinically Meaningful Change in 6MW to Predict Clinically Meaningful Similar Directional Change in MLHF Physical Domain Score.....	89
Clinically Meaningful Change in 6MW to Predict Clinically Meaningful Similar Directional Change in MLHF Emotional Domain Score.....	91
Agreement Between NYHA Class and 6MWD in Assessing the Severity of Heart Failure .....	91
Predictors of Clinically Meaningful Change in 6MWD .....	93
Predictors of Clinically Meaningful Change in 6MW .....	96
Predictors for Clinically Meaningful Change in MLHF Total Score .....	104
Clinical Outcomes.....	108

Mortality.....	109
DISCUSSION .....	113
Summary of Results .....	113
Importance of Clinically Meaningful Change .....	113
Need for Adjustment of Regression to the Mean.....	114
Clinically Meaningful Change in Functional Capacity and Quality of Life.....	115
Effect of Technology-Driven Heart Failure Management on Functional Capacity .....	115
Effect of Technology-Driven Heart Failure Management on Quality of Life ...	116
Effect of Technology-Driven Heart Failure Management on Clinical Outcomes .....	116
Agreement Between Objective Assessment of Functional Capacity and Classification of Heart Failure NYHA Criteria .....	117
Prediction of Change in Quality of Life by Change in Functional Capacity .....	118
Predictors of Change in Physical and Emotional Dimension Scores by Change in Functional Capacity .....	120
Predictors of Clinically Meaningful Change for Functional Capacity .....	121
Predictors of Change in Quality of Life.....	122
Comparison of Existing Studies.....	122
Strengths of the Study .....	123
Limitations of the Study.....	125
Post-Hoc Power Analysis.....	126
Conclusions and Implications .....	128

Future Directives for Research .....	129
REFERENCES .....	131
APPENDIX.....	157
CURRICULUM VITAE.....	160

## LIST OF TABLES

TABLE	PAGE
1. NYHA Functional Classification of Heart Failure .....	3
2. AHA Prognostic Classification .....	4
3. Tests of Normality for Baseline Target Variables.....	60
4. Estimated Skewness and Kurtosis.....	62
5. Baseline Characteristics .....	64
6. Target Variables After 6 Months for Standard Care vs. Technology-Driven Management.....	65
7. Mean Change in Functional Capacity and Quality of Life Over 6 Months .....	66
8. Partial Correlations Between Target Variables at Baseline and Change From Baseline to 6 Months.....	71
9. Standard Error of Measurement for Target Variables.....	73
10. Standard Error of Measurement for Target Variable After Anchoring by NYHA Class.....	74
11. Clinically Meaningful Change in 6MWD in Standard Care vs. Technology-Driven Disease Management.....	75
12. Clinically Meaningful Change in 6MWD for Standard Care vs. Technology-Driven Disease Management After Stratification .....	76
13. Clinically Meaningful Change in 6MW for Standard Care vs. Technology-Driven Disease Management.....	77



14. Clinically Meaningful Change in 6MW for Standard Care vs. Technology-Driven Disease Management After Stratification .....	78
15. Clinically Meaningful Change in MLHF Total Score for Standard Care vs. Technology-Driven Disease Management .....	79
16. Clinically Meaningful Change in MLHF Total Score for Standard Care vs. Technology-Driven Disease Management After Stratification.....	80
17. Clinically Meaningful Change in MLHF Physical Domain Score for Standard Care vs. Technology-Driven Disease Management.....	81
18. Clinically Meaningful Change in MLHF Physical Domain Score for Standard Care vs. Technology-Driven Disease Management After Stratification.....	82
19. Clinically Meaningful Change in MLHF Emotional Domain Score for Standard Care vs. Technology-Driven Disease Management.....	83
20. Clinically Meaningful Change in MLHF Emotional Domain Score for Standard Care Group vs. Technology-Driven Disease Management After Stratification .....	84
21. Clinical Meaningful Change in 6MWD vs. Clinically Meaningful Change in MLHF Total Score .....	86
22. Clinically Meaningful Change in 6MWD vs. Clinically Meaningful Change in Physical Dimension Score.....	87
23. Clinically Meaningful Change in 6MWD vs. Clinically Meaningful Change in Emotional Dimension Score.....	88
24. Clinically Meaningful Change in 6MW vs. Clinically Meaningful Change in MLHF Score.....	89

25. Clinically Meaningful Change in 6MW vs. Clinically Meaningful Change in Physical Dimension Score.....	90
26. Clinically Meaningful Change in 6MW vs. Clinically Meaningful Change in Emotional Dimension Score.....	92
27. Agreement Between NYHA Class and Functional Capacity .....	92
28. Case Processing Summary for Predictors of Clinically Meaningful Change in 6MWD.....	95
29. Model Fitting Information for Predictors of Clinically Meaningful Change in 6MWD.....	96
30. Likelihood Ratio Tests for Predictors of Clinically Meaningful Change in 6MWD.....	97
31. Parameter Estimates for Predictors of Clinically Meaningful Change in 6MWD .....	98
32. Classification Accuracy of Predictors for Clinically Meaningful Change in 6MWD.....	99
33. Case Processing Summary for Predictors of Clinically Meaningful Change in 6MW.....	100
34. Model Fitting Information for Predictors of Clinically Meaningful Change in 6MW.....	101
35. Likelihood Ratio Tests for Predictors of Clinically Meaningful Change in 6MW ...	102
36. Parameter Estimates for Predictors of Clinically Meaningful Change in 6MW .....	103
37. Classification Accuracy for Predictors of Clinically Meaningful Change in 6MW.....	104

38. Case Processing Summary for Predictors of Clinically Meaningful Change in	
MLHF Total Score .....	106
39. Model Fitting Information for Predictors of Clinically Meaningful Change in	
MLHF Total Score .....	106
40. Likelihood Ratio Tests for Predictors of Clinically Meaningful Change in	
MLHF Total Score .....	107
41. Parameter Estimates for Predictors of Clinically Meaningful Change in MLHF	
Total Score.....	107
42. Classification Accuracy for Predictors of Clinically Meaningful Change in	
MLHF Total Score .....	108
43. Clinical End Points for Technology-Driven Management Group vs. Standard	
Care Group.....	110

## LIST OF FIGURES

FIGURE	PAGE
1. Adjustment of regression to the mean and calculation of clinically meaningful change.....	53
2. Sample attrition.....	56
3. Frequency distribution of 6-minute walk distance at baseline.....	57
4. Frequency distribution of 6-minute work at baseline.....	57
5. Frequency distribution of MLHF scores at baseline.....	58
6. Frequency distribution of MLHF physical dimension scores at baseline.....	58
7. Frequency distribution of MLHF emotional dimension scores at baseline.....	59
8. Scatter plot of change in 6-minute walk distance against baseline.....	68
9. Scatter plot of change in 6-minute work against baseline.....	68
10. Scatter plot of change in MLHF total score (QOL) against baseline.....	69
11. Kaplan-Meier Curve for Death from all Causes in 6 months.....	111
12. Kaplan-Meier Curve for Death from all cardiovascular causes.....	111
13. Kaplan-Meier Curve for Death from Heart Failure.....	112

## CHAPTER I

### INTRODUCTION

#### Heart Failure

Heart failure (HF) is a clinical syndrome in which the cardiac pump fails to pump an adequate amount of blood to meet systemic requirements, either due to reduced cardiac contractility (systolic failure) or impaired cardiac relaxation and filling (diastolic dysfunction) or both (Johnson, Parker, & Patterson, 2002). Initially, HF triggers various neuro-hormonal mechanisms to compensate. However, as HF progresses, the body's compensatory and maladaptive mechanisms become imbalanced, which leads to clinical deterioration (Johnson et al., 2002). Common signs and symptoms of HF include fatigue, shortness of breath, difficulty breathing (especially at night, when lying down, or during physical exertion), cough, weight gain (from fluid retention), and swelling of the feet and ankles (National Heart, Lung, and Blood Institute [NHLBI], 2004).

#### Disease Impact

HF is a major, growing public health problem in the United States affecting 4.9 million people (American Heart Association [AHA], 2009). Each year, 550,000 new cases are diagnosed (AHA, 2009). The increasing population of older Americans and the prolongation of the lives of cardiac patients by modern therapeutic strategies have led to the growing incidence of HF (Kannel & Belanger, 1991). Among Americans aged 65 or older, the incidence of HF is increasing and is approaching 1 in 100 (AHA, 2005, 2009).

HF accounts for 12 to 15 million U.S. office visits and 6.5 million hospital days each year, with an average length of stay of 7.7 days for each admission (O'Connell & Bristow, 1994). Between 1979 and 2005, hospital discharges for HF increased by 157% (AHA, 2005), and between 1979 and 2009, they rose by 171% (AHA, 2009). However, despite advanced therapies and technology, the mortality rate of HF remains high. The number of total reported deaths from HF was 287,000 in 1995 and 292,000 in 2005 (AHA, 2009). The overall death rate for HF in 2005 was 52.3 per 100,000 (AHA, 2009). Moreover, HF is mentioned on the death certificate for one in eight deaths (AHA, 2009).

In spite of the commitment to healthcare in the United States, the economic burden of HF is staggering. In 2005, almost 4.8 million Americans were afflicted with HF, accounting for 15% of the total cost of heart disease (AHA, 2006). In 2005, the total direct and indirect cost of HF in the United States was equal to \$27.9 billion (AHA, 2005), and the estimated direct and indirect cost for 2009 is \$37.2 billion (AHA, 2009). Almost 75% of costs associated with a typical HF-related hospitalization accumulate within the first 48 hours (O'Connell & Bristow, 1994). Hospital charges for HF management are approximately \$10,000 per discharge based on a mean length of stay of 5.5 to 6.5 days (AHA, 2008).

According to the Medicare program, Center for Medicare and Medicaid Services (CMS) expenditures for HF in 1991 were higher than those for cancer and those for myocardial infarction (Massie & Shah, 1997). HF is the most common Medicare diagnosis-related group (DRG), and more Medicare dollars are spent on diagnosis and treatment of HF than on any other diagnosis (Massie & Shah, 1997). HF costs accounted

for 37% of Medicare spending (AHA, 2008), and in 1999, Medicare reported paying \$3.6 billion for care of HF (AHA, 2005).

### Heart Failure and Clinical Status Changes

The severity of HF can be classified symptomatically by using a scheme such as the New York Heart Association's (NYHA) functional classification, which groups patients according to the amount of effort needed to produce HF symptoms (Criteria Committee NYHA, 1964, p. 114; see Table 1).

Table 1

#### NYHA Functional Classification of Heart Failure

NYHA grading	<u>Characterisitics</u>	Metabolic equivalent <sup>a</sup>
Class I	No limitations. Ordinary physical activity doesn't cause undue fatigue, dyspnea, or palpitations (asymptomatic LV dysfunction)	>7
Class II	Slight limitation of physical activity. Ordinary physical activity results in fatigue, palpitation, dyspnea or angina pectoris (mild congestive heart failure [CHF])	5
Class III	Marked limitation of physical activity. Less than ordinary physical activity leads to symptoms (moderate CHF)	2–3
Class IV	Unable to carry on any physical activity without discomfort. Symptoms of CHF present at rest (severe CHF)	1.6

<sup>a</sup>Metabolic equivalent is defined as the resting VO<sub>2</sub> for a 40-year-old 70 kg man. 1 MET = 3.5 ml O<sub>2</sub>/min/kg body weight.

NYHA classification is not a good prognostic indicator, because symptom severity can rise and fall despite constant pharmacologic and non-pharmacologic treatments (Hunt, Baker, & Chin, 2001).

#### American Heart Association Prognostic Classification

The staging system to define the prognosis of patients with HF was developed by the AHA in 1999 (Hunt et al., 2001; see Table 2).

Table 2

#### AHA Prognostic Classification

Stage	Explanation
A	Patients with risk factors for the development of structural heart disease or for the development of overt HF
B	Presence of structural heart disease (e.g., myocardial infarction or left ventricular dysfunction) without symptoms of HF
C	Patients with structural heart disease with current or prior symptoms of HF amenable to therapy
D	HF refractory to conventional treatment requiring ventricular assist device, transplantation, or palliative care

Note. Prognosis worsens if disease progresses from Stage A through D



## Heart Failure and Functional Capacity

The hallmark of HF is exercise intolerance and activity restriction most commonly due to impaired breathing and fatigue. These symptoms are debilitating and result in low functional capacity, progressive physical disability, hospitalization, medical management, and follow-up. Moreover, as HF advances, functional capacity deteriorates.

The most popular clinical exercise tests, in order of increasing complexity, are stair climbing, 6-minute walk test (6MWT), shuttle-walk test, cardiac stress test (e.g., Bruce protocol), and cardiopulmonary exercise test (Wasserman, Hansen, Sue, Casaburi, & Whipp, 1999; Weisman & Zeballos, 1994). A recent review of functional walking tests concluded that “the 6MWT is easy to administer, better tolerated, and more reflective of activities of daily living than the other walk tests” (Solway, Brooks, Lacasse, & Thomas, 2001, p. 256).

The 6MWT is a simple and noninvasive measure of the distance covered in consecutive 30-meter or 100-foot laps over a 6-minute period (American Thoracic Society [ATS], 2002). The ability to walk the distance strongly and independently predicts morbidity and mortality (Bittner et al., 1993; Guyatt, Sullivan, et al., 1985; Lipkin, Scriven, & Poole-Wilson, 1986; Poole-Wilson, 2000). However, distance covered doesn't take into account differences in bodyweight that are known to influence exercise capacity (Carter et al., 2003). Carter et al. stated that the 6-minute distance times bodyweight product (6-minute walk work expressed in kilograms per meter of work) is an improved outcome measure for estimating functional capacity in patients with chronic obstructive pulmonary disease (COPD; Carter et al., 2003).

## Heart Failure and Quality Of Life

Poor physical condition, emotional distress, and frequent hospital admissions contribute to decreased quality of life (QOL) among individuals with HF (Blyth et al., 1997; Cafagna, Ponte, & Burri1, 1997), which can lead to social impairments, depression, and psychological disorders (Majania et al., 1999; Steptoe, Mohabir, Mahon, & McKenna, 2000). Depression has been reported in 24%–42% of patients suffering from congestive HF (Havranek, Ware, & Lowes, 1999; Skotzko et al., 2000). Vaccarino et al. (2001) and Jiang et al. (2001) found that depression reduced functional status and increased readmission rates and mortality in patients with CHF. Depressed patients required a greater number of inpatient hospital days and at 1 year, showed greater mortality (Freedland et al., 1991).

QOL deteriorates more rapidly and seriously for patients with chronic HF than for those with other chronic disease like arthritis and COPD (Stewart et al., 1989). Gott et al. (2006) investigated the predictors of quality of life in patients with HF and found that QOL for older people with HF could be described as challenging and difficult. Among 542 people with HF under the age of 60, the following factors were predictive of reduced QOL: being female, being in an older age group, showing evidence of depression, being NYHA Class III and IV, having two or more comorbidities, and being from a low socioeconomic group (Gott et al., 2006). These factors could help clinicians to identify those at risk of reduced QOL and appropriately target interventions.

Poor QOL may have a negative effect on compliance with medical treatment and behavioral regimens (Jaarsma et al., 1999) and, thus, result in further impairment of exercise tolerance, prognosis, and QOL. To interrupt this vicious cycle, an integrated and

comprehensive HF intervention program comprising adequate pharmacologic treatment and careful home monitoring is needed.

### Heart Failure and Clinical Management

The standard of care for HF includes pharmacologic and non-pharmacologic measures. Pharmacologic management of HF includes diuretics, angiotensin-converting enzyme (ACE) inhibitors, angiotensin receptor blockers, beta-blockers, digitalis, and vasodilators. Diuretics, especially loop diuretics (e.g., Furosemide), are used to promote the elimination of excess sodium and water by the kidneys (Brater, 1997). ACE inhibitors are recommended for patients with left ventricular dysfunction and have been shown to reduce mortality from HF; improve HF symptoms, exercise tolerance, and left ventricular ejection fraction; and reduce emergency room visits and hospitalizations (Cohn et al., 1986; Pitt et al., 1991; Yusuf et al., 2000).

Angiotensin receptor blockers may be an alternative for patients who are unable to tolerate ACE inhibitors (Yusuf et al., 2003). Hydralazine and nitrates in combination are effective after load and preload reducing agents used in ACE-intolerant patients (Cohn et al., 1991). Beta blockers have been shown to slow the progression of HF and to reduce hospitalization and mortality by blocking sympathetic stimulation (Hjalmarson et al., 2000; Packer et al., 1999). Digoxin is a weak inotrope, whose main effect is neurohormonal modulation; it has been shown to reduce symptoms, improve physical function and QOL, and decrease the rate of hospitalization in patients with HF, but has not been shown to have an effect on mortality (Digitalis Investigation Group, 1997).

Inotropic agents like dobutamine or milrinone are useful in select patients with acute exacerbations of hypotensive HF or shock (Felker et al., 2001). Spironolactone, an

aldosterone antagonist, showed mortality benefit when added to standard therapy for patients with HF (in NYHA Classes III–IV; Pitt et al., 1999). Intravenous nitroglycerin or nitroprusside are recommended for the management of patients with acute pulmonary edema (“ACC/AHA Guidelines,” 2001).

Patients with HF are more prone to sudden cardiac death. Internal cardioverter defibrillators (ICDs) improve longevity for survivors of cardiac arrest; patients with sustained ventricular tachycardia, inducible ventricular tachycardia, and left ventricular ejection fraction (LVEF) less than 30% following a myocardial infarction (Gollob & Seger, 2002; Moss, Zareba, Hall, Klein, & Wilber, 2002); and dilated cardiomyopathy patients with LVEF < 30% and NYHA Class II (Kadish et al., 2004).

Left ventricular assist devices (LVAD) are used for patients with cardiogenic shock who are unresponsive to inotropic therapy and intra-aortic balloon counter pulsation, or they may be used as a bridge therapy for patients with cardiac transplantation (Delgado et al., 2002; Rose et al., 2001). Cardiac transplantation is reserved for otherwise healthy patients with end-stage congestive heart failure (CHF) with severely impaired function despite optimal medical therapy (Hunt, 1998).

Nonpharmacologic measures include regular exercise for patients with stable HF (Pina et al., 2003), restriction of dietary sodium intake, restriction of excess fluid intake, smoking cessation, restriction of dietary fat intake, abstinence from alcohol and illicit drugs, and treatment of lipid disorders (“ACC/AHA Guidelines,” 2001; NHLBI, 2004). Despite these measures, the prevalence of chronic HF is increasing, and the situation of patients with HF will deteriorate unless new management strategies are developed (Cleland, 2000).

The effectiveness of multidisciplinary nonpharmacological approaches for improving outcomes for patients with HF is under evaluation (McAlister, Stewart, Ferrua, & McMurray, 2004). Considering the barriers to healthcare access related to funding and geographic location, development of remote monitoring models for delivering care must be considered (McAlister et al., 2004). Remote monitoring models can include communication technology, such as telemonitoring (transfer of physiological data such as blood pressure, weight, electrocardiographic (ECG) details, and oxygen saturation through telephone or digital cable from home to healthcare provider) or regular structured telephone contacts between patients and healthcare providers, which may or may not include the transfer of physiological data (McAlister et al., 2004).

#### Heart Failure and Telemonitoring Systems

Telemonitoring permits home monitoring of patients using special telecare devices in conjunction with a telecommunication system. Telecare is increasingly used by care providers in various specialties to support chronically ill patients at home using existing telecommunications systems (Ahring, Ahring, Joyce, & Farid, 1992; Billard et al., 1991). Telemonitoring systems can provide diagnostic information, which can be transmitted manually or automatically, to allow experts to evaluate patients one or more times per day, or even to continuously monitor them. In view of the rising costs of hospital care, limitation of functional capacity, and QOL in controlled settings, interest in telecare has grown, and rapid advances in communication and technology have facilitated development of patient-friendly telecare equipment (Louis et al., 2003).

### Telemonitoring Acceptability and Patient Satisfaction

Six observational studies (Deering, Baines, Christianson, & Milner, 2002; de Lusigman et al., 2000; de Lusigman, Meredith, Wells, Leatham, & Johnson, 1999; Knox, Mueller, Vuckovic, & Acker, 2002; Wang, Yu, Chau, & Lamm, 2002; Williams, Keiler, Sprang, & Mehan, 1998) and one randomized study (Woodend et al., 2002) assessed patients' acceptance and compliance with telemonitoring (see also Louis et al., 2003). De Lusigman et al. (1999) measured acceptability and QOL scores while monitoring blood pressure (BP), heart rate (HR), and weight and video conferencing weekly. De Lusigman et al. concluded that patients using telemonitoring showed good acceptability and improved QOL scores after 3 months.

De Lusigman et al. (2000) monitored heart rate, respiratory rate, ECG, and temperature for a 2-hour period and concluded that telemonitoring had good acceptability and a reliability of 95%. Deering et al. (2002) monitored weight and symptoms for an average follow-up of 5.3 months and concluded that telemonitoring improved patient satisfaction for 94% of patients. Knox et al. (2002) monitored weight, pulse rate, and symptoms for 8 weeks and found that telemonitoring reduced clinic visits for optimization of beta-blocker therapy.

Williams et al. (1998) monitored weight and symptoms for 4 months and concluded that telemonitoring systems showed high compliance and patient satisfaction in 86% of cases. In a randomized, controlled trial comparing telemonitoring with usual care, Woodend et al. (2002) found high patient satisfaction and improvement in QOL through video conferencing with a nurse and telemonitoring of vital signs and ECG.

### Observational Studies of Telemonitoring and Hospitalizations and Readmission Rates

Twelve nonrandomized studies assessed the effect of telemonitoring on hospitalizations and readmission rates (see also Louis et al. 2003). Shah, Der, Ruggerio, Heidenrich, and Massie (1998) found a 50% reduction in cardiovascular admissions and 80% acceptability after 8.5 months of patient education, automated reminders, and weight and blood pressure monitoring for NYHA Class II–IV patients. Cordisco, Beniaminovitz, Hammond, and Mancini (1999) found reduced emergency room (ER) visits and hospitalization rates after 12 months of telemonitoring weight and symptoms. Chrysegolos, Gemme, Coleman, Cheng, & Meyer (1999) monitored weight, BP, and ECG in NYHA Class III patients for 6 months and concluded that telemonitoring reduced readmission rates and ER visits. After 12–18 months, telemonitoring of weight, BP, HR, and O<sub>2</sub> saturation led to a significant decrease in the number of admissions and length of hospital stays in NYHA Class III patients (Bondmass, Malhotra, Castro, & Avitall, 1999).

Heidenrich, Ruggerio, and Massie (1999) monitored weight, BP, and HR and provided patient education through weekly telephone calls for 7.4 months and found significantly reduced hospitalizations for NYHA Class II–III patients. Roglieri et al. (1997) used vital sign monitoring and patient education as an intervention for 3 months and found reduced 30- and 90-day readmission rates, length of stay, and number of emergency room visits. Macropoulos and Selna (2002) monitored weight and symptoms for 12 months and found reduced admissions rates.

Kesinger, Gilani, and Jennison (2002) used electronic home monitoring of weight, symptoms, and medication compliance for NYHA Class III–IV for 7 months and found a

50% reduction in hospitalizations. Ertle and Litman (2002) found a 73% reduction in inpatient costs after 6 months of weight and symptom monitoring. Scalvini, Zanelli, Volterrani, and Benigno (2002) monitored symptoms and ECG for 12 months and found a significant reduction in hospitalizations with no difference in mortality compared to a control group. Lapworth and Dibiase (2000) monitored weight and symptoms of NYHA Class III patients for 8 months and found a 75% reduction in hospitalizations and a 47% reduction in length of hospital stay.

#### Randomized Studies of Telemonitoring and Hospitalizations and Mortality

Five randomized studies assessed the effect of telemonitoring on hospitalizations and mortality rates. Bondmass, Benatar, Castro, and Avitall (2001) found that telemonitoring weight, HR, BP, and O<sub>2</sub> saturation resulted in fewer readmissions and reduced length of hospital stay for a group of telemonitored patients, compared to a group of patients given nurse care visits. Jerant, Azari, and Nesbitt (2001) performed two-way video conferencing integrated with an electronic stethoscope for 6 months with NYHA Class II patients and found an 86% reduction in readmission charges and an 84% reduction in the telemonitored group.

Goldberg et al. (2002) telemonitored weight and symptoms of NYHA Class III–IV patients for 6 months and found no difference in readmission rates compared to patients receiving standard care; however, there was a significant reduction in mortality for women under age 65. Johnson, Wheeler, Deuser, and Sousa (2000) found no difference in cost of healthcare after 10–12 months for patients who used video consultation with a nurse and patients who received standard care. Massie, West, Van Ostaeyen, and Salbalvaro (2001) found no significant difference in outcomes between



groups that did and did not receive telephonic monitoring of weight vital signs and symptoms.

### Technology-Driven Home Monitoring of Patients With Heart Failure

The Health Care Financing Administration (HCFA) Heart Failure Demonstration Project Study Group of 2000 conducted a trial to assess the impact of a technology-driven HF home monitoring system on clinical and economic outcomes of Medicare beneficiaries who were recently hospitalized for HF. The study was conducted in Billings, Montana (rural); Louisville, Kentucky; and Philadelphia, Pennsylvania. Study subjects were randomized and received either standard HF care (as described above) or standard HF care supplemented by home telemonitoring.

The purpose of this substudy is to (a) evaluate the effect of home telemonitoring systems on functional capacity (6-minute walk distance [6MWD] and 6-minute work [6MW]) and QOL (QOL score); (b) determine clinically meaningful difference in 6MWD or 6MW and Minnesota Living With Heart Failure (MLHF) scores; and (c) determine the correlation between clinically meaningful differences in 6MWD or 6MW and QOL (MLHF score) in patients with chronic HF.

### Rationale for the Study

Congestive heart failure (CHF) is a major health problem with increasing incidence and a poor prognosis. It is often accompanied by multiple comorbidities, disability, depression, and cognitive impairment and leads to deterioration in functional capacity and QOL (Stott, 2002; Wielenga et al., 1997). Efforts must be made to develop novel strategies to reduce the rising cost of care for patients with HF without compromising the standard of care.

Louis et al. (2003) conducted a systematic review of telemonitoring studies and concluded that telemonitoring might have an important role as part of a strategy for delivery of healthcare for patients with HF. Several multidisciplinary programs and remote monitoring models were evaluated for effectiveness in terms of a statistically significant difference in outcomes, but none of the studies made definitive conclusions about the clinical significance of such strategies for improving outcomes like functional capacity and QOL (Louis et al., 2003). Thus, integration of technology-driven HF monitoring into standard care and assessing its impact in terms of clinically significant improvement in QOL and functional capacity is imperative to designing sound multidisciplinary disease management models.

Because of the complexity and cost associated with implementing many disease management programs for patients with HF, there is a need for a simpler, less personnel-intensive, more cost-effective approach to the longitudinal care of these patients. Despite the high prevalence of HF, there have been very few studies of the impact of CHF on functional capacity and QOL, suggesting the need for epidemiological studies on functional capacity and QOL (Davis et al., 1999).

In the context of preventive and therapeutic cardiology, functional capacity and QOL are important outcome measures for evaluating the effectiveness of treatment strategies and the course of a disease (Bullinger, 1997; Wenger, 1989). Therefore, functional capacity and QOL are major goals in the context of preventive and therapeutic cardiology. It is, therefore, imperative to test the impact of technology-driven therapeutic modalities of HF management on functional capacity and QOL of patients with CHF. With increased rates of patient adherence, compliance, and acceptance of telemonitoring

technology to monitor HF symptoms (Capomolla, Pinna, La Rovere, Maestri, & Ceresa, 2004; de Lusignan et al., 1999, 2000), various investigators are in the process of developing and evaluating several telemonitoring systems. For example, the HCFA Demonstration Project Study Group investigated the effectiveness of a technology-driven home telemonitoring system for Medicare beneficiaries with chronic HF.

### Significance of the Study

The purpose of this study is to evaluate the effect of a technology-driven home telemonitoring system on functional capacity using 6MWD or 6MW and on QOL using MLHF. Small differences in functional capacity and QOL may be statistically significant but clinically unimportant. Therefore, it is necessary to determine the smallest difference considered worthwhile or clinically important (Hays & Woolley, 2000). The purpose of this study is to determine clinically meaningful differences in the 6MWD or 6MW and MLHF scores of patients with chronic HF.

The MLHF scale is superior and more responsive to changes in QOL compared to the Chronic Heart Failure Questionnaire (CHQ) and the General Health Survey Short Form 12 and 36 (SF-12 and SF-36) instruments (Bennet et al., 2002; Ni, Toy, Burgess, & Wise, 2000). Therefore, the purpose of this study is to determine changes in MLHF scores that are indicative of clinically meaningful change in the QOL of patients with HF.

In a randomized control study, Guyatt, Townsend, Keller, Singer, and Nogradi (1991) concluded that 6MWD better correlated with formal measures of QOL for patients with chronic lung disease. Carter et al. (2003) stated that 6-minute distance times bodyweight product (6MW expressed in kilograms per meter) was an improved outcome measure to estimate the functional capacity of COPD patients. The purpose of this study

is to determine the relationship between clinically meaningful change in 6MWD or 6MW and MLHF scores of patients with CHF.

## CHAPTER II

### LITERATURE REVIEW

CHF is a major health problem with increasing incidence and a poor prognosis. It is often accompanied by multiple comorbidities, disability, depression, and cognitive impairment and leads to deterioration in functional capacity and QOL (Stott, 2002; Wielenga et al., 1997). Therefore, functional capacity and QOL should be included as important outcome measures when developing disease management models for patients with HF. Davis et al. (1999) stated that there was a lack of data on the effect of disease monitoring systems on functional capacity and QOL for patients with HF.

#### Functional Capacity

Functional capacity refers to the performance of, or the capacity to perform, regular daily activities (Leidy, 1994). Such activities require the integrated efforts of the heart, lungs, and circulation to deliver oxygen to the metabolically active muscle mass to perform work (Gibbons et al., 1997). The capacity of an individual to perform work is defined by the maximal oxygen consumption ( $VO_{2\text{ max}}$ ), the product of cardiac output (CO), and arteriovenous oxygen ( $AVO_2$ ) difference at exhaustion (Rowell, 1988). Functional capacity is often expressed in metabolic equivalents (METs); 1 MET represents resting energy expenditure, which is approximately  $3.5 \text{ mL O}_2 \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$  (Fleg & Lakatta, 1988; Rowell, 1988). Functional capacity is affected by age, sex, the

presence of disease, and medications that affect aerobic exercise (Fleg & Lakatta, 1988; Pollock et al., 1998).

### Measuring Functional Capacity

In the hospital setting, functional capacity can be directly measured by determining  $\text{VO}_2$  max or can be estimated using the highest treadmill or stationary cycle ergo meter work rate achieved (Jerome et al., 2000). Peak oxygen uptake (peak  $\text{VO}_2$ ) is a reference parameter, or gold standard, in the assessment of functional capacity of patients with HF; however, the procedure for measuring peak oxygen uptake is complex, invasive, and expensive. Outside the hospital setting, the most widely used methods for assessing functional capacity, the 6MWT and shuttle walk test, are simple, noninvasive, and inexpensive. For the 6MWT, patients are instructed to walk as far as they can in 6 minutes. In the shuttle walk test, patients are pressured to meet multiple deadlines through audio cassette beeps. The 6MWT is self-paced and less likely than the shuttle walk test to push patients beyond their endurance or through angina or other pain (Paul, 2003).

### 6-Minute Walk Test

The 6MWT is a practical, simple test that requires a 100-foot hallway but no exercise equipment or advanced technician training. The test measures the distance that a patient can quickly walk on a flat, hard surface in a period of 6 minutes, known as the 6MWD. Functional capacity, or exercise capacity, is determined by the integrated response of multiple systems, including the pulmonary, cardiovascular, hematopoietic, neuromuscular, and metabolic (ATS, 2002). The 6MWT does not provide specific

information on the function of each organ and system involved in the exercise mechanism; instead, it serves as a surrogate estimate of the collective mechanism.

Most activities of daily living are performed at a sub maximal level. The self-paced 6MWT assesses the sub maximal level of functional capacity. Often, patients do not achieve maximum exercise capacity during the 6MWT; instead, they choose their own intensity of exercise and are allowed to stop and rest during the test (ATS, 2002). The test has been shown to be useful in assessing prognosis and exercise capacity in several studies (Bittner et al., 1993; Cahlin, Mathier, Semigran, Dec, & Disavo, 1996; Guyatt, Thompson, et al., 1985; Roul, Germain, & Bareiss, 1998; Zugck et al., 1998). In clinical practice, the 6MWT is used for measuring the response to medical interventions for patients with moderate to severe heart or lung disease (ATS, 2002).

There is a moderate correlation between distance walked during the 6MWT and peak  $\text{VO}_2$  ( $r = 0.70$ ,  $P < .001$ ) and close correlation between distance walked during the shuttle walk test and peak  $\text{VO}_2$  ( $r = 0.83$ ,  $P < .001$ ; Morales et al., 1999). Solway et al. (2001) performed a qualitative review of measurement properties of the shuttle walk test and 6MWT and concluded that exercise performed during the shuttle walk test was similar to a symptom-limited, maximal, incremental treadmill test. The shuttle walk test is more difficult to administer, requires more equipment, and is less reflective of activities of daily living (ATS, 2002; Solway et al., 2001). The 6MWT is easy to administer, better tolerated, and more reflective of activities of daily living. Therefore, currently, the 6MWT is the test of choice when administering a functional walk test for clinical or research purposes (ATS, 2002; Solway et al., 2001).

As a test of submaximum exercise capacity with high validity and minimal potential for cardiovascular problems (Demers et al., 2001), the 6MWT mimics daily activity (Solway et al., 2001) and is more relevant to both doctor and patient than the maximum exercise test (ATS, 2002). Demers et al. (2001) investigated the reliability, validity, and responsiveness of the 6MWT for patients with HF and concluded that the 6MWT was highly reproducible (intraclass correlation coefficient [ICC] = 0.90), moderately and inversely correlated to NYHA functional class ( $r = -0.43$ ,  $p = 0.001$ ), and weakly inversely correlated to cumulative QOL (MLHF score;  $r = -0.26$ ,  $p = 0.0001$ ).

In an observational study, O’Keffe, Lye, Donnellan, and Carmichael (1998) concluded that the 6MWD was more responsive to deterioration than improvement of HF symptoms. There are several factors that influence the 6MWT: a shorter 6MWD may be associated with shorter height (short legs), older age, greater body weight, female gender, impaired cognition, shorter walking corridor (more turns), acute or chronic lung diseases, cardiovascular diseases, chronic degenerative diseases, and musculoskeletal injuries or deformities (Paul, 2003). Factors associated with a longer 6MWD are taller height (longer legs), male gender, high motivation, and medications increasing aerobic exercise. The effect of these factors must be considered when drawing conclusions about functional capacity. The strongest indication of the 6MWT is for measuring the response to medical interventions of patients with moderate to severe heart or lung disease (ATA, 2002). The 6MWT has also been used as a measure of functional status of patients as well as a predictor of morbidity and mortality (ATS, 2002).



## 6-Minute Work

Distance covered during the 6MWT doesn't take into account differences in bodyweight that are known to influence exercise capacity. Chuang, Lin, and Wasserman (2001) investigated the calculation of the product of bodyweight and walking distance as an alternative method for assessing functional capacity. This calculation accounts for bodyweight differences, and estimates work and energy expenditure through the formula Force x Distance (Chuang et al., 2001). Carter et al. (2003) stated that 6MWD (in meters) x bodyweight (in kilograms), or 6MW expressed in kilograms per meter of work, was an improved outcome measure to estimate functional capacity for COPD patients.

For COPD patients, 6MW yielded higher correlation coefficients than 6MWD when correlated with pulmonary function indices, including lung diffusion for alveolar ventilation, DLco ( $r = 0.6$  vs.  $0.46$ ,  $p = 0.0001$ ), forced expiratory volume in one minute ( $r = 0.52$  vs.  $0.38$ ,  $p = 0.0001$ ), forced vital capacity ( $r = 0.48$  vs.  $0.38$ ,  $p = 0.0001$ ), vital capacity ( $r = 0.48$  vs.  $0.40$ ,  $p = 0.0001$ ), and cycle ergometry gas exchange indices, including work (in watts;  $r = 0.79$  vs.  $0.59$ ,  $p = 0.0001$ ), peak oxygen uptake ( $r = 0.81$  vs.  $0.54$ ,  $p = 0.0001$ ), peak minute ventilation ( $0.59$  vs.  $0.46$ ,  $p = 0.0001$ ), and peak tidal volume ( $r = 0.57$  vs.  $0.43$ ,  $p = 0.0001$ ). The ROC curve demonstrated that 6MW had a significantly larger calculated area under the curve ( $0.782$  vs.  $0.708$ ,  $p < 0.05$ ) than 6MWD (Carter et al., 2003).

The pathophysiology of HF is entirely different from that of COPD. Functional capacity of patients with HF is determined by degree of debilitation from frequent decompensation. Fluctuations in weight may be very rapid and unpredictable. However, the applicability and correlation of 6MW to HF indices has not been evaluated. Even if a

statistically significant difference exists between correlation coefficients and the area under the curves of 6MWD and 6MW, the clinical significance of those differences has not been determined.

### Quality of Life

#### Definition

The World Health Organization (WHO, 1946) defined health as “a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity”. Health status is typically used to represent the patient’s subjective (or perceived) appraisal of his or her state of physical and mental health (Revicki et al., 2000). The concept of QOL lacks clarity and uniform definition. Health status, functional status, and QOL are three concepts often used interchangeably when referring to health.

Gill and Feinstein (1994) stated that QOL incorporates patient values, judgments, and preferences and that these are subjective experiences, states, and perceptions of one’s overall well-being, including aspects of the physical, psychological, social, economic, and political environment (Revicki et al., 2000). Guyatt (1993) used the term health-related QOL (HRQOL) because many widely valued aspects of life, such as income, freedom, and quality of the environment, are not generally considered health-related.

Testa and Simonson (1996) defined HRQOL as the “physical, psychological and social domains of health, seen as distinct areas that are influenced by a person’s experiences, beliefs, expectations and perceptions” (p. 835). QOL is a broad term that comprises subjective domains as assessed by patients—physical functioning, psychological functioning, and social functioning—and objective domains as assessed by

healthcare and social support personnel and measured by laboratory or diagnostic tests—psychopathology (CID-10/DSM-IV-TR), socioeconomic status, and social support (Rui et al., 2005). Therefore, the goals of disease management for patients with chronic HF should be directed at improving symptoms, stabilization or improvement in functional abilities, and improved comfort for the duration of life (Wenger, 1989).

For asymptomatic patients with HF, the management strategies should aim to improve ventricular function in order to improve long-term survival and eliminate interference with QOL, which will improve compliance and enhance long-term outcome advantages (Wenger, 1989). Majania et al. (1999) found that subjective domains were highly correlated with a global sense of well-being, and objective domains were less correlated with a global sense of well-being. QOL reflects subjective perceptions of harmony within the body and between a patient and the external world (Katsching, 1997).

### Measures of Quality of Life

Measures of QOL fall into two categories: (a) generic measures, which attempt to provide a summary of a patient's perceived QOL, and (b) disease-specific measures, which focus on problems associated with particular disease states and patient groups (Guyatt, 1993). Generic measures can be standardized and applied to a wide variety of illnesses to allow for comparisons (Dempster & Donnelly, 2000). Early QOL research used psychological well-being scales, including the Affect Balance Scale by Bradburn (1969), the Quality of Well-Being Scale by Kaplan and Anderson (1988), and the Psychological General Well-Being Index by Dupuy (1984). Later research included the Sickness Impact Profile by Gilson et al. (1975); the Nottingham Health Profile by Hunt et al. (1980) and the Rand SF-36 Health Status Profile by Ware et al. (1993).

Generic measures lack the range, sensitivity, and flexibility to deal with a particular illness (Dempster et al., 2000). Therefore, disease-specific QOL instruments were developed. Disease-specific instruments are more responsive and more sensitive to changes (Spertus, Winder, Dewhurst, Deyo, & Stephan, 1994). Some examples of disease-specific instruments to measure QOL in patients with HF are the Chronic Heart Failure Questionnaire (Guyatt et al., 1989), the Minnesota Living with Heart Failure Questionnaire (Rector, Kubo, & Cohn, 1992), and the Kansas City Cardiomyopathy Questionnaire (Green, Porter, Bresnahan, & Spertus, 2000).

Although numerous instruments are available to measure QOL, the most commonly used instruments in HF research are generic instruments (SF-12 and SF-36) and disease-specific instruments (CHQ and MLHF). Bennett et al. (2002) conducted a study to compare the psychometric properties of these instruments and concluded that all instruments were reliable and valid; however, CHQ and MLHF were more sensitive and responsive to changing conditions than generic measures were. Considering the feasibility of administering these instruments and their responsiveness to changes in QOL, MLHF is superior to CHQ, SF-12, and SF-36 (Bennet et al., 2002; Ni et al., 2000).

#### Minnesota Living With Heart Failure Questionnaire

The MLHF questionnaire is a disease-specific instrument designed to measure the effects of HF and treatments for HF on QOL (Rector et al., 1992). It is composed of 21 items and three subscales measuring the following dimensions: (a) the physical functioning dimension (eight items), (b) the emotional functioning dimension (five items), and (c) an overall QOL score (all 21 items). Eight separate items not assessing an

underlying construct or dimension of HQOL measure social and economic impairment for patients due to HF and are part of the overall score.

To measure the effects of symptoms, functional limitations, and psychological distress on an individual's QOL, the MLHF questionnaire asks each person to indicate using a 6-point Likert scale (0 to 5) to what extent each of the 21 facets prevent them from living as they desire. The questionnaire results in a total score between 0 (no impairment) and 105 (severe impairment), which represents the physical, socioeconomic, and psychological impairments of HF that affect QOL.

Factor analysis identifies two components on the questionnaire—physical and emotional. The associations between patients' overall ratings of how much HF prevented them from living as they wanted ( $r = 0.8$ ,  $p < 0.01$ ) and their NYHA classification ( $r = 0.60$ ,  $p < 0.01$ ) suggest that the MLHF score is a valid representation of patient impairment (Rector et al., 1992). The test–retest reliability of the MLHF score was  $r = 0.93$  ( $p < 0.001$ ) for the total score,  $r = 0.89$  ( $p < 0.001$ ) for the physical dimension subscore, and  $r = 0.88$  ( $p < 0.001$ ) for the emotional dimension subscore (Rector & Cohn, 1992).

### Clinically Meaningful Difference

#### Definition

Jaeschke, Singer, and Guyatt (1989) defined minimal important difference as

the smallest difference in score in the domain of interest which patients perceive as beneficial and which would mandate, in the absence of troublesome side-effects and excessive cost, a change in the patient's management. (p. 407)

Small numerical differences in mean scores might give statistically significant results when large sample sizes are used, but statistical significance is not equivalent to

clinical significance (Osoba et al., 1998). Currently, there are no guidelines or standards to help physicians determine whether or not the treatment provided to patients with CHF results in a clinically meaningful change in their functional capacity and QOL. The purpose of this study is to define the level of change in functional capacity (measured by 6MWD or 6MW) and QOL (measured by MLHF) that is clinically meaningful for patients with CHF.

### Measuring Clinically Meaningful Difference

There are two approaches for determining meaningful change: (a) the distribution based-approach and (b) the anchor-based approach.

#### Distribution-Based Approach

Distribution-based approaches to determining clinically meaningful change are based on the statistical characteristics of the obtained sample. To estimate the effect size (Cohen, 1988; Kazis et al., 1989) or standard error of measurement (SEM; McHorney & Tarlov, 1995; Wywich, Tierney, & Wolinsky 1999), distribution-based methods use the empirical distribution of a measure and its psychometric characteristics. Therefore, distribution-based methods rely only on statistical and psychometric properties of a measure.

Distribution-based measures fall under three broad categories. First, measures based on statistical significance evaluate change in relation to the probability that change occurred by random variation. Examples are paired  $t$ -statistic (Husted et al., 2000) and growth curve analyses (Speer & Greenbaum, 1995). Second, measures based on sample variation evaluate change in relation to sample variation. Some examples are effect size (Cohen, 1988; Kazis, Anderson, & Meenan, 1989), standardized response mean (Stucki

et al., 1995), and responsiveness statistic (Guyatt, Bombardier, & Tugwell, 1986). The final category of measures includes those based on the measurement precision of the instrument. Some examples are SEM (Wywich et al. 1999) and reliable change index (Jacobson & Truax, 1991).

According to Crosby, Kolotkin, and Williams (2003),

the most promising distribution-based measures for establishing clinically meaningful change are those based on the measurement precision of the instrument (Standard Error of Measurement, SEM and Reliable Change Index, RCI). Because these measures quantify the amount of error inherent in the instrument and the amount of random variation that can be expected in repeated administrations. In addition, they are not influenced to a large degree by variability in the sample at baseline (as is the effect size), variability of the observed change (as are the Standardized Response Mean, SRM and the Responsiveness Statistic), or the sample size (as are the *t*-statistic and growth curve analysis). These precision measures can be used to establish cutoffs based on a desired confidence level. (p. 400)

#### Anchor-Based Approach

Anchor-based methods define a clinical standard for comparison using the patient's or physician's perception of change as an external anchor to estimate the clinical meaningfulness of corresponding change (Jaeschke et al., 1989). Anchor-based approaches have been used to determine clinically meaningful change via cross-sectional and longitudinal methods. Various cross-sectional methods described in the literature include comparison to disease-related criteria (Deyo et al., 1992; Johnson, Goldman, & Orav, 1995), comparison to non-disease-related criteria (Testa & Lenderking, 1992; Testa & Simonson, 1996), preference ratings (Llewellyn-Thomas, 1996) and comparison to known populations (Jacobson & Truax, 1991).

One of the limitations of cross-sectional comparisons is that groups are likely to differ in many relevant variables besides HRQOL. Samsa et al. (1999) suggested statistically controlling these variables using regression methods and calculating effect

sizes based on adjusted mean differences. The most commonly used anchor-based approaches for establishing clinically meaningful change in longitudinal studies focus on global ratings of change (Jaeschke et al., 1989; Stucki et al., 1995), prediction of prognosis of future events (Mossey & Shapiro, 1982), and changes in disease-related outcomes (Kolotkin, Crosby, & Williams, 2002).

According to Crosby et al. (2003),

Longitudinal anchor-based methods are more directly linked with change when compared to cross-sectional methods because the former are more directly linked with change. Hence longitudinal anchor-based methods are preferable to cross-sectional methods. In considering these longitudinal methods, patient global ratings are especially well suited for assessing patient's perception of change and are recommended for that purpose. However, when using global ratings, it is important to assess the reliability of these ratings. Clinician global ratings and longitudinal disease-related measures of outcome are the most suitable methods of determining meaningful change from the clinical perspective. (p. 399)

#### Integration of Distribution- and Anchor-Based Methods

It's not clear whether it is better to use anchor-based or distribution-based methods for determining clinically meaningful change. There have been some attempts to integrate anchor- and distribution-based approaches. Jacobson and Truax (1991) proposed that individuals should be considered improved or deteriorated only when they meet both the anchor- and distribution-based criteria for change. Another study by Cella et al. (2002) to determine clinically meaningful change on the Functional Assessment of Cancer Therapy–Lung (FACT-L) questionnaire described that clinically meaningful change determined by anchor-based methods (e.g., differences between clinically distinguishable groups) was confirmed by distribution-based methods (e.g., effect size, SEM). The agreement across these methods was high, with kappa ranging from 0.71 to 1.0. Individuals with the greatest impairments at baseline have the greatest opportunity for improvement compared to individuals with less extreme scores.



Failure to take RTM (RTM) into account may lead to erroneously concluding that those with severe impairments have shown clinically meaningful improvement, when much of that change is due to RTM (Barnett, Van Der Pol, & Dobson, 2005). Cella et al. (2002), Guyatt and Jaeschke (1997), Hays and Woolley (2000), Juniper, Guyatt, Willan, and Griffith (1994), and Ware et al. (1993) reported that when a patient is improving, a smaller amount of change may be considered clinically important than when the patient is worsening. Therefore, direction of change should be accounted for when defining clinically meaningful change.

Kolotkin et al. (2002) described a method to determine clinically meaningful change in obesity-specific QOL using the combined information from anchor-based (weight loss) and distribution-based (SEM corrected for RTM) methods, taking into account baseline impairment and direction of change. Speer (1992) mentioned that while defining clinically meaningful difference, if regression to the mean is not present, then the Jacobson and Truax (1991) method is more appropriate.

Jacobson and Truax (1991) proposed a two-step method. The first step is to define cutoff points that separate a functional population from a dysfunctional population. For example, Cutoff A specifies the functional population as those with post-therapy scores that are two standard deviations (SDs) or more from the pretreatment mean. The second step compares individual's change from pre- to post-therapy to the standard error (SE) of measurement of the outcome ( $\pm 1.96$  SE), referred to as reliable change index (RCI).

These two steps classify individuals into four categories: recovered (individual has passed Cutoff A and RCI in the positive direction), improved (passed RCI in the

positive direction but not Cutoff A), unchanged (passed neither criterion), or deteriorated (passed RCI in the negative direction). If RTM is present, the Edwards–Nunnally (EN) method is more appropriate than the Jacobson and Truax method (Edwards et al., 1978). The EN method addresses regression to the mean by shrinking pretherapy scores toward the pretherapy mean using the reliability of the measure. The estimated true score is then placed at the center of a confidence interval so that estimates can be made of the significance of post-therapy change, or two SEs from the adjusted center.

Crosby, Kolotkin, and Williams (2004) designed an integrated method that combined information from anchor- and distribution-based methods using data aggregated from weight-loss studies. A total of 1,476 weight-loss study participants were evaluated at baseline and at 6 months using the Impact of Weight on Quality of Life–Lite (IWQOL–Lite). Then, Crosby et al. used the EN method and corrected SEM for regression to the mean.

#### Clinically Meaningful Change in 6-Minute Walk Distance

Using distribution-based methods, G. Claire, Martin, Joseph, Tamiza, and Stephen (2005) conducted a study of pulmonary hypertension patients to explore minimally important difference (MID) for 6MWD and SF-36 domains. They concluded the MID for 6MWD was 39 m, and for SF-36, the MID for physical functioning was 11 m; role physical, 21; social functioning, 16; and vitality, 13.

Redelmier, Bayoumi, Goldstein, and Guyatt (1997) conducted a study of patients with chronic lung disease and found that the 6MW was significantly correlated with patients' ratings of their walking ability relative to other patients ( $r = 0.59$ , 95% confidence interval [CI]: 0.54 to 0.63). Distances must differ by 54 m for the average

patient to stop rating themselves as about the same and start rating themselves as either a little bit better or a little bit worse (95% CI: 37 to 71 m).

Cooper (2001) conducted a study of patients with chronic pulmonary disease and considered a 54-meter change in 6MWD to be clinically meaningful based on the conclusions of Redelmier et al. (1997). Perera, Mody, Woodman, and Studenski (2006) conducted a secondary analysis of data from an observational study and clinical trials of community-dwelling older adults and subacute stroke survivors. Perera et al. calculated the effect size using distribution-based methods (Cohen, 1988; Kazis et al., 1989; Testa, 1987) and concluded that small meaningful change in 6MWD was 20 m and substantial change was 50 m. In these trials, RTM and reliability of the instrument were not taken into account when defining meaningful change in 6MWD.

#### Clinically Meaningful Change in MLHF Score

According to Rector (2005), clinically meaningful refers to an effect that would prompt patients and physicians to consider using a treatment if the risk and costs were acceptable. So far, there has been no gold standard to determine cutoff values for changes in QOL in terms of improvement or deterioration. Rector et al. (1995) found that an improvement of 5 points in the MLHF score was sufficient to be clinically effective for a majority of patients to take a medication that had no side effects or costs. According to Bennet et al. (2002), a change of 5 points was the minimum clinically meaningful change based on the mean change in the MLHF score observed in a group of patients who rated their change in overall condition as  $\pm 2$  or  $\pm 3$  on a  $-7$  to  $+7$  scale.

Cohen (1988) provided benchmarks to serve as a guide for interpreting effect size, where effect size = (mean at baseline – mean at follow-up)/standard deviation at baseline.

For small effects, the benchmark was 0.2; for moderate effects, 0.5; and for large effects, 0.8. Based on a comprehensive review of the literature, an effect size of 0.2 should serve as an appropriate definition of the minimum clinically important difference (Samsa et al., 1999). Using effect size calculations, conclusions of clinically meaningful change at the individual level may be influenced by the heterogeneity of the sample. In addition, the effect size doesn't take into account the variability of change, the test-retest reliability of the instrument, and RTM.

Wyrwich, Nienaber, et al. (1999) suggested that a change of 1 SEM was clinically meaningful. The 1-SEM (baseline) threshold is equivalent to 0.71 SEM change, and approximately 52% of the area under the normal curve is between  $Z = -0.71$  and  $Z = 0.71$  (Wyrwich, Nienaber, et al. 1999). To decide if an improvement or deterioration has occurred at the individual level, a 51% level of certainty may be applied (Donaldson & Moinpour, 2002). A higher level of certainty (90%–95%) is needed to compare statistically significant differences at the group level but not at the individual level. Therefore, at the individual level, change over time exceeding these limits of the  $\pm$  SEM baseline are likely to reflect a minimum important change rather than merely demonstrating measurement error or trivial fluctuations of chronic disease measures (Wyrwich, Nienaber, et al., 1999).

SEM is an estimate of error for use in interpreting individual test scores. A test score is an estimate of a person's true test performance. Using a reliability coefficient and the test's SD, this value may be calculated as follows:

$$\text{SEM} = \text{SD} \cdot [\sqrt{1 - r}] \quad (1)$$

Where  $SD$  = the standard deviation for the test, and  $r$  = the reliability coefficient for the test.

The higher a test's reliability coefficient, the smaller the test's SEM is. The larger the SEM, the less reliable the test is. There are no studies that define clinically meaningful change in MLHF score using clinical anchors such as functional capacity or NYHA class.

## CHAPTER III

### METHODOLOGY

#### Hypotheses

For patients previously hospitalized with HF and with continued symptoms of HF, addition of technology-driven HF management to standard care may be associated with clinically meaningful effects when compared to standard medical care alone.

#### Primary Hypothesis 1

Research hypothesis: Technology-driven HF management is associated with clinically meaningful effect on functional capacity measured in terms of 6MWD or 6MW after 6-months follow-up in elderly patients with HF

Null hypothesis: Technology-driven HF management is not associated with clinically meaningful effect on functional capacity measured in terms of 6MWD or 6MW after 6-months follow-up in elderly patients with HF.

#### Primary Hypothesis 2

Research hypothesis: Technology-driven HF management is associated with clinically meaningful effect on QOL measured in terms of MLHF score after 6-month follow-up in elderly patients with HF

Null hypothesis: Technology-driven HF management is not associated with clinically meaningful effect on QOL measured in terms of MLHF score after 6-month follow-up in elderly patients with HF.

### Secondary Hypothesis 1

Research hypothesis: There is a correlation between clinically meaningful change in functional and the QOL in elderly patients with HF after 6-months follow-up.

Null hypothesis: There was no correlation between clinically meaningful change in functional capacity and the QOL in elderly patients with HF after 6-months follow-up.

### Secondary Hypothesis 2

Research hypothesis: Clinically meaningful change in QOL is responsive to clinically meaningful change in functional capacity after 6-months follow-up in elderly patients with HF.

Null hypothesis: Clinically meaningful change in QOL is not responsive to clinically meaningful change in functional capacity after 6-months follow-up in elderly patients with HF.

### Study Design

The HCFA Demonstration Project Study Group designed a randomized trial for adding a home telemonitoring system to the standard of care for Medicare beneficiaries with a primary diagnosis of chronic HF and with continued HF symptoms and who were discharged from an acute care hospital within the previous 6 months. The study was designed to assess the incremental effect of adding a home telemonitoring system to standard care on health outcomes, resource utilization, and costs and processes of care for patients with HF.

Subjects were enrolled and randomized between January 2001 and January 2003. To examine whether the initial effectiveness of telemonitoring system would persist or decay over time, a novel method of randomization was implemented. Six months after

initial randomization and follow-up, subjects who had been randomly assigned to follow a home monitoring system were rerandomized to either continue with the home telemonitoring system or standard care alone. Both groups were followed for an additional 6 months.

The end points of the HCFA study included cardiovascular rehospitalization or cardiovascular death within 6 months following study randomization. The secondary end points explored potential continued benefit and attenuation of benefit over 12 months after randomization. Using the HF demonstration project data from the HCFA Demonstration Project Study Group, the effect of adding technology-driven HF management to standard care on functional capacity (assessed using 6MWD or 6MW) and QOL (assessed using MLHF score) for patients with chronic HF was assessed in this sub-study. The purpose of this study was to determine the clinically meaningful change in 6MWD or 6MW and QOL (MLHF scores) and to explore the correlation among these variables (if any).

#### Study Setting

Participants for the main study were recruited from the followings sites: St. Vincent's Hospital and Health Center in Billings, Montana; Jewish Hospital University of Louisville School of Medicine in Louisville, Kentucky; and University of Pennsylvania Health System, Philadelphia, Pennsylvania.

#### Eligibility Criteria

Participants must be Medicare beneficiaries aged 65 years or older who were discharged from an acute care hospital with a primary diagnosis of HF (DRG 127) within the preceding 6 months. The following criteria also had to be met: (a) evidence of HF



(left ventricular ejection fraction < 50% documented on echocardiography, radionuclide ventriculography, or cardiac catheterization) and currently experiencing symptoms of HF; (b) ability to stand for 20 s without holding the wall; (c) bodyweight less than 400 pounds; (d) ability to speak English or Spanish (required for the use of the home telemonitoring system); (e) ability to complete a 6MWT; (f) ability to provide witnessed, written informed consent for all aspects of the study, including permission to provide or access the clinical, health status, and medical service utilization data required by the study protocol prior to enrollment and randomization.

The following subjects were excluded from the study: (a) those participating in other ongoing HF research or demonstration studies or having prior experience with the Alere DayLink Monitoring System; (b) those lacking a phone line or accessible phone jack in the home, as required by the home telemonitoring system; (c) Medicare managed care participants; (d) patients with chronic dialysis, a serum creatinine level over 3.0 in the past 30 days, and anticipated initiation of dialysis within 6 months; (e) patients who had a heart transplant; (f) patients with uncorrected thyroid disease; and (g) patients with end-stage or terminal illness such as metastatic malignancy or AIDS with anticipated life expectancy less than 6 months in addition to HF, or a score less than or equal to 25 on a minimal status exam.

#### Subject Identification and Enrollment

At each study site, clinical research coordinators (CRCs) identified physicians who were interested in recruiting patients for the study. After obtaining patients' permission, CRCs reviewed current and past admission records. All personnel had HIPAA training and strictly adhered to HIPAA procedures. Patients hospitalized with a

primary diagnosis of HF who were Medicare beneficiaries aged 65 or older were identified. After reviewing their hospital records, CRCs identified potentially eligible discharged patients who had been recently hospitalized. After screening eligible patients, CRCs contacted attending physicians to confirm eligibility and obtain the attending physician's consent to recruit the potential participants.

After confirming eligibility and obtaining the attending physician's consent, CRCs contacted potential participants by phone and explained the study. Patients who expressed interest were scheduled for a baseline visit within 2 weeks to obtain written, signed, dated, and witnessed informed consent, at which time the data coordination center (DCC) was contacted to begin the randomization process and baseline data collection.

Upon enrollment, the following procedures were performed.

1. Study coordinators obtained the results of MUGA, ECHO, or angiography, if performed, from the past year. If such a test had not been performed, the study coordinator requested the physician prescribe one to be completed within 6 weeks of enrollment, especially to assess ejection fraction.
2. Study coordinators obtained demographic data, medical history, medications, physician assessment data at baseline (screening visit), and baseline laboratory data including serum sodium, creatinine, blood urea nitrogen (BUN), total bilirubin, potassium, digoxin trough level, and ECG results. If laboratory data were not available, they were obtained within 2 weeks of baseline visit.
3. Patients completed a baseline 6MWT, and study coordinators documented the results.

4. Patients completed QOL questionnaires with the assistance of the coordinators.
5. Study coordinators provided participants with HF diaries and educational material, explained the importance of self-monitoring and daily recording of weight in the HF dairy, emphasized the importance of bringing the diary to each physician visit, and scheduled a 6-month follow-up visit.

### Randomization

Eligible subjects who provided written informed consent were randomized using a 1:1 ratio to the control group (standard HF care) or the intervention group (standard HF care plus home telemonitoring). Randomization was performed at each site to control practice-site-specific bias or confounding effects. Randomization was performed prior to collection of baseline data as soon as eligibility was confirmed to avoid selection bias. Each enrolled participant was assigned a unique randomization number and unique study number. These numbers were maintained in a central log at the DCC. All enrollment and randomization logs were maintained by local CRCs.

### Intervention

The HCFA HF Demonstration Project Study Group replicated as much standard medical care as possible among participants. Therefore, standard HF care was provided as background therapy for participants in both groups. The only difference between the two study groups was the addition of a technology-driven HF home telemonitoring system to the standard of care for the intervention group.

The Alere DayLink Monitoring System was chosen for this study because Alere was the first company to offer in-home telemonitoring and the system was used in the

Weight Monitoring in Heart Failure (WHARF) trial (Goldberg et al., 2002). Upon completion of the randomization process, the study coordinator faxed patient demographics, history, and medications at the time of enrollment to Alere Medical Inc. At the time of the study, Alere was licensed to do business in California and had its corporate headquarters in San Francisco with satellite centers in 25 U.S. states. It supported more than 30,000 patients, and its call center was staffed with trained personnel 7 days a week, 365 days a year, from 7:00 a.m. to 5:00 p.m. Pacific Standard Time.

After receiving patients' information, Alere shipped the devices on the same day so that patients would receive them the next day at their homes. Installation of the device was simple, and instructions were mailed to participants along with the devices. By following instructions, participants connected the three color-coded cables to the required jacks. After the unit was connected to a power source, the set-up was completed automatically. The black box (console) prompted participants throughout the installation process by instructing them to press yes or no keys on the console. By following the prompts, the patient or family member could complete the set-up. After completing the set-up, the device was ready to use.

Per protocol, twice daily, the console visibly and audibly prompted participants to step onto a scale for weighing. After recording their weight, the console compared patients' weight to the preset weight, or dry weight. Then the console prompted participants to answer several preset questions. After answering the last question, the console acknowledged the patient with a thank you and automatically transmitted the information over the phone line to the telemonitoring station.

Call center nurses performed twice-daily monitoring (morning and afternoon) of patients' weight and symptoms using the data transmitted through the telemonitoring system. When a patient's weight was acceptable and the patient was not symptomatic, monitoring continued without changes to therapy, and a summary report was sent to the appropriate physician's office. If a patient's weight was not within an acceptable range (i.e., it increased by more than 3–5 pounds in 2 days), the patient did not weigh in, or the patient was symptomatic, call center nurses called the patient or a nurse practitioner to alert them to the situation. The call center nurses would obtain further information from the patient and adjust the medical therapy or suggest an ER visit or hospital admission.

In addition to periodic assessments, the intervention group participants received health education through weekly telephone calls from call center nurses. The nurses placed special emphasis on medication changes, changes in health status or condition, and the need for lifestyle modifications. Patients' physicians selected the telemonitoring and therapy parameters per individual patients' clinical status.

All enrolled participants received standardized HF education and teaching materials irrespective of whether they were randomized to standard HF care or standard HF care plus telemonitoring. All patients received a comprehensive educational handbook, instructions about HF from registered nurses, and a diary with standardized instructions to record daily weight and symptoms.

Per study protocol, participants were required to attend a follow-up visit at 6 months for data collection for primary and secondary endpoints and for rerandomization. Participants were also required to attend a follow-up visit at 12 months for data collection for tertiary endpoints. At the 6-month protocol visit, after data collection, participants

who used the telemonitoring System were then rerandomized to continue telemonitoring or to standard care. If the participant was rerandomized to standard care therapy, he or she was provided with a digital home scale and instructed to obtain and record daily weight and HF symptoms in a diary.

Due to the nature of the intervention, the intervention group was not blinded. This might have introduced some bias into the trial; however, the study group agreed that the design would maximize consistent HF baseline care for every patient independent of randomization, which would allow the group to determine the true clinical effect of telemonitoring.

To minimize bias, research staff was blinded to the study hypothesis and group assignment. They obtained all data through the telephone (details are described in the Data Collection section). Similarly, event adjudication was determined by a clinical endpoint committee, which was blinded to the study group.

#### Group 1: Standard Care

Group 1 participants received a CHF standard care plan individualized through routine patient care provided by a primary care physician or cardiologist or both. Patients were given diaries to chart their daily weight and instructed to take the diary with them to each physician visit.

#### Group 2: Standard Care Plus Technology-Driven HF Monitoring

In addition to receiving the same CHF standard care that Group 1 received, Group 2 participants were given a telemonitoring System to monitor their weight and symptoms daily through telephonic transmission of data. The data were used to alert caregivers or physicians and to reinforce HF management.

### Data Collection

All data were collected through standardized procedures and data collection instruments. Data coordination was performed at the Center for Epidemiology and Clinical Investigation Sciences, University of Louisville, under the direction of Dr. Carlton Hornung. CRCs at each site collected baseline data and administered the 6MWT at the time of enrollment and at 6 and 12 months after enrollment. Regardless of group, the patients were contacted every 6 weeks (i.e., at study weeks 6, 12, 18, 30, 36, and 42) by telephone to obtain supplemental data on social support assessment and any ER visits or hospitalizations.

At 12 weeks and 36 weeks, additional data on QOL were collected through completion of QOL questionnaires via telephone. Research staff who were blinded to the study hypotheses and the intervention group participants, collected all of the telephone survey data. The data collected included socio-demographic variables; social and family support; clinical status; physical, functional, and mental health status; subjective health rating; HRQOL; and satisfaction with care. After each telephonic contact, research assistants mailed additional diaries and a self-addressed stamped envelope to participants. After completion of 6 and 12 month protocol visits, research assistants reminded participants to return the diaries in the self-addressed envelope. Medicare utilization data were obtained through CMS service utilization files.

The study endpoints were measured through standardized instruments used in earlier and ongoing studies, as mentioned in chapter 2. Functional capacity was assessed using the 6MWT. QOL for patients with chronic HF was measured using the MLHF

questionnaire. Social support assessment was performed using the Krause–Marchetti survey.

### Study Endpoints

This sub-study was limited to a 6-month follow-up. The purpose of the sub-study was to evaluate the impact of adding telemonitoring to standard care on functional capacity (in terms of 6MWD or 6MW) and QOL (MLHF score) at the study end point.

### Data Entry

The HCFA HF Demonstration Project Study Group designed a customized computer-based system for data entry. The DCC created a separate database for each site with separate data entry files. Each site had access to its own data entry files but did not have access to the files at other sites. The site-specific CRCs collected and entered the data into electronic case report forms during each patient visit and telephone survey.

CRC data transmissions were stored in a temporary database, and integrity of the data was verified against source documents using the quality assurance keys developed by the DCC. Separate and secure databases were created on a server at each site and were password protected; only authorized personnel had access to the data. The DCC prepared a standard manual of operations to guide CRCs and each clinical site monitoring team in order to maintain the quality and integrity of the data.

### Data Quality Monitoring

Prior to the first month of study, standardized pre-study training sessions were held at the University of Pennsylvania for the principal investigators (PIs) and CRCs from all sites. These sessions were to train PIs and CRCs in all aspects of study protocol, methods, data collection, and data entry. The first line of data quality monitoring was



performed at the point of data collection. The second line was performed at the time of data entry. Only after completion of all pre-specified data fields were the files merged into a site-specific database.

The computer data entry system had a built-in data quality/integrity check system, which verified the integrity and quality of the data at each site. The DCC maintained an 800-number telephone hotline for CRCs to call with any questions with regard to data entry or editing. The DCC, project manager, and CRCs had weekly conference calls to identify and address data collection and quality problems. The director and the DCC conducted periodic reviews of the collected data and monitored the type and frequency of data entry errors, edits, and changes at each site.

The DCC merged the electronic case report forms, telemonitoring System data, and CMS data for each site separately and created individual and site-specific data files for monitoring the quality of data and tracking of errors at each site. If a high error rate occurred at any site, notices were sent to the clinical site monitoring team and CRCs. The integrity of the data was verified with source documents, and errors were rectified at each site.

### Data Management

The data coordination center was located in the Center for Epidemiology and Clinical Investigation Sciences at the University of Louisville (one of the sites of the project). The center had PhD-trained biostatisticians, epidemiologists, database management experts, a web master, and supporting personnel. Database management experts merged the data from individual sites and prepared a master data file for analysis.

### Sample Size and Power Calculations

In the initial protocol, power analysis calculations were performed based on the primary end point of reduction in hospital admission rates.

It was estimated that 175 patients per arm will be needed for the 6 month data collection period for a primary analysis, a total of 35% patients to detect a 40% reduction in cardiovascular hospital admission rates (two-sided alpha of 0.05, 80% power, assuming a 6-month readmission rate for the control group (Krumholz 1997) of 35% and 21% in the interventional arm. Assuming the 20% reduction in the sample size during the first 6 months, it was estimated that 220 patients per arm, for total of 440 participants who meet all eligibility criteria and consent to enroll in the study. (HCFA Heart Failure Demonstration Project, 2000)

According to a recent review published by ATS (2002), optimal reference equations for healthy population-based samples using standardized 6MWT methods were not yet available. Miyamoto et al. (2000) conducted a study in which the median 6MWD was approximately 580 m for 117 healthy men and 500 m for 173 healthy women. Another study reported a mean 6MWD of 630 m in 51 healthy older participants (Stevens et al., 1999).

Cardiac rehabilitation in patients with various heart diseases increased the 6MWD by a mean of 170 m, or 15% (Bittner et al., 1993). In an observational study of 45 older patients with HF, the smallest difference in 6MWD that was associated with a noticeable difference in the global rate of worsening was a mean of 43 m (O’Keffe et al., 1998). However, there have not been any studies in the literature that defined the degree of change in 6MWD or 6MW and QOL from a baseline to the point of measurement and that should be considered clinically meaningful or as demonstrating clinically significant improvement or deterioration. Also, no randomized studies have found any correlation between clinically meaningful improvement or deterioration in 6MWD or 6MWW correlated with and clinically meaningful improvement or deterioration in QOL.

The focus of this sub-study is to define clinically meaningful cutoffs and use those cutoffs to estimate the efficacy of technology-driven HF management systems in improving the functional capacity and QOL for patients with chronic HF. Given extremely uncertain estimates of the cutoff values for the study variables, it is not possible to offer a specific power calculation. After obtaining the data, a post-hoc power analysis may be performed. As the intervention is not invasive and carries no risk, according to the available data (as mentioned in chapter 2), there is no clinical rationale to say that the telephonic management of patients with HF is inferior to standard care.

#### Data Analysis

SPSS version 17.0 software was used for the data analysis. The primary analysis included all randomized patients using the principle of intent to treat (International Conference on Harmonization of Technical Requirement for Registration of Pharmaceuticals for Human Use, 1999). The randomized patients were excluded only upon objective detection of violation of eligibility criteria. In addition, the subset of participants who complied with the protocol were analyzed; compliance was determined by completion of at least 1 month of care under the treatment to which the participant was randomized, availability of the primary outcome, and absence of any major protocol violation.

#### Handling of Missing Data

Individual missing items can be a nuisance to statistical packages based on the assumption that balanced data are available. If records are deleted where individual items are missing, there may be both loss of power and bias in estimation. The reasons for

missing covariate data (e.g., refusal, problems with data collection, rehospitalization, and death) have been documented.

Many statistical methods may be used to estimate or impute the most likely values for the missing items. If a patient died prior to the assessment or was hospitalized and not available for assessment, then 0 m was assigned for 6MWD; however, some of the patients who were hospitalized with CHF exacerbation were able to walk 0 to 25 feet (5–10 m), which was considered significant impairment in functional capacity. Therefore, these patients were classified as members of a clinically meaningful deterioration in functional capacity group.

The MLHF questionnaire consists of 21 items and three subscales measuring the following dimensions: (a) the physical function dimension (eight items), (b) the emotional function dimension (five items), and (c) an overall QOL score (all 21 items). Eight separate items not assessing an underlying construct or dimension of HQOL measure social and economic impairments for patients due to HF and are part of the overall score. To measure the effects of symptoms, functional limitations, and psychological distress on an individual's QOL, the MLHF questionnaire asks participants to indicate to what extent each of 21 facets prevented them from living as they desired using a 6-point (0 to 5) Likert scale.

The MLHF questionnaire results in a total score between 0 (no impairment) and 105 (severe impairment), which represents the physical, socioeconomic, and psychological impairments that affect QOL. The 6-point score is divided into three quartiles: The scores 0 and 1 belong to the first quartile (mild impairment); 2 and 3, the second quartile (moderate impairment); and 4 and 5, the third quartile (severe

impairment). For each question, a score of 5 was assigned to those patients who died prior to assessment, and a score of 4 was assigned to those patients who were hospitalized and not available for assessment. In both cases, the patients were classified into a clinically significant deteriorated group.

If a participant filled in most of the questions on the MLHF questionnaire except for a few (not more than three), then the missing data were estimated using the multiple means imputation method (Curran et al., 1998). The objective of imputation is to replace the missing data with estimated values to preserve the relationships between items and to reflect, to the extent possible, the most likely true value. The dropouts that were unrelated to outcomes were ignored in the analysis (Heyting, Tolboom, & Essers, 1992).

#### Adjustment for Confounders

Although participants were randomly assigned to intervention and control groups, there was a possibility of unequal distribution of patients due to potential confounders. Adjustment for covariates associated with the outcome is essential (Altman & Dore, 1991; Hauck, Anderson, & Marcus, 1998; Senn, 1994). Therefore, baseline variables that were correlated with outcomes were included in multivariable models in the comparison of two groups.

Potential variables are socio-demographic factors (age, gender, race/ethnicity, education); heart failure severity and case mix (ejection fraction, NYHA class, number of co-morbid conditions, pacemaker/ICD etc) and variables potentially associated with processes of care (primary care vs. cardiologist involvement in care). Confounding is assessed by a change in the effect size of treatment caused by inclusion of the potential confounder as a covariate. (HCFA Heart Failure Demonstration Project, 2000)

#### Generalizability of Study Sample

To facilitate interpretation of study findings and estimate the generalizability of findings, potentially eligible Medicare beneficiaries were screened for the study. Out of

those screened and found eligible were included in the study. Eligible enrollees were described in detail and compared to eligible non-enrollees according to all baseline data provided by the refusers. However, the measures of statistical significance for these comparisons were not reliable owing to the limited power.

### Preliminary Analyses

#### Univariate

Because of random assignment, the two groups were not expected to differ on any variable. If they differed on any variable related to an outcome or correlated highly with an outcome, its confounding effect was adjusted using the variable as a covariate in subsequent analyses. Univariate statistics (mean, median, SD, and SE) were computed for all quantitative variables; for all dichotomous variables, counts and percentages, along with SE, were computed.

#### Bivariate

Associations among all variables were examined for potential redundancy or colinearity for subsequent regression analysis. The method of assessing the association varied depending on the nature of the variables and the shape of their frequency distributions. If there was an association, then the variables were adjusted in subsequent analyses by stratification as covariates or by forced entry into regression equations.

The effects of the intervention were statistically tested using independent sample *t*-tests, paired sample *t*-tests, Mann–Whitney test, chi-square statistic, and Fisher’s exact tests, which, in turn, depended on the nature of the outcome variable and the shape of its frequency distribution.

## Analysis of Endpoints

Integrated methods of distribution- and anchor-based methods were used. RTM is a statistical phenomenon that occurs when repeated measurements are made on the same subject or unit of observation (Barnett et al., 2005). Values are observed with random error; it is rare to observe data without random error, which makes RTM a common phenomenon (Barnett et al., 2005). RTM is a statistical phenomenon that may make natural variation in repeated data appear to be real change (Barnett et al., 2005). The effect of RTM in a sample becomes more noticeable when measurement error increases due to repeated measurements in the sample (Barnett et al., 2005).

Change over 6 months from the baseline was adjusted for RTM using the EN method (Edwards et al., 1978; Speer et al., 1992). Regression to the normative group mean was assessed by finding Spearman's or Pearson's correlation coefficients (depending on the distribution) between the baseline and 6-month measurements, between the baseline and change from the baseline to 6-month measure, and between the absolute difference between the baseline and normative mean and the baseline to 6-month change in measurement. If the RTM was significant, then the change from baseline to 6 months was adjusted for RTM using the EN method. If the RTM was not significant, then the observed change from baseline to 6 months was taken into account without any adjustment.

### Adjustment for Regression to the Mean

The EN method classifies pre-post intervention changes as improved or deteriorated based on 95% confidence intervals calculated using the EN index. If there

was RTM, an individual's true score or measurement would be closer to the mean of the group than to the actual pretest score or measurement (Barnett et al., 2005).

$$\text{EN index} = (t' - X_2)/\text{SEM} \quad (2)$$

where,  $t'$  is the true score or measurement

$$t' = r_t(X_1 - M_G) + M_G \quad (3)$$

where,  $r_t$  is the reliability coefficient of the measure or instrument, and  $M_G$  is the mean of the normative group toward which the scores or measurements are assumed to regress (Crosby et al., 2004).

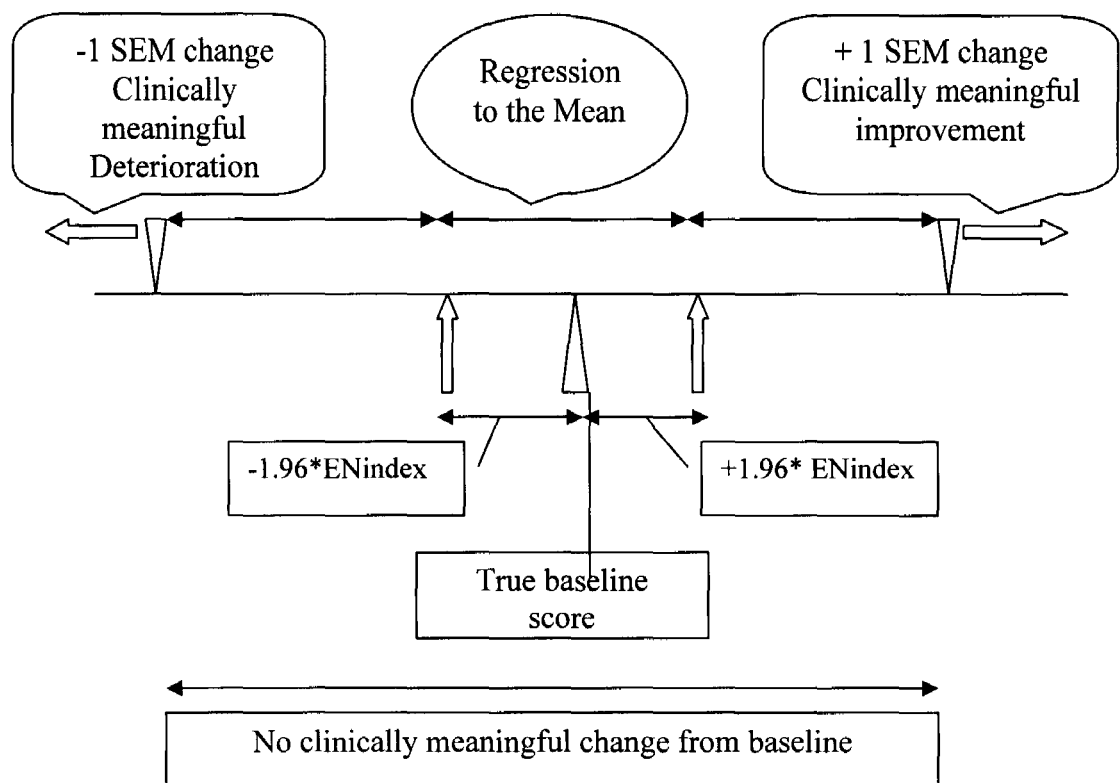
The EN CI is asymmetrical around the actual pretest score, whereas in the SEM method, the CI centers on the pretest score.

To calculate the mean of the normative group, the sample was anchored to NYHA Class II, III, and IV groups by stratification. Normative group means, SD, and SEM were calculated for each group. Using the above formulae, each individual's true score, EN index, and 95% confidence interval around the true score were calculated to establish threshold values of cutoffs (true score  $\pm 1.96 \times \text{EN index}$ ). If the post-intervention score or measurement after 6 months fell beyond the cutoffs, then the change from the baseline was considered true change toward improvement or deterioration after adjusting for RTM. To define whether this change was clinically meaningful or not, the change was measured in terms of SEM (i.e., true change after adjusting for RTM/SEM; see Figure 1).

#### Standard Error of Measurement

SEM was calculated using the formula  $\text{SD} \times \text{SQRT}(1-r_t)$ , in which SD is the standard deviation of the baseline measurement and SQRT is the square root (Anstassi & Urbina, 1997; Redelmier et al., 1997). If the magnitude of change from the baseline after





**Figure 1.** Adjustment of regression to the mean and calculation of clinically meaningful change.

correction to RTM was greater than or equal to 1 SEM, then that change was defined as clinically meaningful (Wyrwich, Tierney, et al., 1999). Based on the direction and magnitude of change in terms of  $\pm$  SEM units, individuals were grouped into clinically meaningful improvement, deterioration, or no change from baseline.

#### Assessment of the Effect of Technology-Driven Heart Failure Management System

Clinically meaningful change at the individual level was assessed by the SEM method after adjusting the change over time to RTM. The statistical significance of clinically meaningful change with respect to the standard care group and the standard

care plus telemonitoring group was tested through a chi-square statistic; odds ratios were calculated using a Mantle–Haenszel test. After defining groups with clinically meaningful improvement, deterioration, or no change in functional capacity or QOL, the group with no change was considered a reference group.

Multinomial regression analysis was performed to establish the best predictors of clinically meaningful improvement and deterioration in functional capacity and QOL for patients with HF. Both the control group and intervention group were controlled for all possible potential confounders. Sensitivity and specificity of clinically meaningful change in 6MWD and 6MW for predicting clinically meaningful change in MLHF score were calculated.

## CHAPTER IV

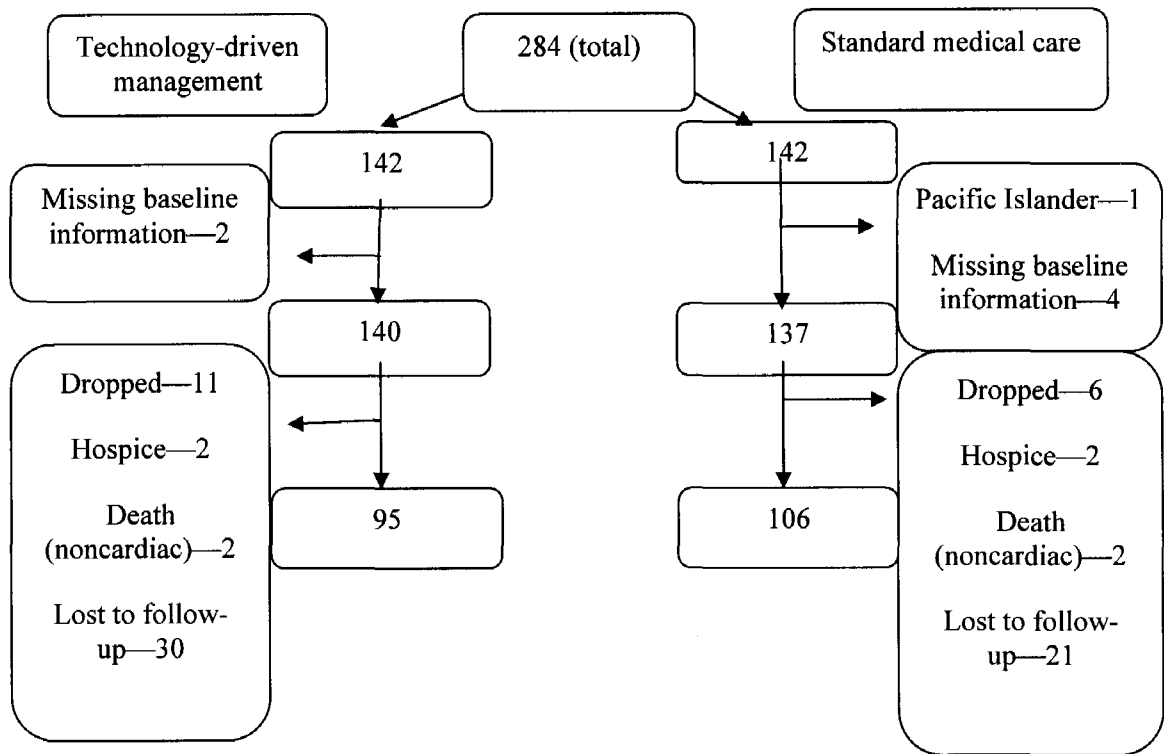
### RESULTS

#### Sample Description

A total of 284 participants who met eligibility criteria were randomized into either a standard care group ( $n = 142$ ) or a technology-driven HF management group ( $n = 142$ ). Out of 142 participants in the standard care group, 36 were excluded from analysis due to missing data (1 was a Pacific Islander, 4 lacked baseline characteristics documentation, 6 dropped out, 2 entered hospice, 2 suffered noncardiac deaths, and 21 were lost to follow-up). Out of 142 participants in the technology-driven HF management group, 47 were excluded from analysis due to missing data (2 lacked baseline characteristics documentation, 11 dropped out, 2 entered hospice, 2 suffered noncardiac deaths, and 30 were lost to follow-up).

Baseline characteristics, such as demographic variables, cardiovascular risk factors, past medical history of cardiovascular diseases, and HF disease characteristics (see Table 4), of the sample included in the study ( $n = 201$ ) were compared with those of the sample excluded from the analysis ( $n = 83$ ) due to loss to follow-up or missing information. The sample included in the analysis was not different from the group excluded from the analysis. Distribution of baseline characteristics in the standard medical care group and the technology-driven HF management group were tested. Tests of normality showed baseline target variables (6MWD, 6MW, MLHF total score,

emotional domain score, and physical domain score) were not normally distributed (see Figures 2–6 and Table 3).



**Figure 2.** Sample attrition.

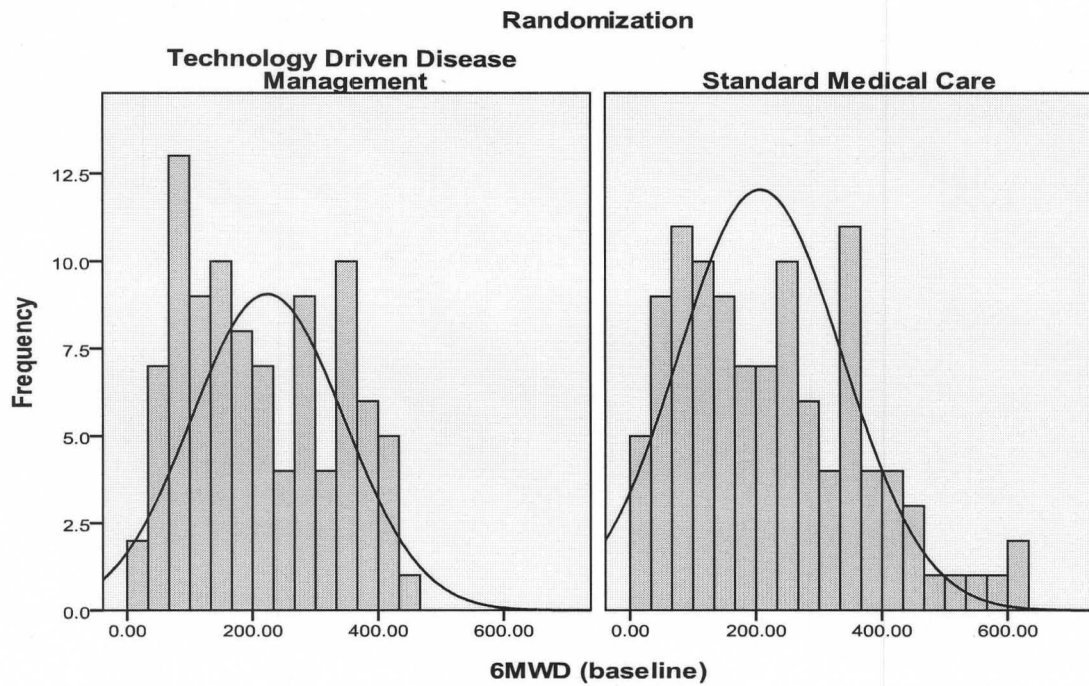


Figure 3. Frequency distribution of 6-minute walk distance at baseline.

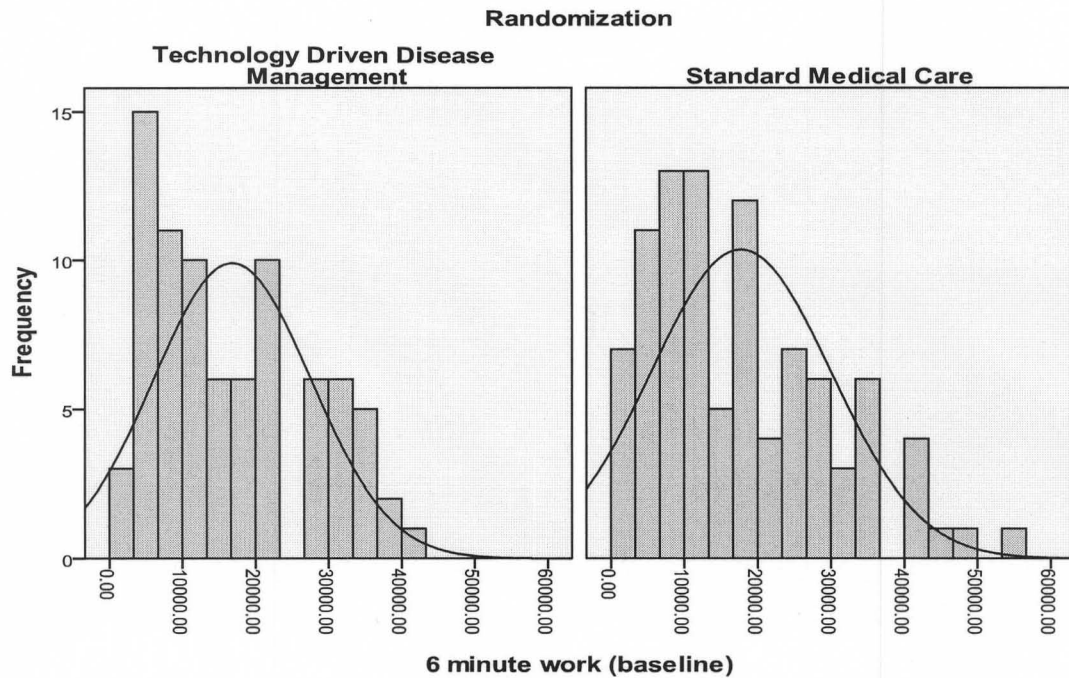


Figure 4. Frequency distribution of 6-minute work at baseline.

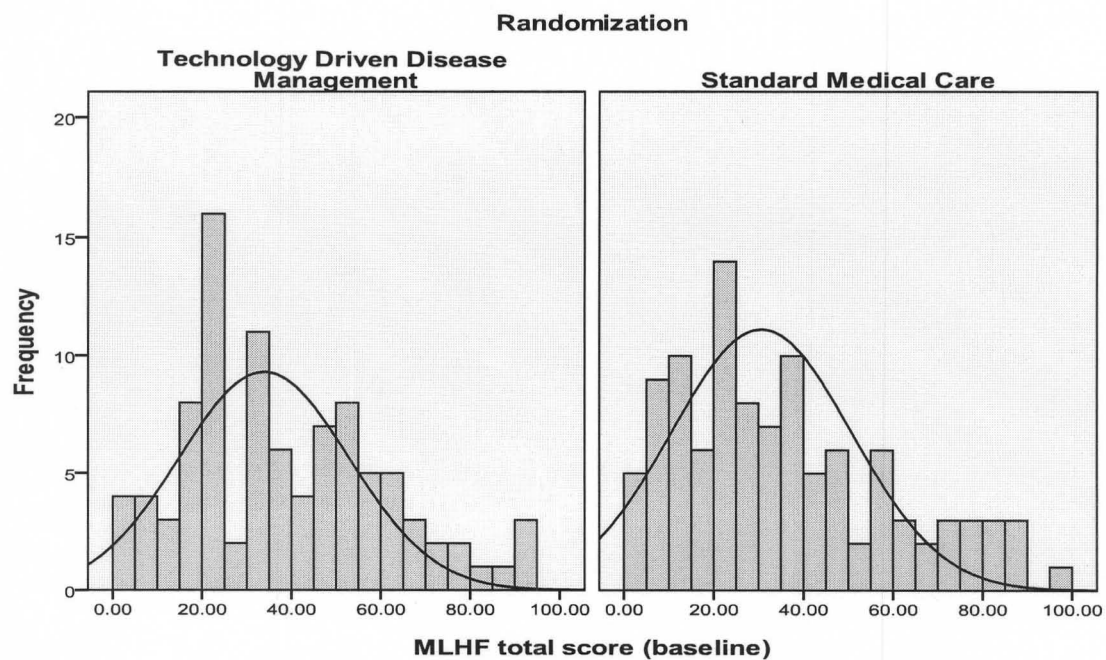


Figure 5. Frequency distribution of MLHF scores at baseline.

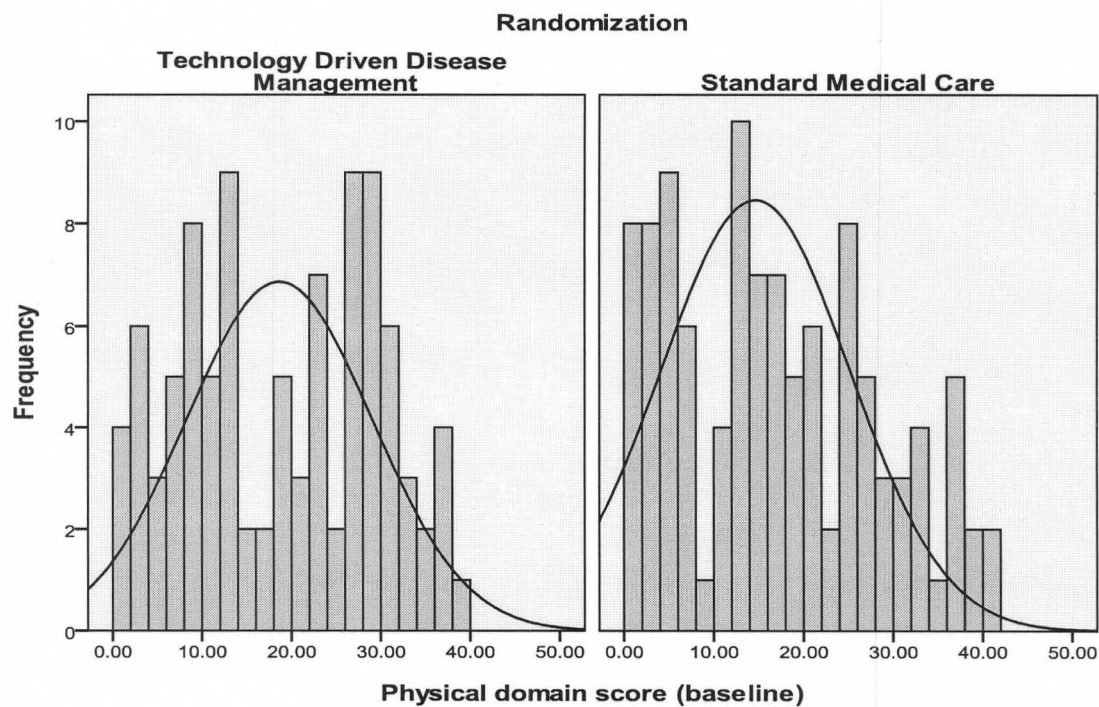


Figure 6. Frequency distribution of MLHF physical dimension scores at baseline.

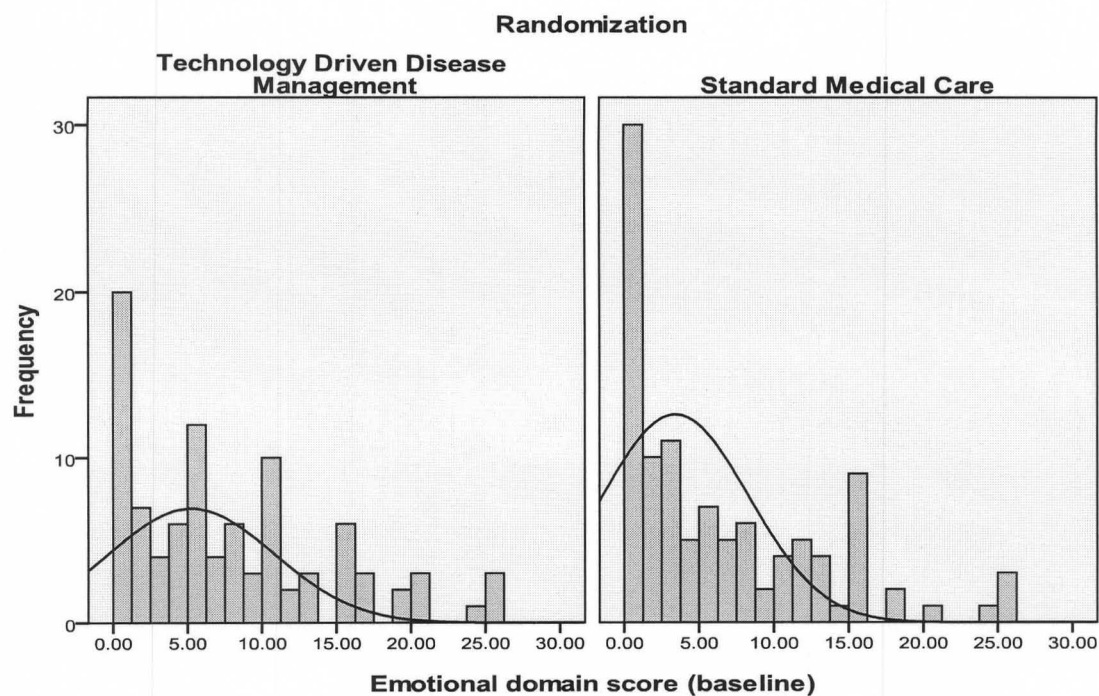


Figure 7. Frequency distribution of MLHF emotional dimension scores at baseline.

Table 3

Tests of Normality for Baseline Target Variables

		Shapiro–Wilk Test		
	<u>Randomization</u>	<u>Static</u>	<u>df</u>	<u>p (2-tailed)</u>
6MWD (baseline)	Technology-driven disease management	.941	95	.000
	Standard medical care	.957	106	.002
MLHF total score (baseline)	Technology-driven disease management	.965	95	.013
	Standard medical care	.943	106	.000
Physical domain score (baseline)	Technology-driven disease management	.950	95	.001
	Standard medical care	.952	106	.001
Emotional domain score (baseline)	Technology-driven disease management	.909	95	.000
	Standard medical care	.862	106	.000



### Estimation of Skewness and Kurtosis for Baseline Target Variable Distribution

The skewness and kurtosis of frequency distributions in the baseline 6MWD, 6MW, and QOL (MLHF) scores of the intervention vs. control groups were estimated (see Table 4). Kurtosis is a measure of the peakedness or flatness of a distribution. A kurtosis value near zero indicates a shape close to normal. A negative value indicates a distribution that is more peaked than normal, and a positive kurtosis indicates a shape that is flatter than normal. An extreme positive kurtosis indicates a distribution in which more of the values are located in the tails rather than around the mean.

Skewness is the extent to which a distribution of values deviates from symmetry around the mean. A value of zero means the distribution is symmetric. A positive skewness indicates a greater number of smaller values, and a negative value indicates a greater number of larger values.

In the Classic Principles of Statistics (1965), M.G. Bulmer suggests rule of thumbs. If skewness is less than  $-1$  or greater than  $+1$ , the distribution is highly skewed. If skewness is between  $-1$  and  $-\frac{1}{2}$  or between  $+\frac{1}{2}$  and  $+1$ , the distribution is moderately skewed. If skewness is between  $-\frac{1}{2}$  and  $+\frac{1}{2}$ , the distribution is approximately symmetric. The reference standard is a normal distribution, which has a kurtosis of 3. In token of this, often the excess kurtosis is presented:  $\text{excess kurtosis} = \text{kurtosis} - 3$ . A normal distribution has kurtosis exactly 3 (excess kurtosis exactly 0). Any distribution with  $\text{kurtosis} \approx 3$  (excess  $\approx 0$ ) is called mesokurtic. A distribution with  $\text{kurtosis} < 3$  (excess kurtosis  $< 0$ ) is called platykurtic. Compared to a normal distribution, its central peak is lower and broader, and its tails are shorter and thinner. A distribution with  $\text{kurtosis} > 3$  (excess kurtosis  $> 0$ ) is called leptokurtic. Compared to a normal distribution,

its central peak is higher and sharper, and its tails are longer and fatter (Bulmer, 1965; see Table 4).

Table 4

Estimated Skewness and Kurtosis

Target variable	Skewness (SE)	Kurtosis (SE)
Technology-driven care		
6-MWD	0.268 (0.250)	-1.239 (0.495)
6-MW	0.504 (0.250)	-0.897 (0.495)
MLHF total score	0.498 (0.250)	-0.382 (0.495)
Physical dimension score	-0.029 (0.250)	-1.217 (0.495)
Emotional dimension score	0.847 (0.250)	-0.048 (0.495)
Standard care		
6-MWD	0.609 (0.236)	-0.233 (0.467)
6-MW	0.690 (0.236)	-0.152 (0.467)
MLHF total score	0.675 (0.236)	-0.346 (0.467)
Physical dimension score	0.321 (0.236)	-0.890 (0.467)
Emotional dimension score	1.075 (0.236)	0.432 (0.467)

Note. If the skewness or kurtosis statistic falls within the range of  $\pm 2$  times the SE, then it falls within the expected range of chance fluctuations (Tabachnick & Fidell, 1996).

The 6MW, MLHF total score, and emotional dimension score were moderately skewed in both the standard care and technology-driven care groups, whereas the 6MWD values were skewed only in standard care group.

#### Comparison of Baseline Characteristics

The baseline characteristics of the intervention group receiving technology-driven HF management and the control group receiving standard medical care were analyzed using a chi-square test for equality of proportions or independent sample *t*-tests for equality of means or a Mann–Whitney (U) test, when appropriate. Both groups were well balanced with respect to their baseline characteristics, except for revascularization procedures (see Table 5). There was statistically significant difference between the standard care group and the technology-driven HF management group with respect to history of revascularization procedures. About 50% of the participants in the technology-driven HF management group had undergone revascularization procedures, whereas in the standard medical care group, only 36% had undergone revascularization procedures ( $X^2 = 4.409$ ,  $df = 1$ ,  $p = 0.036$ ; see Table 5).

#### Description of Target Variables After 6 Months

Functional capacity (6MWD and 6MW) and QOL scores (MLHF scores) after 6 months for the group receiving technology-driven HF management were compared with those of standard care group. The Mann–Whitney test statistic (U) and Z scores were calculated. There was no significant difference between the two groups after 6 months (see Table 6).

Table 5

Baseline Characteristics

Characteristic	Standard care (N = 106)	Technology- driven HF management (N = 95)	p value (2-tailed)
Demographic variables			
Age in years (Mean)	75 $\pm$ 8	74 $\pm$ 7	0.179
BMI (Mean)	27 $\pm$ 6	28 $\pm$ 6	0.183
Gender			0.825
Male (%)	59	58	
Female (%)	41	42	
Race			0.170
Caucasians (%)	80	87	
Others (%)	20	13	
Marital status			0.155
Married (%)	50	60	
Unmarried (%)	50	40	
Cardiovascular risk factors			
Smoking (%)	38	31	0.294
Hyperlipidemia (%)	60	63	0.686
Hypertension (%)	76	80	0.539
Diabetes (%)	44	40	0.534
Peripheral vascular diseases (%)	22	27	0.350
History of			
Cerebrovascular accident/transient ischemic attack (%)	19	22	0.690
Myocardial Infarction (%)	46	48	0.756
Revascularizations (%)	36	50	0.036 <sup>a</sup>
Valvular Abnormalities (%)	81	81	0.989
Disease characteristics			
Duration of HF in months (mean)	59 $\pm$ 74	59 $\pm$ 78	0.990
Left Ventricular Ejection Fraction (Mean)	32 $\pm$ 14	32 $\pm$ 14	0.817
Type of HF			
Systolic (%)	58	63	0.564
Diastolic (%)	13	13	1.000
Systolic and diastolic (%)	28	24	0.526
Type of Cardiomyopathy			0.458
Ischemic (%)	49	54	
Non-ischemic (%)	51	46	

Table 5 (continued)

Characteristic	Standard care (N = 106)	Technology- driven HF management (N = 95)	p value (2-tailed)
NYHA Class			0.509
II (%)	35	31	
III & IV (%)	65	69	
Disease management			
Cardiologist on board	75	69	0.380
Pacemaker (%)	16	15	0.710
ICD (%)	25	28	0.638
Beta-blockers (%)	81	85	0.608
ACE-I/ARBs (%)	93	88	0.217
Target variables			
6MWD in meters (mean)	226 ±144	213 ±118	0.784 <sup>b</sup>
6MW in kg/m (mean)	17931 ±12187	17225 ±10768	0.877 <sup>b</sup>
MLHF total score (mean)	35 ±23	38 ±22	0.219 <sup>b</sup>
Emotional dimension score (mean)	7 ±7	8 ±7	0.127 <sup>b</sup>
Physical dimension score (mean)	17 ±11	18 ±11	0.216 <sup>b</sup>

<sup>a</sup>Statistically significant difference observed. <sup>b</sup>Mann–Whitney (U) test.

Table 6

Target Variables After 6 Months for Standard Care vs. Technology-Driven Management

Target variable	Standard group (mean)	Technology- driven management group (mean)	Mann–Whitney test		
			<u>U</u>	<u>Z</u>	<u>p (2- tailed)</u>
6 MWD (m)	189.14	189.16	5022	-0.320	0.975
6 MW (kg/m of work)	15184.88	14342.54	3673	-0.538	0.590
MLHF total score	39.00	38.00	4930	-0.255	0.799
Physical domain score	18.50	17.90	5006	-0.069	0.945
Emotional domain score	7.50	7.60	4649	-0.945	0.345

The mean of the changes in functional capacity and QOL scores after 6 months for the standard management group vs. the technology-driven HF management group were not statistically significant (see Table 7).

Table 7

Mean Change in Functional Capacity and Quality of Life Over 6 Months

Target variable mean change	Standard group (mean)	Technology driven management group (mean)	Mann–Whitney test		
			<u>U</u>	<u>Z</u>	<u>p (2- tailed)</u>
6 MWD (m)	33.48	20.44	4731.5	-0.737	0.461
6 MW (kg/m of work)	2574.21	2482.28	3807.0	0.000	1.000
MLHF total score	-3.40	0.74	4814.0	-0.537	0.591
Physical domain score	-1.50	0.93	4577.5	-1.112	0.266
Emotional domain score	-0.84	0.27	4981.0	-0.132	0.895

Correlations Among Changes in Target Variables After 6 Months

After controlling for the treatment group assignment, the change in 6MWD after 6 months was correlated with changes in 6MW ( $r = 0.892$ ,  $p < 0.01$ ), MLHF total score ( $r = -0.537$ ,  $p < 0.01$ ), physical domain score ( $r = -0.492$ ,  $p < 0.01$ ), and emotional domain score ( $r = -0.468$ ,  $p < 0.01$ ) after 6 months. The change in 6MW after 6 months was also correlated with changes in MLHF total score ( $r = -0.418$ ,  $p < 0.01$ ), physical domain score ( $r = -0.364$ ,  $p < 0.01$ ), and emotional domain score ( $r = -0.350$ ,  $p < 0.01$ ) after 6 months. The change in physical domain score after 6 months was correlated with the change in emotional domain score after 6 months ( $r = 0.762$ ,  $p < 0.01$ ).

### Regression to the Mean

RTM is a statistical phenomenon that occurs when repeated measurements are made on the same subject or unit of observation. RTM may make natural variation in repeated data appear to be real change. It is rare to observe data without random error, which makes RTM a common phenomenon (Barnett et al., 2005). The effect of RTM in a sample becomes more noticeable when measurement error increases due to repeated measurements in the sample.

The scatter plot of change in target variables over 6 months against baseline measurements helped to identify RTM. The effect of RTM should be adjusted with appropriate statistical techniques. The scatter plots of change in target variable from baseline to 6 months vs. the baseline target variable (e.g., 6MWD, 6MW, and QOL [MLHF total score]) are depicted in Figures 8–10. These scatter plots provide visual input on the effect of RTM on baseline target variables.

### Assessment of Regression to Mean

Regression toward the normative mean was assessed by drawing partial correlations between baseline and 6-month measurements, between baseline measurements and change over 6 months, and between the absolute difference between baseline and normative group mean and the absolute difference between baseline and 6-month measurements of target variables (6MWD, 6MW, MLHF total score, MLHF physical domain score, and MLHF emotional domain score).

#### Regression to Mean for Change in 6-Minute Walk Distance Over 6 Months

There were partial correlations between 6MWD at baseline and after 6 months ( $r = 0.73$ ,  $p < 0.01$ ) and between baseline 6MWD and change in 6MWD from baseline to 6

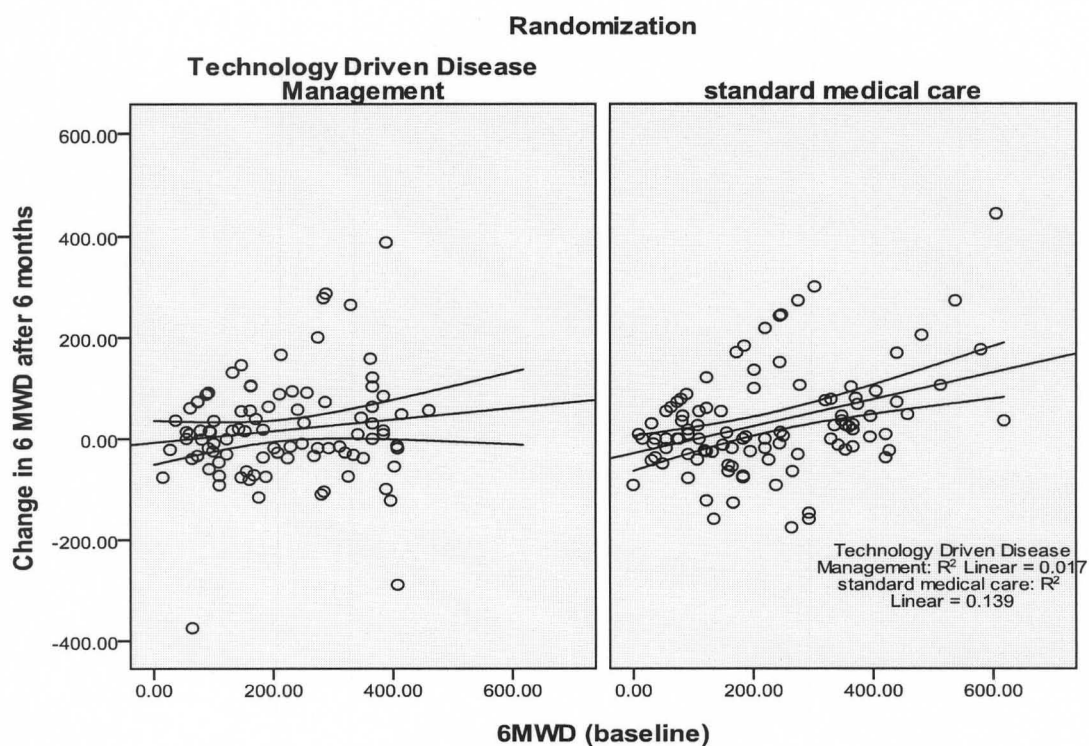


Figure 8. Scatter plot of change in 6-minute walk distance against baseline.

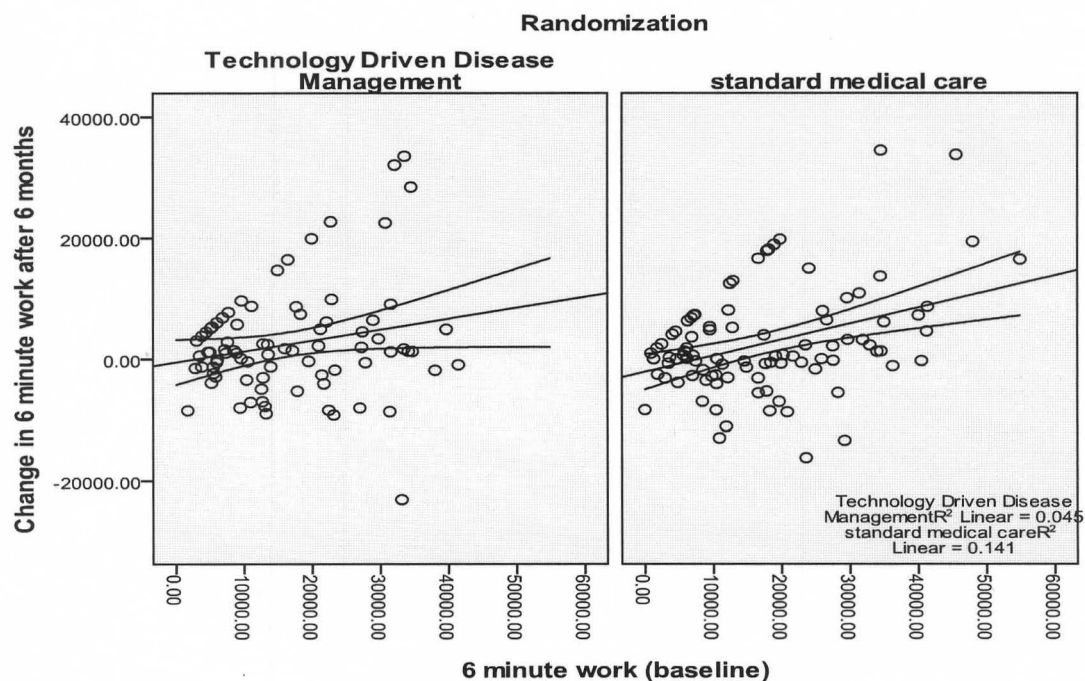


Figure 9. Scatter plot of change in 6-minute work against baseline.



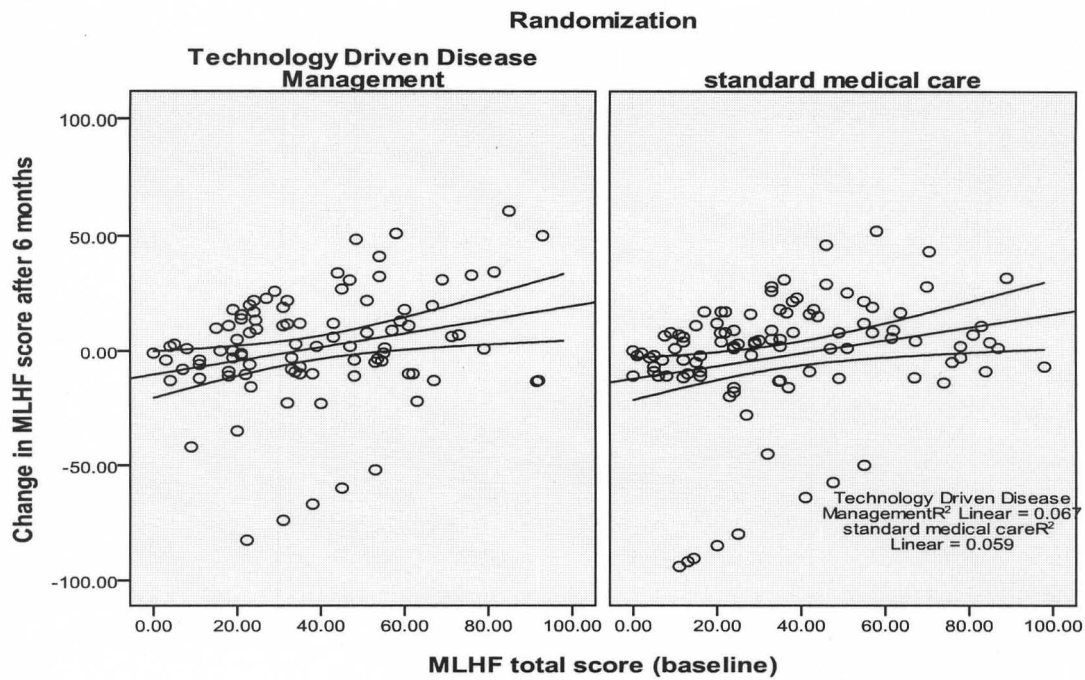


Figure 10. Scatter plot of change in MLHF total score (QOL) against baseline.

months ( $r = 0.27$ ,  $p < 0.01$ ). The absolute difference between the baseline 6MWD and normative mean and the absolute difference between the baseline 6MWD and 6MWD after 6 months ( $r = 0.15$ ,  $p < 0.01$ ) were statistically significant, indicating that RTM for change in 6MWD after 6 months was statistically significant.

#### Regression to Mean for Change in 6-Minute Work after 6 Months

Pearson's correlations were made between 6MW at baseline and after 6 months ( $r = 0.73$ ,  $p < 0.01$ ) and between baseline 6MW and change in 6MW after 6 months ( $r = 0.30$ ,  $p < 0.01$ ). The absolute difference between the baseline 6MW and normative mean, and the absolute difference between the baseline 6 MW and 6 MW after 6 months ( $r = 0.28$ ,  $p < 0.01$ ) were statistically significant, indicating that RTM for change in 6MW after 6 months was statistically significant.

### Regression to Mean for Change in MLHF Score after 6 Months

Partial correlations were made between the MLHF total score at baseline and after 6 months ( $r = 0.539$ ,  $p < 0.01$ ) and the baseline MLHF total score and change in MLHF total score after 6 months ( $r = 0.25$ ,  $p < 0.01$ ). The absolute difference between the baseline MLHF total score and the normative mean and the absolute difference between the baseline MLHF total score and the MLHF total score after 6 months ( $r = -0.07$ ,  $p = 0.30$ ) were statistically significant, indicating that RTM for change in MLHF total score from baseline to 6 months was statistically significant.

These findings suggested that the effect of RTM on target variable measurements at baseline and 6 months was statistically significant. Individuals with more severe impairment in functional capacity or QOL at baseline exhibited greater change from baseline to 6 months, and individuals with less impairment at baseline exhibited less change from baseline to 6 months. Therefore, the change in 6MWD, 6MW, MLHF total score, physical domain score, and emotional domain score from baseline to 6 months should be adjusted to RTM.

### Adjustment of RTM for Change in Target Variables from Baseline to 6 Months

If the regression toward the group's normative was significant, then the differences between baseline and after 6 months values of target variables might have understated or overstated the true change. Therefore, target variable change from baseline to 6 months was adjusted to RTM using the EN method. According to the EN method, the change in target variables from baseline to 6 months was classified as improved, no change, or deteriorated based on confidence intervals calculated using SEM. The change in measurements over time was adjusted for regression to the mean by

establishing threshold values of  $\pm 1.96 \times$  EN index (95% confidence interval) as cutoff values to determine real or true change from baseline. Then clinically meaningful change from baseline was calculated using SEM method (described later in this chapter; see Table 8).

Table 8

Partial Correlations Between Target Variables at Baseline and Change From Baseline to 6 Months

Variables	Partial correlations	p (2-tailed)
6MWD		
6MWD at baseline and after 6 months	0.725	0.00
Baseline 6MWD and change in 6MWD after 6 months	0.270	0.00
Absolute difference between baseline 6MWD and normative mean and absolute difference between baseline 6MWD and 6MWD after 6 months	0.153	0.03
6MW		
6MW at baseline and after 6 months	0.725	0.00
Baseline 6MW and change in 6MW after 6 months	0.302	0.00
Absolute difference between baseline 6MW and normative mean and absolute difference between baseline 6MW and 6MW after 6 months	0.277	0.00
QOL (MLHF) total score		
MLHF score at baseline and after 6 months	0.539	0.00
Baseline MLHF score and change in MLHF score after 6 months	0.250	0.00
Absolute difference between baseline MLHF score and normative mean and absolute difference between baseline MLHF score and MLHF score after 6 months	-0.072	0.30

For each participant, the EN index was calculated for each baseline measurement by using Equation 2. The true score or measurement at baseline was calculated using Equation 3. The normative group mean, SD, and SEM were calculated by anchoring the study group to NYHA Class II, III, or IV. To determine the magnitude of true change from the baseline, the cutoff values for change in target variables from baseline to 6 months (95% confidence interval) were determined by calculating  $\pm 1.96 * (\text{EN index})$ . The cutoff values (95% confidence interval) to determine true change from the baseline values for 6MWD, MLHF total score, physical domain score, and emotional domain score were determined by calculating  $\pm 1.96 * (\text{EN index})$  for each patient.

Therefore, if the change in 6MWD, MLHF total score, physical domain score, and emotional domain score fell within the range of cutoff values  $\pm 1.96 * (\text{EN index})$ , then the change was not real and was due to the effect of RTM. If the magnitude of change from the baseline to 6 months fell beyond the cutoff values, then the change was real after adjusting to RTM.

#### Determination of Clinically Meaningful Change in Target Variables From Baseline to 6 Months

If the magnitude of change from the baseline cutoff values was greater than or equal to 1 SEM, then it was defined as clinically meaningful (Wyrwich et al., 1996). Based on the direction and magnitude of change in terms of  $\pm$  SEM units, individuals were grouped according to clinically meaningful improvement, deterioration, or no change from the baseline. SEM was calculated using the formula  $\text{SD} * \text{SQRT}(1 - r_t)$ , in which SD is standard deviation of the baseline measurement and SQRT is the square root (Anstasi et al., 1997; Redelmier et al., 1997). After adjusting the change from baseline to

RTM, the cutoff values for true change in the target variables from baseline to 6 months (6MWD, 6MW, MLHF total score, MLHF physical domain score, and MLHF emotional domain score) were determined in terms of  $\pm 1$  unit of SEM.

If the magnitude of true change in the target variables fell beyond  $\pm 1$  unit of SEM, then it was clinically meaningful. If the magnitude of true change in target variables fell within the range of cutoff values ( $\pm 1$  SEM unit), then it was true change but not clinically meaningful. Change in 6MWD by 42 m (or 39 m–42 m), in 6MW by 3,668 kg/m of work (or 3,063–3,820), in MLHF total score by 6.0 points (or 5.0–7.0), in physical domain by 4.0 points (or 3.0–4.0), and emotional domain by 2.0 points (or 2.0–3.0) were considered clinically meaningful change in patients with chronic CHF (see Tables 9 and 10).

Table 9

Standard Error of Measurement for Target Variables

Target variable	Mean	SD	Reliability coefficient of measurement (rt)	SEM
6MWD (m)	220.43	133.00	0.90	42.06
6MW (kg/m of work)	17326.28	14857.16	0.90	3667.61
MLHF total score	36.70	23.22	0.93	6.14
Physical domain score	17.55	11.12	0.89	3.69
Emotional domain score	7.09	6.64	0.88	2.30

Table 10

Standard Error of Measurement for Target Variable After Anchoring by NYHA Class

Target variable	Mean	SD	Reliability coefficient of measurement (rt)	SEM
NYHA Class II				
6MWD (m)	307.74	125.21	0.90	39.59
6MW (kg/m of work)	24968.17	11236.05	0.90	3553.15
MLHF total score	26.74	19.91	0.93	5.27
Physical domain score	12.00	9.54	0.89	3.16
Emotional domain score	5.24	5.59	0.88	1.94
NYHA Class III				
6MWD (m)	183.23	111.41	0.90	35.23
6MW (kg/m of work)	14363.49	9688.37	0.90	3063.73
MLHF total score	41.34	22.79	0.93	6.03
Physical domain score	20.18	10.89	0.89	3.61
Emotional domain score	7.8	6.76	0.88	2.34
NYHA Class IV				
6MWD (m)	144.25	132.90	0.90	42.03
6MW (kg/m of work)	11177.32	12090.92	0.90	3823.48
MLHF total score	42.95	26.38	0.93	6.98
Physical domain score	20.79	10.97	0.89	3.64
Emotional domain score	9.16	8.09	0.88	2.80

Clinically Meaningful Change in 6MWD for Standard Care Vs. Technology-DrivenDisease Management

Based on the direction and magnitude of change in terms of  $\pm 1$  SEM units for target variables, individuals were divided into three groups: (a) clinically meaningful improvement, (b) no change, and (c) clinically meaningful deterioration from the baseline. Participants who died from cardiovascular disease were included in the

deterioration group. In comparing the standard care vs. technology-driven HF management groups, there was no statistically significant difference with respect to clinically meaningful change in 6MWD after 6 months (see Table 11).

Table 11

Clinically Meaningful Change in 6MWD in Standard Care vs. Technology-Driven Disease Management

			Clinically meaningful change in 6MWD			Total
			Deteriorated	No change	Improved	
Randomization	Technology driven-disease management	Count	21	41	33	95
		% within randomization	22.1%	43.2%	34.7%	100.0%
	Standard medical care	Count	21	45	40	106
		% within randomization	19.8%	42.5%	37.7%	100.0%
Total		Count	42	86	73	201
		% within randomization	20.9%	42.8%	36.3%	100.0%

Note.  $X^2 = 0.256$ ,  $df = 2$ ,  $p$  (2-tailed) = 0.880.

Addition of technology-driven telemonitoring to standard therapy didn't provide any additional clinically meaningful impact on functional capacity measured in terms of 6-MWD. After stratification of both groups into NYHA Class II and Class III/IV, the effect of telemonitoring was still not significant (see Table 12).

Table 12

Clinically Meaningful Change in 6MWD for Standard Care vs. Technology-DrivenDisease Management After Stratification

NY HA clas s				Clinically meaningful change in 6MWD			Total
				Deteri- orated	No change	Im- proved	
III & IV	Randomi- zation	Technology	Count	15	28	23	66
		driven- disease	% within randomi- zation	22.7%	42.4%	34.8%	100.0%
		management					
	Standard medical care	Technology	Count	14	32	23	69
		driven- disease	% within randomi- zation	20.3%	46.4%	33.3%	100.0%
		management					
II	Randomi- zation	Technology	Count	6	13	10	29
		driven- disease	% within randomi- zation	20.7%	44.8%	34.5%	100.0%
		management					
	Standard medical care	Technology	Count	7	13	17	37
		driven- disease	% within randomi- zation	18.9%	35.1%	45.9%	100.0%
		management					
	Total	Technology	Count	29	60	46	135
		driven- disease	% within randomi- zation	21.5%	44.4%	34.1%	100.0%
		management					

Note. For NYHA Class II,  $\chi^2 = 0.235$ ,  $df = 2$ , and  $p$  (2-tailed) = 0.889. For NYHA Classes III and IV,  $\chi^2 = 0.936$ ,  $df = 2$ , and  $p$  (2-tailed) = 0.626.



Clinically Meaningful Change in 6MW from Baseline to 6 Months for Standard Care Vs.

Technology-Driven Disease Management

In comparing standard care and technology-driven disease management groups, there was no statistically significant difference with respect to clinically meaningful change in 6MW after 6 months (see Table 13). Addition of technology-driven telemonitoring to standard medical therapy didn't provide any additional clinically meaningful benefits in terms of functional capacity as measured by 6MW. After stratification of both groups into NYHA Class II and Class III/IV, the effect of telemonitoring was still not significant (see Table 14).

Table 13

Clinically Meaningful Change in 6MW for Standard Care vs. Technology-Driven Disease Management

			Clinically meaningful change in 6MW			Total
			Deteri- orated	No change	Im- proved	
Randomi- zation	Technology driven- disease management	Count	16	38	26	80
		% within randomization	20.0%	47.5%	32.5%	100.0%
	Standard medical care	Count	18	43	33	94
		% within randomization	19.1%	45.7%	35.1%	100.0%
Total		Count	34	81	59	174
		% within randomization	19.5%	46.6%	33.9%	100.0%

Note.  $\chi^2 = 0.131$ ,  $df = 2$ ,  $p$  (2-tailed) = 0.937.

Table 14

Clinically Meaningful Change in 6MW for Standard Care vs. Technology-Driven DiseaseManagement After Stratification

NYHA class				Clinically meaningful change in 6 MW			Total
				Deterio- -rated	No change	Im- proved	
III & IV	Randomization	Technology driven-disease management	Count	11	25	20	56
			% within randomization	19.6%	44.6%	35.7%	100.0%
		Standard medical care	Count	13	31	20	64
			% within randomization	20.3%	48.4%	31.3%	100.0%
	Total		Count	24	56	40	120
			% within randomization	20.0%	46.7%	33.3%	100.0%
II	Randomization	Technology driven-disease management	Count	5	13	6	24
			% within randomization	20.8%	54.2%	25.0%	100.0%
		Standard medical care	Count	5	12	13	30
			% within randomization	16.7%	40.0%	43.3%	100.0%
	Total		Count	10	25	19	54
			% within randomization	18.5%	46.3%	35.2%	100.0%

Note. For NYHA Class II,  $\underline{X}^2 = 1.463$ ,  $\underline{df} = 2$ , and  $\underline{p}(2\text{-tailed}) = 0.481$ . For NYHA Class III and IV,  $\underline{X}^2 = 1.186$ ,  $\underline{df} = 2$ , and  $\underline{p}(2\text{-tailed}) = 0.359$ .

Clinically Meaningful Change in QOL From Baseline to 6 Months for Standard Care Vs.  
Technology-Driven Disease Management

In comparing standard care and technology-driven disease management, there was no statistically significant difference with respect to clinically meaningful change in MLHF total score after 6 months (see Table 15). Addition of technology-driven telemonitoring to standard medical therapy didn't provide any additional clinically meaningful benefit for QOL measured in terms of MLHF score. After stratification of both groups into NYHA Class II and Class III/IV, the effect of telemonitoring was still not significant (see Table 16).

Table 15

Clinically Meaningful Change in MLHF Total Score for Standard Care vs. Technology-Driven Disease Management

			Clinically meaningful change in MLHF total score			Total
			Deteri- orated	No change	Improved	
Randomization	Technology driven- disease management	Count	38	30	27	95
		% within randomization	40.0%	31.6%	28.4%	100.0 %
	Standard medical care	Count	30	46	30	106
		% within randomization	28.3%	43.4%	28.3%	100.0 %
Total		Count	68	76	57	201
		% within randomization	33.8%	37.8%	28.4%	100.0 %

Note.  $\chi^2 = 3.877$ ,  $df = 2$ ,  $p$  (2-tailed) = 0.144.

Table 16

Clinically Meaningful Change in MLHF Total Score for Standard Care vs. Technology-Driven Disease Management After Stratification

NYHA Class				Clinically meaningful change in MLHF total score			Total
				Deteriorated	No change	Improved	
III & IV	Randomization	Technology driven-disease management	Count	24	21	21	66
			% within randomization	36.4%	31.8%	31.8%	100.0%
	Standard medical care	Count	20	30	19	69	
		% within randomization	29.0%	43.5%	27.5%	100.0%	
	Total		Count	44	51	40	135
			% within randomization	32.6%	37.8%	29.6%	100.0%
II	Randomization	Technology driven-disease management	Count	14	9	6	29
			% within randomization	48.3%	31.0%	20.7%	100.0%
	Standard medical care	Count	10	16	11	37	
		% within randomization	27.0%	43.2%	29.7%	100.0%	
	Total		Count	24	25	17	66
			% within randomization	36.4%	37.9%	25.8%	100.0%

Note. For NYHA Class II,  $X^2 = 3.174$ ,  $df = 2$ , and  $p(2\text{-tailed}) = 0.205$ . For NYHA Class III and IV,  $X^2 = 1.986$ ,  $df = 2$ , and  $p(2\text{-tailed}) = 0.370$ .

Clinically Meaningful Change in Physical Domain Score from Baseline to 6 Months for  
Standard Care Vs. Technology-Driven Disease Management

In comparing standard care vs. technology-driven disease management, there was no statistically significant difference with respect to clinically meaningful change in MLHF physical dimension score after 6 months (see Table 17). Addition of technology-driven telemonitoring to standard medical therapy didn't provide any additional clinically meaningful benefit for the physical domain of MLHF score. After stratification of both groups into NYHA Class II and Class III/IV, the effect of telemonitoring was still not significant (see Table 18).

Table 17

Clinically Meaningful Change in MLHF Physical Domain Score for Standard Care vs.  
Technology-Driven Disease Management

			<u>Clinically meaningful change in physical domain score</u>			Total
			Deteri- orated	No change	Improved	
Randomi- zation	Technology driven- disease management	Count	19	61	15	95
		% within randomization	20.0%	64.2%	15.8%	100.0%
	Standard medical care	Count	18	71	17	106
		% within randomization	17.0%	67.0%	16.0%	100.0%
Total		Count	37	132	32	201
		% within randomization	18.4%	65.7%	15.9%	100.0%

Note.  $\chi^2 = 0.309$ ,  $df = 2$ ,  $p$  (2-tailed) = 0.857.

### Clinically Meaningful Change in MLHF Physical Domain Score for Standard Care vs. Technology-Driven Disease Management After Stratification

**Note.** For NYHA Class II,  $\bar{X}^2 = 0.839$ ,  $df = 2$ , and  $p$  (2-tailed) = 0.658. For NYHA Class III and IV,  $\bar{X}^2 = 0.127$ ,  $df = 2$ , and  $p$  (2-tailed) = 0.938.

Clinically Meaningful Change in MLHF Emotional Domain Score From Baseline to 6 Months for Standard Care Vs. Technology-Driven Disease Management

In comparing standard care and technology-driven disease management, there was no statistically significant difference with respect to clinically meaningful change in MLHF emotional dimension score after 6 months (see Table 19). Addition of technology driven telemonitoring to standard medical therapy didn't provide any additional clinically meaningful benefit for the emotional domain of MLHF score. After stratification of both groups into NYHA Class II and Class III/IV, the effect of telemonitoring was still not significant (see Table 20).

Table 19

Clinically Meaningful Change in MLHF Emotional Domain Score for Standard Care vs. Technology-Driven Disease Management

			Clinically meaningful change in emotional domain score			Total
			Deteri- orated	No change	Im- proved	
Randomization	Technology driven- disease management	Count	5	87	3	95
		% within randomization	5.3%	91.6%	3.2%	100.0%
	Standard medical care	Count	6	94	6	106
		% within randomization	5.7%	88.7%	5.7%	100.0%
Total		Count	11	181	9	201
		% within randomization	5.5%	90.0%	4.5%	100.0%

Note.  $\chi^2 = 0.762$ ,  $df = 2$ ,  $p(2\text{-tailed}) = 0.683$   
Table 20

# Clinically Meaningful Change in MLHF Emotional Domain Score for Standard Care

## Group vs. Technology-Driven Disease Management After Stratification

NYHA Class				Clinically meaningful change in emotional domain score			Total
				Deteriorated	No change	Improved	
III & IV	Randomization	Technology driven-disease management	Count	4	59	3	66
			% within randomization	6.1%	89.4%	4.5%	100.0%
		Standard medical care	Count	6	57	6	69
			% within randomization	8.7%	82.6%	8.7%	100.0%
	Total		Count	10	116	9	135
			% within randomization	7.4%	85.9%	6.7%	100.0%
II	Randomization	Technology driven-disease management	Count	1	28		29
			% within randomization	3.4%	96.6%		100.0%
		Standard medical care	Count	0	37		37
			% within randomization	.0%	100.0%		100.0%
	Total		Count	1	65		66
			% within randomization	1.5%	98.5%		100.0%

Note. For NYHA Class II,  $\underline{X^2} = 1.295$ ,  $\underline{df} = 2$ , and  $\underline{p}(2\text{-tailed}) = 0.439$ . For NYHA Class III and IV,  $\underline{X^2} = 1.368$ ,  $\underline{df} = 2$ , and  $\underline{p}(2\text{-tailed}) = 0.504$ .



Clinically Meaningful Change in 6MWD to Predict Clinically Meaningful Similar

Directional Change in MLHF Score

1. Clinically meaningful improvement in 6MWD to predict clinically meaningful improvement in MLHF score: sensitivity = 56.1%, specificity = 71.5%, positive predictive value = 43.8%, negative predictive value = 80.5%, agreement = 67%.
2. Clinically meaningful deterioration in 6MWD to predict clinically meaningful deterioration in MLHF score: sensitivity = 35.3 %, specificity = 86.5%, positive predictive value = 57.1%, negative predictive value = 72.3%, agreement = 69%.
3. No clinically meaningful change in 6MWD to predict no clinically meaningful change in MLHF score: sensitivity = 44.7 %, specificity = 58.4%, positive predictive value = 39.5%, negative predictive value = 63.4%, agreement = 53.2% (see Table 21).

Clinically Meaningful Change in 6MWD to Predict Clinically Meaningful Similar

Directional Change in MLHF Physical Domain Score

1. Clinically meaningful improvement in 6MWD to predict clinically meaningful improvement in MLHF physical domain score: sensitivity = 68.8%, specificity = 69.8%, positive predictive value = 30.1%, negative predictive value = 92.2%, agreement = 69.7%.
2. Clinically meaningful deterioration in 6MWD to predict clinically meaningful deterioration in MLHF physical domain score: sensitivity = 35.1 %, specificity = 82.3%, positive predictive value = 31.0%, negative predictive value = 85.0%, agreement = 73.6%.

Table 21

Clinical Meaningful Change in 6MWD vs. Clinically Meaningful Change in MLHF TotalScore

		Clinically meaningful change in MLHF total score			Total
		Deteriorated	No change	Improved	
Clinically meaningful change in 6MWD	Deteriorated	24	15	3	42
		57.1%	35.7%	7.1%	100.0%
	No change	30	34	22	86
		34.9%	39.5%	25.6%	100.0%
	Improved	14	27	32	73
		19.2%	37.0%	43.8%	100.0%
Total		68	76	57	201
		33.8%	37.8%	28.4%	100.0%

Note.  $\chi^2 = 24$ ,  $df = 4$ ,  $p < 0.001$ ; Spearman correlation = 0.345,  $p < 0.001$ ; Kappa = 0.169,  $p < 0.001$ .

- No clinically meaningful change in 6MWD to predict no clinically meaningful change in MLHF physical domain score: sensitivity = 46.2 %, specificity = 63.8%, positive predictive value = 71.0%, negative predictive value = 38.3%, agreement = 52.2% (see Table 22).

Clinically Meaningful Change in 6MWD to Predict Clinically Meaningful SimilarDirectional Change in MLHF Emotional Domain Score

- Clinically meaningful improvement in 6MWD to predict clinically meaningful improvement in MLHF emotional domain score: sensitivity = 100%, specificity = 66.7%, positive predictive value = 12.3%, negative predictive value = 100%, agreement = 68.1%.

Table 22

Clinically Meaningful Change in 6MWD vs. Clinically Meaningful Change in Physical Dimension Score

		Clinically meaningful change in physical domain score			Total
		Deteriorated	No change	Improved	
Clinically meaningful change in 6MWD	Deteriorated	13	27	2	42
		31.0%	64.3%	4.8%	100.0%
	No change	17	61	8	86
		19.8%	70.9%	9.3%	100.0%
	Improved	7	44	22	73
		9.6%	60.3%	30.1%	100.0%
Total		37	132	32	201
		18.4%	65.7%	15.9%	100.0%

Note.  $\chi^2 = 22.37$ ,  $df = 4$ ,  $p < 0.001$ ; Spearman correlation = 0.31,  $p < 0.001$ ; Kappa = 0.161,  $p < 0.001$ .

2. Clinically meaningful deterioration in 6MWD to predict clinically meaningful deterioration in MLHF emotional domain score: sensitivity = 63.6%, specificity = 81.6%, positive predictive value = 16.7%, negative predictive value = 97.5%, agreement = 80.6%.
3. No clinically meaningful change in 6MWD to predict no clinically meaningful change in MLHF emotional domain score: sensitivity = 45.3 %, specificity = 80%, positive predictive value = 95%, negative predictive value = 14%, agreement = 49% (see Table 23).

Table 23

Clinically Meaningful Change in 6MWD vs. Clinically Meaningful Change in Emotional Dimension Score

		Clinically meaningful change in emotional domain score			Total
		Deteriorated	No change	Improved	
Clinically meaningful change in 6MWD	Deteriorated	7	35	0	42
		16.7%	83.3%	.0%	100.0%
	No change	4	82	0	86
		4.7%	95.3%	.0%	100.0%
	Improved	0	64	9	73
		.0%	87.7%	12.3%	100.0%
Total		11	181	9	201
		5.5%	90.0%	4.5%	100.0%

Note.  $\chi^2 = 30.02$ ,  $df = 4$ ,  $p < 0.001$ ; Spearman correlation = 0.35,  $p < 0.001$ ; Kappa = 0.127,  $p < 0.001$ .

Clinically Meaningful Change in 6MW to Predict Clinically Meaningful Similar

Directional Change in MLHF Score

1. Clinically meaningful improvement in 6MW to predict clinically meaningful improvement in MLHF score: sensitivity = 56%, specificity = 75%, positive predictive value = 47.5%, negative predictive value = 81%, agreement = 69.5%.
2. Clinically meaningful deterioration in 6MW to predict clinically meaningful deterioration in MLHF score: sensitivity = 32.8 %, specificity = 87%, positive predictive value = 56%, negative predictive value = 72%, agreement = 69%.
3. No clinically meaningful change in 6MW to predict no clinically meaningful change in MLHF score: sensitivity = 53 %, specificity = 57.4%, positive

predictive value = 43.2%, negative predictive value = 66.7%, agreement = 55.7%

(see Table 24).

Table 24

Clinically Meaningful Change in 6MW vs. Clinically Meaningful Change in MLHF

Score

		Clinically meaningful change in MLHF total score			Total
		Deteriorated	No change	Improved	
Clinically meaningful change in 6MW	Deteriorated	19	13	2	34
		55.9%	38.2%	5.9%	100.0%
	No change	26	35	20	81
		32.1%	43.2%	24.7%	100.0%
	Improved	13	18	28	59
		22.0%	30.5%	47.5%	100.0%
Total		58	66	50	174
		33.3%	37.9%	28.7%	100.0%

Note.  $X^2 = 22.8$ ,  $df = 4$ ,  $p < 0.001$ ; Spearman correlation = 0.33,  $p < 0.001$ ; Kappa = 0.200,  $p < 0.001$ .

Clinically Meaningful Change in 6MW to Predict Clinically Meaningful Similar

Directional Change in MLHF Physical Domain Score

1. Clinically meaningful improvement in 6MW to predict clinically meaningful improvement in MLHF physical domain score: sensitivity = 69.3%, specificity = 72.3%, positive predictive value = 30.5%, negative predictive value = 93%, agreement = 72%.

2. Clinically meaningful deterioration in 6MW to predict clinically meaningful deterioration in MLHF physical domain score: sensitivity = 34.4 %, specificity = 83.4%, positive predictive value = 32.5%, negative predictive value = 85.0%, agreement = 74.7%.
3. No clinically meaningful change in 6MW to predict no clinically meaningful change in MLHF physical domain score: sensitivity = 51 %, specificity = 62%, positive predictive value = 73%, negative predictive value = 38.7%, agreement = 54.6% (see Table 25).

Table 25

Clinically Meaningful Change in 6MW vs. Clinically Meaningful Change in Physical Dimension Score

		Clinically meaningful change in physical domain score			Total
		Deteriorated	No change	Improved	
Clinically meaningful change in 6MW	Deteriorated	11	22	1	34
		32.4%	64.7%	2.9%	100.0%
	No change	15	59	7	81
		18.5%	72.8%	8.6%	100.0%
	Improved	6	35	18	59
		10.2%	59.3%	30.5%	100.0%
Total		32	116	26	174
		18.4%	66.7%	14.9%	100.0%

Note.  $\chi^2 = 21.7$ ,  $df = 4$ ,  $p < 0.001$ ; Spearman correlation = 0.32,  $p < 0.001$ ; Kappa = 0.18,  $p < 0.001$ .

### Clinically Meaningful Change in 6MW to Predict Clinically Meaningful Similar

#### Directional Change in MLHF Emotional Domain Score

1. Clinically meaningful improvement in 6MW to predict clinically meaningful improvement in MLHF emotional domain score: sensitivity = 89%, specificity = 69%, positive predictive value = 13.6%, negative predictive value = 99%, agreement = 70%.
2. Clinically meaningful deterioration in 6MW to predict clinically meaningful deterioration in MLHF emotional domain score: sensitivity = 45.4%, specificity = 82%, positive predictive value = 14.7%, negative predictive value = 95.7%, agreement = 80%.
3. No clinically meaningful change in 6MW to predict no clinically meaningful change in MLHF emotional domain score: sensitivity = 48.7 %, specificity = 70%, positive predictive value = 92.6%, negative predictive value = 15%, agreement = 51% (see Table 26).

#### Agreement Between NYHA Class and 6MWD in Assessing the Severity of Heart Failure

Of participants, 4% had mildly impaired functional capacity, but 33% were classified under NYHA Class II. Only 25% had moderately impaired functional capacity, but 58% were classified under NYHA Class III. For 71% of the participants functional capacity was severely impaired, but only 10% were classified under NYHA Class IV. Objectively measured functional capacity in terms of 6MWD was weakly correlated with physician-assessed severity (NYHA class) of HF condition (spearman correlation (S) = 0.365,  $p < 0.05$ ). The agreement between these two methods was very poor (18%) and not statistically significant (Kappa = -0.38,  $p = 0.173$ ; see Table 27).

Table 26

Clinically Meaningful Change in 6MW Vs. Clinically Meaningful Change in Emotional Dimension Score

		Clinically meaningful change in emotional domain score			Total
		Deteriorated	No change	Improved	
Clinically meaningful change in 6MW	Deteriorated	5	29	0	34
		14.7%	85.3%	.0%	100.0%
	No change	5	75	1	81
		6.2%	92.6%	1.2%	100.0%
	Improved	1	50	8	59
		1.7%	84.7%	13.6%	100.0%
Total		11	154	9	174
		6.3%	88.5%	5.2%	100.0%

Note.  $\chi^2 = 18.28$ ,  $df = 4$ ,  $p < 0.001$ ; Spearman correlation = 0.293,  $p < 0.001$ ; Kappa = 0.114,  $p < 0.001$ .

Table 27

Agreement Between NYHA Class and Functional Capacity

			6MWD at baseline			Total
			> 450 m	301–450 m	< 300 m	
NYHA class at baseline	II	Count	7	29	30	66
		% of total	3.5%	14.4%	14.9%	32.8%
	III	Count	1	17	98	116
		% of total	.5%	8.5%	48.8%	57.7%
	IV	Count	0	4	15	19
		% of total	.0%	2.0%	7.5%	9.5%
Total		Count	8	50	143	201
		% of total	4.0%	24.9%	71.1%	100.0
						%

Note.  $\chi^2 = 34.68$ ,  $df = 1$ ,  $p < 0.05$ ; Kappa = -0.038,  $p = 0.173$ ; agreement = 18%.



There is lack of agreement between objective assessment of NYHA functional class and objective measuring of functional capacity because either patients or physicians are under- or over-estimating functional capacity compared to objectively measured values. The method of assessing NYHA class is arbitrary, and intraoperator variability is 66% (Claire et al., 2007). When assessing NYHA class, many clinicians simply ask patients with HF exertion symptoms how far they can walk or exert themselves before symptom onset.

On the other hand, 6MWD (measured in meters) may not be a sole indicator of severity of CHF condition, especially when comparing individuals with different ages, genders, BMIs, and comorbid conditions. The NYHA system and 6MWD are regularly used as outcome measures in clinical trials and are even included in HF management guidelines. Thus, the lack of agreement and consistency between these two classification methods suggests the need for development of novel classification methods by designing norm-referenced equations comprising age, gender, BMI, comorbid conditions, symptoms, and signs of HF.

#### Predictors of Clinically Meaningful Change in 6MWD

Clinically relevant independent variables were selected from baseline characteristics to predict 6MWD. These variables were age, BMI, ejection fraction, history of diabetes, history of coronary artery disease, history of coronary revascularizations, history of peripheral vascular disease, diastolic HF, beta-blocker usage, ACE/ARBs usage, history of ICD implantation, and telemonitoring usage. Multinomial logistic regression analysis was performed by entering variables through an entry method into the main effects model. The category of participants with no

meaningful change in 6MWD was selected as the reference category. Multinomial regression analysis was performed on 181 participants (20 were missing data; see Table 28).

The probability of the model's chi-square ( $\chi^2 = 48.699$ ,  $df = 24$ ,  $p = 0.002$ ) was statistically significant ( $p < 0.05$ ). The null hypothesis that there was no difference between the model without independent variables and the model with independent variables was rejected. The existence of a relationship between the independent variables and the dependent variable was supported (see Table 29).

The benchmark to characterize this multinomial logistic regression model as useful was a 25% improvement over the rate of accuracy achievable by chance alone. The proportional by-chance accuracy rate was computed by calculating the proportion of cases for each group (improved, deteriorated, and no change) based on the number of cases in each group in the marginal percentages of the case processing summary, as previously described, and then squaring and summing the proportion of cases in each group:  $0.377^2 + 0.210^2 + 0.171^2$ . The proportional by-chance accuracy criterion was 44.5% (1.25 x 35.6% = 44.5%). The classification accuracy rate of this model should be greater than the proportional by chance accuracy criterion of 44.5%.

The chi-square statistic was the difference in  $-2 \log$  likelihoods between the final model and a reduced model. The reduced model was formed by omitting an effect from the final model. The null hypothesis was that all parameters of that effect are zero. The reduced model was equivalent to the final model because omitting the effect did not increase the degrees of freedom. The likelihood ratio tests for BMI, ejection fraction, history of diabetes, beta-blocker usage were statistically significant ( $p < 0.05$ ) and

Table 28

Case Processing Summary for Predictors of Clinically Meaningful Change in 6MWD

		<u>N</u>	<u>Marginal percentage</u>
Clinically meaningful change in 6MWD	Improved	68	37.6%
	Deteriorated	38	21.0%
	No change	75	41.4%
Diabetes mellitus–II (DM–II)	Yes	75	41.4%
	No	106	58.6%
History of CAD	Yes	139	76.8%
	No	42	23.2%
Revascularization procedures	Yes	77	42.5%
	No	104	57.5%
Peripheral vascular disease (PVD)	Yes	45	24.9%
	No	136	75.1%
Diastolic HF	Yes	25	13.8%
	No	156	86.2%
Beta-blocker usage	Yes	150	82.9%
	No	31	17.1%
ACE–I/ARBs usage	Yes	165	91.2%
	No	16	8.8%
ICD implantation	Yes	53	29.3%
	No	128	70.7%
Randomization	Technology-driven disease management	82	45.3%
	Standard medical care	99	54.7%
Valid		181	100.0%
Missing		20	
Total		201	
Subpopulation		181 <sup>a</sup>	

<sup>a</sup>The dependent variable had only one value observed in 181 (100.0%) subpopulations.

Table 29

Model Fitting Information for Predictors of Clinically Meaningful Change in 6MWD

Model	Model-fitting		Likelihood ratio tests		
	criteria				
	<u>-2 log</u> <u>likelihood</u>	<u>Chi-square</u>	<u>df</u>	<u>Sig.</u>	
Intercept only	383.923				
Final	335.224	48.699	24	.002	

suggestive of significant prediction of clinically meaningful change in 6MWD (see Table 30).

BMI, LVEF, diabetes, and beta-blocker usage were significant predictors for differentiating clinically meaningful improvement in 6MWD vs. no clinically meaningful change in 6MWD, whereas beta-blocker usage was the only variable significant enough to differentiate clinically meaningful deterioration in 6MWD vs. no clinically meaningful change in 6MWD (see Table 31).

The classification accuracy for clinically meaningful improvement in 6MWD was 65% and for no meaningful change in 6MWD, was 69%. The classification accuracy for clinically meaningful deterioration in 6MWD was 16%. The overall prediction accuracy of this model was 56.4%, which was greater than that of classification accuracy by chance alone (i.e., 44.5%; see Table 32). Therefore, this prediction model was valid to predict clinically meaningful change in 6MWD.

Predictors of Clinically Meaningful Change in 6MW

Clinically relevant independent variables were selected from baseline characteristics to predict 6MW. The variables included age, BMI, ejection fraction,

Table 30

Likelihood Ratio Tests for Predictors of Clinically Meaningful Change in 6MWD

Effect	Model-fitting	Likelihood ratio tests		
	criteria <u>-2 log</u> <u>likelihood</u>	<u>Chi-square</u>	<u>df</u>	<u>Sig.</u>
Intercept	3.352E2	.000	0	.
Age	337.427	2.203	2	.332
Body-mass index (BMI)	346.897	11.673	2	.003
Left ventricular ejection fraction (LVEF)	348.680	13.456	2	.001
Diabetes mellitus-II (DM-II)	342.222	6.998	2	.030
History of CAD	336.610	1.386	2	.500
History of Revascularizations	337.002	1.778	2	.411
History of peripheral vascular disease (PVD)	336.134	.910	2	.635
Diastolic HF	339.441	4.217	2	.121
Beta-blocker usage	341.963	6.739	2	.034
ACE-I/ARBs usage	336.529	1.306	2	.521
History of ICD implantation	340.091	4.867	2	.088
Telemonitoring	336.757	1.533	2	.465

Table 31

Parameter Estimates for Predictors of Clinically Meaningful Change in 6MWD

Clinically meaningful change in 6MWD <sup>a</sup>		B	SE	Wald	Sig.	Exp <sup>b</sup>	95% confidence interval for exp <sup>b</sup>	
							Lower bound	Upper bound
Improved	Intercept	10.246	3.295	9.671	.002			
	Age	-.045	.031	2.136	.144	.956	.901	1.015
	BMI	-.134	.042	9.919	.002	.875	.805	.951
	LVEF	-.071	.020	12.307	.000	.932	.896	.969
	DM-II	.824	.411	4.026	.045	2.281	1.019	5.103
	CAD	-.578	.505	1.307	.253	.561	.208	1.511
	Revasculari- zation	-.437	.429	1.039	.308	.646	.278	1.497
	PVD	.346	.429	.653	.419	1.414	.610	3.276
	Diastolic HF	1.650	.845	3.809	.051	5.205	.993	27.280
	Beta-blocker usage	-1.241	.573	4.695	.030	.289	.094	.888
	ACE-I/ARBs usage	-.269	.652	.170	.680	.764	.213	2.740
	ICD implanta- tion	.774	.436	3.148	.076	2.169	.922	5.100
	Telemonitoring	-.176	.401	.193	.660	.839	.382	1.839
Deterior- ated	Intercept	2.900	3.435	.713	.398			
	Age	-.016	.033	.225	.635	.985	.923	1.050
	BMI	-.026	.041	.396	.529	.975	.900	1.056
	LVEF	-.033	.021	2.387	.122	.967	.927	1.009
	DM-II	-.341	.475	.515	.473	.711	.280	1.805
	CAD	-.144	.543	.071	.790	.866	.299	2.508
	Revasculari- zation	-.567	.481	1.387	.239	.567	.221	1.457
	PVD	-.071	.502	.020	.888	.931	.348	2.493
	Diastolic HF	1.013	.812	1.555	.212	2.753	.560	13.523
	Beta-blocker usage	-1.207	.569	4.498	.034	.299	.098	.912
	ACE-I/ARBs usage	.692	.877	.623	.430	1.998	.358	11.153
	ICD implanta- tion	-.256	.550	.217	.641	.774	.263	2.276
	Telemonitoring	.389	.437	.792	.373	1.475	.627	3.474

<sup>a</sup>The reference category is no change. <sup>b</sup>The standard error of predictor variables in the model was far less than 2.0, indicating that there were no numerical problems with predictor variables in the final model.

Table 32

Classification Accuracy of Predictors for Clinically Meaningful Change in 6MWD

Observed	Predicted			Percent correct
	Improved	Deteriorated	No change	
Improved	44	4	20	64.7%
Deteriorated	11	6	21	15.8%
No change	20	3	52	69.3%
Overall percentage	41.4%	7.2%	51.4%	56.4%

history of diabetes, history of coronary artery disease, history of coronary revascularizations, history of peripheral vascular disease, diastolic HF, beta-blocker usage, ACE/ARBs usage, history of ICD implantation, and telemonitoring usage. Multinomial logistic regression analysis was performed by entering variables through an entry method into the main effects model. The category of participants with no meaningful change in 6MW was selected as the reference category. Multinomial regression analysis was performed on 161 participants (40 were missing data; see Table 33).

The benchmark to characterize this multinomial logistic regression model as useful was a 25% improvement over the rate of accuracy achievable by chance alone. The proportional by-chance accuracy rate was computed by calculating the proportion of cases for each group (improved, deteriorated, and no change) based on the number of cases in each group in the marginal percentages of case processing summary, as previously described, and then squaring and summing the proportion of cases in each group:  $0.342^2 + 0.193^2 + 0.466^2$ . The proportional by-chance accuracy criterion was

Table 33

Case Processing Summary for Predictors of Clinically Meaningful Change in 6MW

		<u>N</u>	<u>Marginal percentage</u>
Clinically meaningful change in 6MW	Improved	55	34.2%
	Deteriorated	31	19.3%
	No change	75	46.6%
DM-II	Yes	67	41.6%
	No	94	58.4%
History of CAD	Yes	126	78.3%
	No	35	21.7%
Revascularization procedures	Yes	68	42.2%
	No	93	57.8%
Peripheral vascular diseases	Yes	41	25.5%
	No	120	74.5%
Diastolic HF	Yes	23	14.3%
	No	138	85.7%
Beta-blockers	Yes	131	81.4%
	No	30	18.6%
ACE inhibitors/ARBs	Yes	146	90.7%
	No	15	9.3%
ICD	Yes	45	28.0%
	No	116	72.0%
Randomization	Technology-driven disease management	72	44.7%
	Standard medical care	89	55.3%
Valid		161	100.0%
Missing		40	
Total		201	
Subpopulation		161 <sup>a</sup>	

<sup>a</sup>The dependent variable has only one value observed in 161 (100.0%) subpopulations.



46.4% ( $1.25 \times 37.1\% = 46.4\%$ ). The classification accuracy rate of this model should be greater than the proportional by-chance accuracy criteria of 46.4%.

The probability of the model's chi-square ( $\chi^2 = 44.177$ ,  $df = 24$ ,  $p = 0.007$ ) was statistically significant ( $p < 0.05$ ). The null hypothesis that there was no difference between the model without independent variables and the model with independent variables was rejected. The existence of a relationship between the independent variables and the dependent variable was supported (see Table 34).

Table 34

Model Fitting Information for Predictors of Clinically Meaningful Change in 6MW

Model	Model-fitting criteria	Likelihood ratio tests		
	-2 log likelihood	Chi-square	df	Sig.
Intercept only	334.875			
Final	290.698	44.177	24	.007

The likelihood ratio tests for BMI, ejection fraction, beta-blocker usage, and ICD implantation were statistically significant ( $p < 0.05$ ), suggestive of significant prediction of clinically meaningful change in 6MW (see Table 35). BMI, LVEF, beta-blocker usage, and ICD implantation were significant predictors for differentiating clinically meaningful improvement in 6MW vs. no clinically meaningful change in 6MW, whereas beta-blocker usage was the only variable significant enough to differentiate clinically meaningful deterioration in 6MW vs. no clinically meaningful change in 6MW (see Table 36).

Table 35

Likelihood Ratio Tests for Predictors of Clinically Meaningful Change in 6MW

Effect	Model-fitting criteria	Likelihood ratio tests		
	<u>-2 log likelihood of reduced model<sup>a</sup></u>	<u>Chi-square</u>	<u>df</u>	<u>Sig.</u>
Intercept	2.907E2	.000	0	.
Age	292.934	2.236	2	.327
BMI	297.758	7.060	2	.029
LVEF	298.792	8.094	2	.017
DM-II	294.636	3.938	2	.140
History of CAD	291.213	.515	2	.773
History of revascularizations	294.878	4.180	2	.124
History of peripheral vascular disease (PVD)	294.099	3.401	2	.183
Diastolic HF	293.209	2.511	2	.285
Beta-blocker usage	297.250	6.551	2	.038
ACE-I/ARBs usage	290.846	.148	2	.929
History of ICD implantation	298.369	7.671	2	.022
Telemonitoring	291.581	.883	2	.643

Note. The chi-square statistic was the difference in -2 log likelihoods between the final model and a reduced model. The reduced model was formed by omitting an effect from the final model. The null hypothesis was that all parameters of that effect were zero.

<sup>a</sup>This reduced model was equivalent to the final model because omitting the effect did not increase the degrees of freedom.

Table 36

Parameter Estimates for Predictors of Clinically Meaningful Change in 6MW

Clinically meaningful change in 6MW <sup>a</sup>		B	SE	Wald	Sig.	Exp <sup>b</sup>	95% confidence interval for exp <sup>b</sup>	
							Lower bound	Upper bound
Improved	Intercept	8.404	3.446	5.948	.015			
	Age	-.048	.033	2.150	.143	.953	.894	1.016
	BMI	-.091	.043	4.426	.035	.913	.838	.994
	LVEF	-.057	.021	7.473	.006	.945	.907	.984
	DM	.741	.429	2.992	.084	2.098	.906	4.860
	CAD	-.348	.520	.447	.504	.706	.255	1.957
	Revascularizations	-.862	.450	3.671	.055	.422	.175	1.020
	PVD	.047	.442	.011	.915	1.048	.441	2.493
	Diastolic HF	1.226	.889	1.901	.168	3.406	.597	19.445
	Beta-blocker usage	-1.261	.577	4.773	.029	.283	.091	.878
	ACE-I/ARBs usage	.058	.686	.007	.933	1.059	.276	4.061
	ICD implantation	1.117	.466	5.748	.017	3.057	1.226	7.619
	Telemonitoring	-.062	.420	.022	.882	.940	.412	2.142
Deteriorated	Intercept	2.270	3.692	.378	.539			
	Age	-.026	.036	.496	.481	.975	.907	1.047
	BMI	.024	.042	.333	.564	1.025	.943	1.113
	LVEF	-.031	.024	1.734	.188	.969	.926	1.015
	DM	-.145	.516	.079	.779	.865	.315	2.378
	CAD	-.007	.617	.000	.991	.993	.296	3.330
	Revascularizations	-.659	.535	1.519	.218	.517	.181	1.476
	PVD	-1.005	.627	2.567	.109	.366	.107	1.252
	Diastolic HF	1.014	.883	1.318	.251	2.757	.488	15.567
	Beta-blocker usage	-1.181	.596	3.928	.048	.307	.095	.987
	ACE-I/ARBs usage	.331	.882	.140	.708	1.392	.247	7.842
	ICD implantation	-.235	.635	.138	.711	.790	.228	2.742
	Telemonitoring	.394	.479	.675	.411	1.483	.579	3.794

<sup>a</sup>The reference category is no change. <sup>b</sup>The SE of predictor variables in the model was far less than 2.0, indicating that there were no numerical problems with predictor variables in the final model.

The classification accuracy for clinically meaningful improvement in 6MW was 62% and for no meaningful change in 6MW, was 77%. The classification accuracy for clinically meaningful deterioration in 6MW was 23%. The overall prediction accuracy of this model was 62%, which was greater than that of classification accuracy by chance alone (i.e., 46.4%; see Table 37). Therefore, this prediction model was valid to predict clinically meaningful change in 6MW.

Table 37

Classification Accuracy for Predictors of Clinically Meaningful Change in 6MW

Observed	Predicted			Percent correct
	Improved	Deteriorated	No change	
Improved	34	0	21	61.8%
Deteriorated	5	7	19	22.6%
No change	17	0	58	77.3%
Overall percentage	34.8%	4.3%	60.9%	61.5%

Predictors for Clinically Meaningful Change in MLHF Total Score

Clinically relevant independent variables were selected from baseline characteristics to predict clinically meaningful change in MLHF total score. These variables included age, BMI, ejection fraction, history of diabetes, history of coronary artery disease, history of coronary revascularizations, history of peripheral vascular disease, diastolic HF, beta-blocker usage, ACE/ARBs usage, history of ICD implantation, and telemonitoring usage. Multinomial logistic regression analysis was performed by entering variables through an entry method into the main effects model. The category of

participants with no clinically meaningful change in MLHF total score was selected for the reference category.

Multinomial regression analysis was performed on 181 participants (20 were missing data). The probability of chi-square statistic for the model with the variables was not statistically significant and did not meet the model-fitting criteria. The independent variables with low chi-square static values were dropped from the model, and the reduced model (meeting the fitting criteria) was deduced. The reduced model, using age, history of diabetes, history of coronary revascularization procedures, and telemonitoring and meeting the model-fitting criteria ( $\chi^2 = 16.34$ ,  $df = 8$ ,  $p = 0.037$ ) was performed on 201 participants (none were missing data; see Table 38).

The benchmark to characterize this multinomial logistic regression model as useful was a 25% improvement over the rate of accuracy achievable by chance alone. The proportional by-chance accuracy rate was computed by calculating the proportion of cases for each group (improved, deteriorated, and no change) based on the number of cases in each group in the marginal percentages of case processing summary, as previously described, and then squaring and summing the proportion of cases in each group:  $0.284^2 + 0.338^2 + 0.378^2$ . The proportional by-chance accuracy criterion was 42.25% ( $1.25 \times 33.79\% = 42.25\%$ ). The classification accuracy rate of this model should be greater than the proportional by-chance accuracy criteria of 42.25% (see Table 39). Though the model fit well, there were no statistically significant individual predictors for clinically meaningful change in MLHF total score (see Tables 40 & 41).

The classification accuracy for clinically meaningful improvement in MLHF total score was 23% and for no meaningful change in MLHF total score, was 58%. The

Table 38

Case Processing Summary for Predictors of Clinically Meaningful Change in MLHFTotal Score

		<u>N</u>	<u>Marginal percentage</u>
Clinically meaningful change in MLHF total score	improved	57	28.4%
	deteriorated	68	33.8%
	no change	76	37.8%
DM	yes	85	42.3%
	no	116	57.7%
Revascularization procedures	yes	86	42.8%
	no	115	57.2%
Randomization	Technology driven disease management	95	47.3%
	Standard medical care	106	52.7%
Valid		201	100.0%
Missing		0	
Total		201	
Subpopulation		129 <sup>a</sup>	

<sup>a</sup>The dependent variable had only one value observed in 93 (72.1%) subpopulations.

Table 39

Model Fitting Information for Predictors of Clinically Meaningful Change in MLHFTotal Score

Model	Model-fitting criteria	Likelihood ratio tests		
	-2 log likelihood	Chi-square	<u>df</u>	<u>Sig.</u>
Intercept only	360.376			
Final	343.992	16.384	8	.037

Table 40

## Likelihood Ratio Tests for Predictors of Clinically Meaningful Change in MLHF Total

Score

Effect	Model-fitting criteria	Likelihood ratio tests		
	-2 log likelihood of reduced model	Chi-square	df	Sig.
Intercept	3.440E2	.000	0	.
Age	349.117	5.125	2	.077
DM	348.498	4.506	2	.105
Telemonitoring	348.407	4.415	2	.110

<sup>a</sup>This reduced model was equivalent to the final model because omitting the effect did not increase the degrees of freedom.

Table 41

## Parameter Estimates for Predictors of Clinically Meaningful Change in MLHF Total

Score

Clinically meaningful change in MLHF total score <sup>a</sup>						95% confidence interval for exp <sup>b</sup>	
						Lower bound	Upper bound
Improved	Intercept	-1.950	1.975	.975	.323		
	Age	.018	.025	.496	.481	1.018	.969 1.069
	Diabetes mellitus	.640	.368	3.030	.082	1.897	.923 3.899
	Revascularization	-.319	.366	.761	.383	.727	.355 1.488
	Telemonitoring	.417	.365	1.304	.254	1.517	.742 3.100
Deteriorated	Intercept	2.459	1.764	1.944	.163		
	Age	-.035	.023	2.330	.127	.966	.924 1.010
	Diabetes mellitus	-.101	.361	.078	.780	.904	.445 1.835
	Revascularization	-.691	.358	3.736	.053	.501	.248 1.010
	Telemonitoring	.729	.353	4.276	.039	2.073	1.039 4.138

<sup>a</sup>The reference category is no change. <sup>b</sup>This parameter is set to zero because it is redundant.

Table 42

Classification Accuracy for Predictors of Clinically Meaningful Change in MLHF Total Score

Observed	Predicted			Percent correct
	Improved	Deteriorated	No change	
Improved	13	14	30	22.8%
Deteriorated	8	29	31	42.6%
No change	14	18	44	57.9%
Overall percentage	17.4%	30.3%	52.2%	42.8%

classification accuracy for clinically meaningful deterioration in MLHF total score was 43%. The overall prediction accuracy of this model was 42.8%, which was greater than that of classification accuracy by chance alone (i.e., 42.28%; see Table 42). Therefore, this prediction model was valid to predict clinically meaningful change in MLHF total score.

Clinical Outcomes

Technology-driven HF management was shown to reduce the number of days of hospitalization from HF as the primary diagnosis, the number of days of hospitalization from other cardiovascular causes, the number of Medicare claims for inpatient admissions, the number deaths from all causes, the number of deaths from all cardiovascular causes, and the number of deaths from HF, compared to standard care. However, these differences were not statistically significant. Technology-driven HF management showed increases in the number of days of hospitalizations from noncardiovascular causes, number of ER visits, number of clinic visits for all causes,



number of clinical visits for all cardiovascular causes, and number of clinical visits from HF ( $p < 0.05$ ) and increased cost of clinic claims ( $p < 0.05$ ).

Therefore, technology-driven HF management reduced the number of days of hospitalization but increased the number of outpatient visits for all cardiovascular causes or HF. The standard care group had a higher number of days of hospitalization and a lower number of outpatient visits from cardiovascular causes or HF. Further analysis by combining these end points and deducing cost effectiveness was beyond the scope of this dissertation. The primary end points of the HCFA Demonstration Project were separately analyzed by primary investigators of the project (see Table 43).

#### Mortality

Out of 201 total patients who were included in final analysis, 6 out of 95 (6.3%) in the technology-driven HF management group and 10 out of 106 (9.4%) in the standard care group died from any cause (OR: 0.647, 95%; CI: 0.226–1.854). In terms of cardiovascular deaths, 5 out of 95 (5.3%) in the technology-driven HF management group and 9 out of 106 (8.5%) in the standard care group died from cardiovascular causes (OR: 0.599, 95%; CI: 0.193–1.854). Finally, 3 out of 95 (3.2%) patients in the technology-driven HF management group and 5 out of 106 (4.7%) in the standard care group died from HF (OR: 0.659, 95%; CI: 0.153–2.833). Technology-driven HF management showed a trend toward decreasing all causes of mortality, mortality from all cardiovascular causes, and mortality from worsening of HF (see Figures 11–13). However, the incremental impact of adding a telemonitoring system to the standard of care was not statistically significant ( $p > 0.05$ ).

Table 43

Clinical End Points for Technology-Driven Management Group vs. Standard Care Group

	Treatment group			Control group			t-test
	<u>N</u>	<u>Sum</u>	<u>Mean</u>	<u>N</u>	<u>Sum</u>	<u>Mean</u>	<u>p-value</u>
Days of hospitalization (all causes)	95	358	3.77	106	263	2.48	0.30
Days of hospitalization (CHF primary diagnosis)	95	116	1.22	106	158	1.49	0.69
Days of hospitalization (other CVD primary diagnosis)	95	172	1.81	106	195	1.84	0.97
Days of hospitalization (noncardiovascular primary diagnosis)	95	186	1.96	106	68	0.64	0.78
In-patient claims	95	68	0.72	106	90	0.85	0.41
ER claims	95	37	0.39	106	25	0.24	0.10
Clinic visits (all causes)	95	290	3.05	106	87	0.82	0.00
Clinic visits (CHF primary diagnosis)	95	81	0.85	106	21	0.20	0.01
Clinic visits (CVD primary diagnosis)	95	176	1.85	106	45	0.42	0.00
Deaths from all causes	95	6	0.06	106	10	0.09	0.41
Deaths from all cardiovascular causes	95	5	0.05	106	9	0.08	0.36
Deaths from HF	95	3	0.03	106	5	0.05	0.57
Cost of in-patient claims	95	620769	6534	106	675596	6373	0.93
Cost of ER claims	95	7437	78	106	7882	74	0.89
Cost of clinic claims	81	31512	389	90	10380	115	0.01

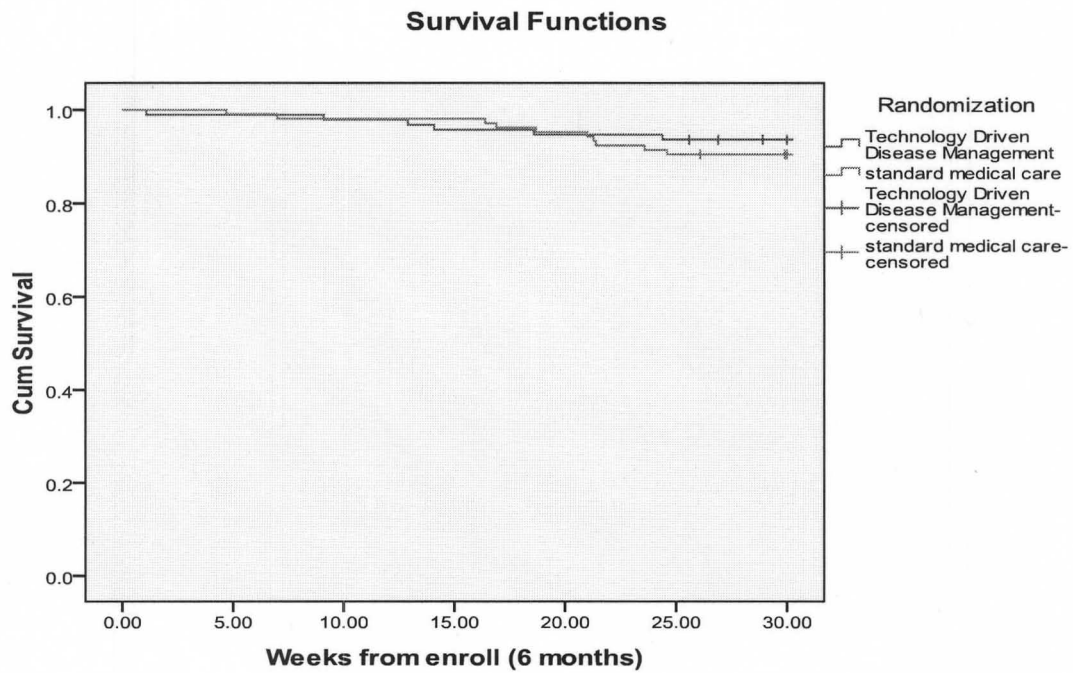


Figure 11. Kaplan-Meier Curve for Death from all Causes in 6 months.

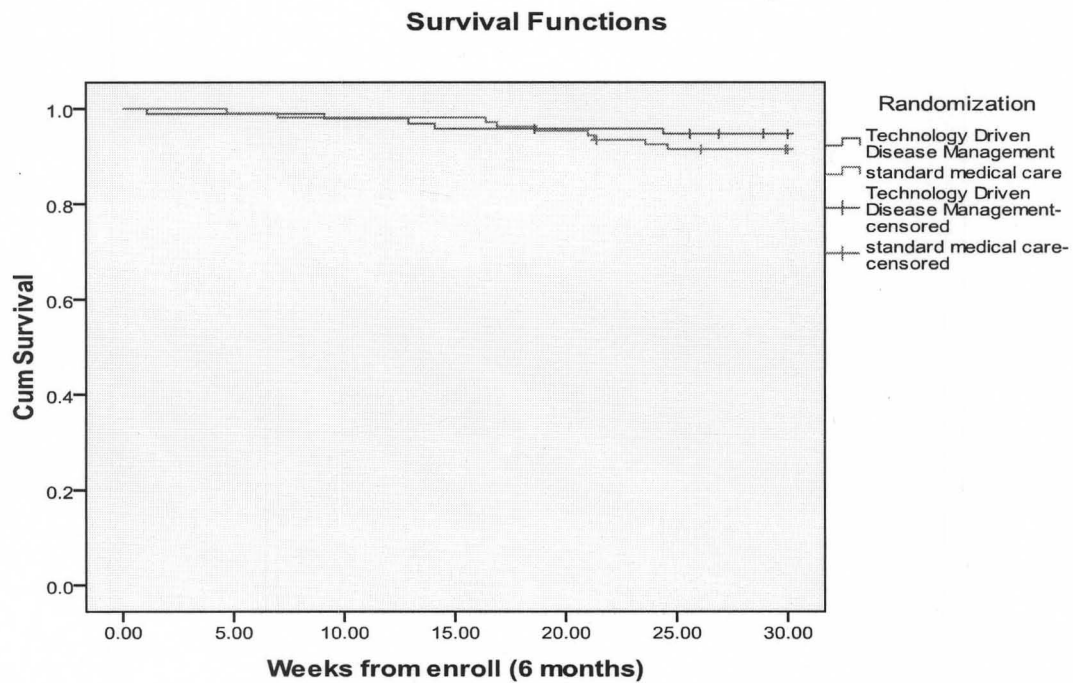


Figure 12. Kaplan-Meier Curve for Death from all Cardiovascular Causes.

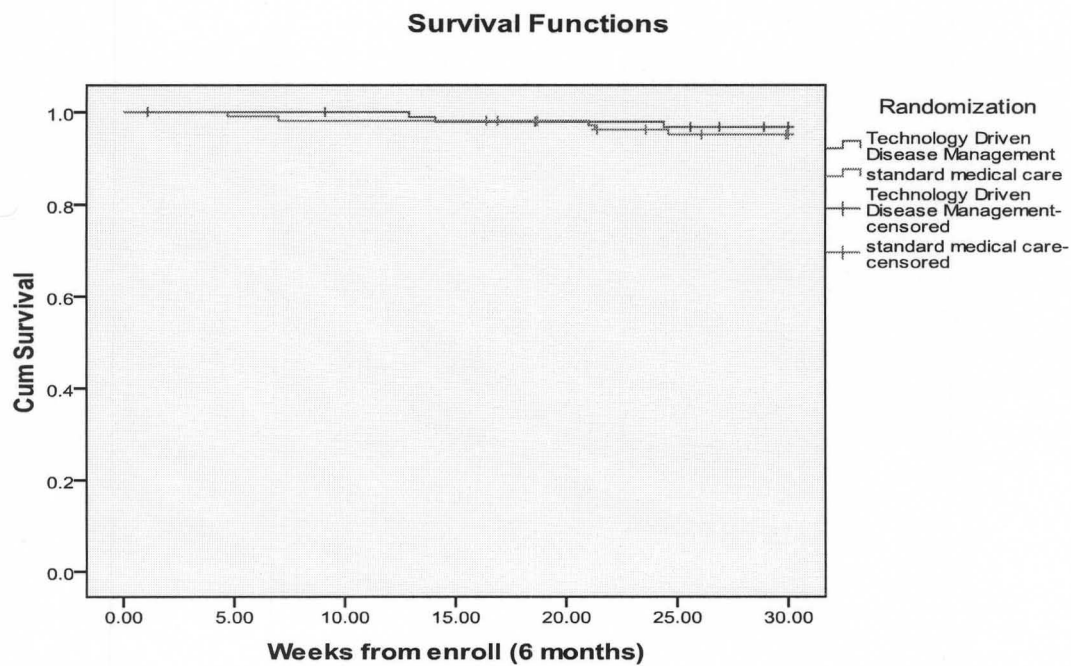


Figure 13. Kaplan-Meier Curve for Death from Heart Failure.

## CHAPTER V

### DISCUSSION

#### Summary of Results

This substudy was conducted based on a multicentered, randomized, phase III clinical trial of Medicare beneficiaries who were previously hospitalized with HF (NYHA Class II–IV) to evaluate the combined clinical and economic outcomes of adding a technology-driven HF monitoring system to standard HF management. The effect of technology-driven HF monitoring on functional capacity and QOL was evaluated by measuring the magnitude and direction of change at a 6-month follow-up. Traditional statistical techniques were translated to derive novel refined methods in order to define clinically meaningful difference in terms of functional capacity and QOL. The agreement between existing objective methods of assessing severity of HF in terms of NYHA class and 6MWD was tested. The clinically relevant factors predictive of clinically meaningful change in functional capacity and QOL were also deduced. Finally, the clinical outcomes of technology-driven HF management vs. standard care groups at the end of 6 months were compared.

#### Importance of Clinically Meaningful Change

Traditionally, treatment effects are evaluated by comparing changes resulting from the treatments under investigation and evaluating the statistical significance of that difference. In clinical trials, treatment effects often show statistically significant difference, but very little attention is paid to clinical significance. Statistical significance

is a function of the alpha level, the variance of the outcome measure, the magnitude of the difference, and the sample size. Clinicians often are interested in the variability of responses between and within groups. However, traditional methods are very limited in terms of determining the variability of response within a group due to lack of translation of traditional statistical methods to precise and refined novel statistical methods for determining clinically meaningful difference.

#### Need for Adjustment of Regression to the Mean

If patients, particularly those with extreme values, are repeatedly measured after treatment, the repeat measurement will be closer to the mean of the whole population than the initial measures were. In clinical trials, this phenomenon is often interpreted as showing the effect of the intervention or treatment; clinical investigators often assume that if they observe differences in baseline measurements among subjects, they can take the difference between the pre- and post-intervention/treatment values as the outcome value. However, it is incorrect to do so because imbalances at the baseline values will be reversed due to RTM. For example, subjects with a low degree of impairment at baseline will show less improvement, and subjects with a higher degree of impairment at baseline will show greater improvement.

In this study, functional capacity and QOL were repeatedly measured in terms of 6MWD or 6MW and MLHF scores on same subjects before and after treatment. At a 6-month follow-up, it was found that RTM was statistically significant, which meant that natural variation due to repeat data collection gave the appearance of real change. Hence, in clinical trials, if functional capacity or QOL is measured pre and post test, RTM must be estimated; otherwise, treatment effects will be overestimated and biased. If RTM is

significant, then RTM should be adjusted using the E–N method to determine true change from baseline.

#### Clinically Meaningful Change in Functional Capacity and Quality of Life

In this study, clinically meaningful difference in functional capacity and QOL was estimated after a 6-month follow-up after adjusting for RTM effects. The change in functional capacity in terms of 6MWD by 42 m (39 m–42 m) or 6MW by 3668 kg/m (3063 kg/m–3820 kg/m) and in QOL in terms of MLHF total score by 6.0 (5.0–7.0), physical dimension score by 4.0 (3.0–4.0), and emotional dimension score by 2.0 (2.0–3.0) from baseline was determined as either clinically meaningful improvement or deterioration.

#### Effect of Technology-Driven Heart Failure Management on Functional Capacity

Addition of twice-daily home telemonitoring of weight and HF symptoms to standard care for Medicare beneficiaries with chronic HF did not result in clinically meaningful change in functional capacity after 6 months. The null hypothesis failed to be rejected and the alternative hypothesis failed to be accepted. After stratification by NYHA class, it was found that technology-driven HF management retarded the change from baseline for a slightly higher proportion of patients in NYHA Class II than in Classes III and IV; a slightly higher proportion of patients in NYHA Classes III and IV than in Class II showed clinically meaningful improvement. In the standard care group, a slightly higher proportion of patients in NYHA Class II showed improvement and a slightly higher proportion of patients in Classes III and IV showed retarded change. Therefore, technology-driven HF management improved functional capacity for a slightly higher proportion of patients in NYHA Classes III and IV and stabilized functional

capacity for a slightly higher proportion of patients in NYHA Class II. However, none of these effects reached statistical significance at a 6-month follow-up. Therefore, there is no clinically meaningful incremental benefit in functional capacity from adding technology-driven HF monitoring to standard care.

#### Effect of Technology-Driven Heart Failure Management on Quality of Life

Addition of twice-daily home telemonitoring of weight and HF symptoms to standard care for Medicare beneficiaries with chronic HF did not show clinically meaningful change in QOL. The null hypothesis failed to be rejected, and the alternative hypothesis failed to be accepted. The standard care group showed retarded change, and the technology-driven HF management group showed clinically meaningful deterioration in a greater proportion of patients than the standard care group. After stratification by NYHA class, these trends persisted irrespective of NYHA class. However, none of these effects reached statistical significance. Therefore, there is no clinically meaningful benefit in QOL from adding technology-driven HF monitoring to standard care. Rather, technology-driven HF monitoring may be counterproductive and lead to deterioration in QOL, possibly by interfering with the daily routine of patients. Hence, patients may perceive technology-driven HF monitoring as impairing their QOL.

#### Effect of Technology-Driven Heart Failure Management on Clinical Outcomes

Technology-driven HF management did not show any statistically significant benefit for reducing the number of Medicare claims for inpatient admissions; the number of days of hospitalization from HF and other cardiovascular causes; and the number of deaths from all causes, all cardiovascular causes, and HF at the end of a 6-month follow-up. However, there was a statistically significant increase in the number of outpatient



clinic visits for all causes, all cardiovascular causes, and HF and the cost of Medicare claims for clinic visits at a 6-month follow-up. After 6 months, technology-driven HF management significantly increased resource utilization of outpatient care, and there was no statistically significant reduction in resource utilization of inpatient care when compared to standard care.

The intervention group received twice-daily monitoring of weight and symptoms by telemonitoring center nurses, which could have led to sending alerts to physicians early in the course of HF decompensation. A substantial proportion of the participants (25%) reported receiving care from a primary care physician without a cardiologist on board. In addition, the cardiologists who provided care to the participants may not have had training in advanced HF management. For these reasons, patients may have overutilized outpatient clinic visits. Participants in the standard care group did not report this type of communication between physician and patient, probably because the progression of the disease was recognized in the later part of decompensation process. Therefore, this group experienced more hospitalizations and total days of hospitalizations; however, these effects didn't reach statistical significance. At the same time, the standard care group experienced a statistically significant reduction in outpatient clinic visits at the end of 6 months compared to the technology-driven HF management group.

#### Agreement Between Objective Assessment of Functional Capacity and Classification of Heart Failure NYHA Criteria

To determine a patient's NYHA class, clinicians ask patients about their HF symptoms and the effect on exertion (i.e., how far they can walk before symptom onset).

However, there is no consistent method for determining NYHA functional class. The method of assessing NYHA class is arbitrary, and intraoperator variability is 66% (R. Claire et al., 2007). As mentioned earlier, 6MWD is a simple, cost-effective means of clinically assessing the functional capacity of patients with cardiac and pulmonary diseases. The NYHA system and 6MWD are regularly used as outcome measures in clinical trials and are even included in HF management guidelines (“ACC/AHA Guidelines,” 2009). Hence, the agreement between these two functional classifications was examined.

There is a lack of agreement between objective assessment of NYHA functional class and objective measurement of functional capacity (6MWD). The correlation coefficient between these two outcome measures was very poor because either patients are under- or over-reporting functional capacity to their physicians or physicians are misinterpreting patient class compared to objectively measured values. Still, 6MWD may not be the single best indicator of severity of HF among groups of patients of different ages, genders, BMIs, and comorbid conditions. This lack of consistency between NYHA class and 6MWD suggests the need to develop novel classification methods by designing norm-referenced equations comprising age, gender, BMI, comorbid conditions, and symptoms and signs of HF.

#### Prediction of Change in Quality of Life by Change in Functional Capacity

In this study, change refers to clinically meaningful change, not statistical change from baseline. The sensitivity and specificity of change in functional capacity to predict change in QOL was low, and Spearman correlation coefficients were very weak. Clinically meaningful improvement in functional capacity to predict clinically

meaningful improvement in QOL and clinically meaningful deterioration in functional capacity to predict clinically meaningful deterioration in QOL, and vice versa, were estimated. If there was improvement in functional capacity (6MWD), the probability of improvement in QOL (MLHF) was 56%. However, if there was improvement in QOL, the probability of improvement in functional capacity was only 44%. If there was deterioration in functional capacity, the probability of deterioration in QOL was 35%. However, if there was deterioration in QOL, the probability of deterioration in functional capacity was 55%.

The probability of change in functional capacity was more responsive to deterioration in QOL than probability of improvement in QOL: If there was no change in functional capacity, the probability of no change in QOL was 35%. If there was no change in QOL, the probability of no change in functional capacity was 31%. If the functional capacity was measured in terms of 6MW, it had similar probabilities of predicting improvement and deterioration patterns. But, if there was no change in functional capacity, the probability of no change in QOL was 53%. If there was no change in QOL, the probability of no change in functional capacity was 43%, which was 8% higher than that of no change in 6MWD.

Overall, the change in functional capacity is not a good predictor of similar directional change in QOL and vice versa. Therefore, if a clinician sees improvement in functional capacity, the probability of improvement in QOL is 56%. If there is deterioration in QOL, the probability of deterioration in functional capacity is 55%. In all other cases, it is difficult to predict the direction of change in QOL with respect to change in functional capacity and vice versa.

## Predictors of Change in Physical and Emotional Dimension Scores by Change in Functional Capacity

If there is change in functional capacity (6MWD), the probability of similar directional change in physical dimension score is 69% for improvement, 35% for deterioration, and 46% for no change; the probability of similar directional change in emotional dimension score is 100% for improvement, 64% for deterioration, and 45% for no change. If there is a change in physical dimension score, the probability of similar directional change in functional capacity is 30% for improvement, 31% for deterioration, and 71% for no change. If there is a change in emotional dimension score, the probability of similar directional change in functional capacity is 12% for improvement, 17% for deterioration, and 95% for no change.

In summary, change in functional capacity is a poor predictor of change in total QOL; however, functional capacity shows good predictive patterns for change in physical and emotional components of QOL. Improvement in functional capacity increases the probability of improvement in physical dimension score but not vice versa. Change in functional capacity increases the probability of similar directional change in emotional dimension score but not vice versa. Surprisingly, change in functional capacity better predicts change in emotional dimension score than physical dimension score. Similar patterns were observed with 6MW. It is not surprising that improvement in functional capacity predicts improvement in physical dimension score. Surprisingly emotional scores are influenced by physical endurance. The sense of accomplishment in improving functional capacity may improve patients' self esteem by providing them with positive feedback, which in turn improves the emotional dimension score.

Perception of QOL involves a complex interplay of multidimensional indices of life. A change in physical, emotional, or socioeconomic indices that leads to a change in the subscore of one dimension does not always translate into change in total score or functional capacity due to fluctuations in other item scores. Therefore, change in functional capacity is a poor predictor of change in total QOL score and vice versa. Specificity and negative predictive value of change in functional capacity are far better than sensitivity and positive predictive values to predict similar directional change in total QOL and its components.

#### Predictors of Clinically Meaningful Change for Functional Capacity

BMI, LVEF, type-2 diabetes, and beta-blocker usage are statistically significant predictors of clinically meaningful change in functional capacity (6MWD). BMI and LVEF are inversely related to improvement in functional capacity (i.e., the regression coefficients are negative). As BMI increases, improvement in functional capacity is less likely. As LVEF improves, functional capacity improves initially; once functional capacity reaches the ceiling effect, improvement in functional capacity is less likely.

Subjects with type 2 diabetes mellitus (OR 2.281, 95% of CI: 1.019 to 5.103) are more likely to improve in functional capacity. Diabetes itself doesn't improve functional capacity, but additional diabetics-related care processes might contribute to improvement in functional capacity. Subjects who use beta-blockers (OR 0.289, 95% of CI: 0.094 to 0.888) are less likely to show improvement and less likely to show deterioration (OR 0.299, 95% of CI: 0.98 to 0.912) in functional capacity. Therefore, beta-blockers play a significant role in stabilizing functional capacity.

BMI, LVEF, and beta-blocker usage show similar significant predictive patterns for predicting clinically meaningful change in 6MW. ICD implantation is also a significant predictor of functional capacity (6MW). Subjects with ICD implantation show clinically meaningful improvement in functional capacity (OR 3.06, 95% of CI: 1.23 to 2.14).

### Predictors of Change in Quality of Life

The reduced regression model, which contained the independent variables of age, type 2 diabetes mellitus, revascularization procedures, and technology-driven HF monitoring, satisfied model fit criteria. None of the independent variables surfaced as significant individual predictors of clinically meaningful improvement in QOL (MLHF score), whereas technology-driven HF monitoring was a statistically significant predictor of clinically meaningful deterioration in QOL. Twice-daily recording of weight, answering automated questions about symptoms, and receiving telephone calls from Alere nurses may interfere with the daily routines of patients with HF, who may perceive technology-driven HF monitoring as impairing their QOL.

### Comparison of Existing Studies

It has to be acknowledged that to date the majority of studies that test telephonic support (Riegel et al, 2002; Laramée, 2003; DeBusk, Miller and Parker, 2004; Gesica Inv, 2005; Riegel et al, 2005) and vital sign monitoring (deLusignan, 2001; Woodend et al, 2003; Capomolla et al, 2004) have not shown a mortality benefit in patients with chronic heart failure. Only a few studies have shown that telephonic support (Cleland et al, 2005) and vital sign monitoring (Cleland et al, 2005; Goldberg et al, 2002) reduced mortality. The difference in mix of patients and provider populations, geographical

settings, treatment modalities, and monitoring strategies might explain different outcomes in different studies.

Telephonic support and vital sign monitoring are two different strategies to monitor HF patients. Cleland (2005) and Goldberg (2002) showed there was a significant absolute reduction in mortality by 10.4% and 16% respectively through electronic monitoring of vital signs and symptoms. In our study the absolute reduction in mortality was only 3% which is not statistically significant. In Cleland's (2005) study primary care physicians delivered standard care for all heart failure patients. In the WHARF (Goldberg et al, 2002) trial, patients were sicker with more advanced heart failure (NYHA class III and IV) than in our sample and care was provided by cardiologists with advanced training in heart failure management. In our study 31% of patients were NYHA class II and a substantial proportion of patient care was provided by cardiologists without advanced training in heart failure or by primary care physicians without a cardiologist on board. However, the proportion of patients on beta-blockers and ACE-I/ARBs in our study was greater than in the WHARF trial (Goldberg et al, 2002) suggesting more widespread acceptance of the American College of Cardiology and the American Heart Association guidelines. Our study's patient population having less advanced heart failure would predict much lower mortality than observed in the WHARF trial and hence less opportunity for electronic monitoring to reduce the absolute death rate.

#### Strengths of the Study

The WHARF trial (Goldberg et al., 2002) was the first multi-centered, prospective trial conducted for patients with advanced HF (NYHA Classes III and IV). Researchers concluded that there was a significant mortality benefit from home-based telemonitoring

with a nurse-staffed call center (Goldberg et al., 2002). At the end of 6 months, 18.4% of participants in the standard care group and 8.0% of those in the telemonitored group died (Goldberg et al., 2002). Thus, there was a 56.2% relative risk reduction in mortality for the telemonitored group of patients with advanced HF (Goldberg et al., 2002). However, there was no significant difference between the two groups with respect to length of stay, number of hospitalizations, and utilization of resources (Goldberg et al., 2002).

The results of the WHARF study were attributable to patients with advanced HF (NYHA Classes III and IV) but not to NYHA Class II patients because the study was conducted for NYHA class III and IV patients only (Goldberg et al., 2002). Only cardiologists with special training in advanced HF management provided care for patients recruited for the study (Goldberg et al., 2002). It is possible that these cardiologists better utilized the data generated through technology-driven HF monitoring than primary care physicians would have. Hence, reproducing a similar benefit in a real-world clinical setting may not be feasible because most HF care is provided by cardiologists without training in advanced HF management and primary care physicians without training in cardiology or without a cardiologist on board.

The HCFA HF Demonstration Project was conducted for Medicare beneficiaries classified as NYHA Class II–IV. Therefore, the findings of the sub-study are attributable to elderly patients with HF who belong to NYHA Class II–IV. HF management was provided by cardiologists with or without training in advanced HF management or by primary care physicians with or without cardiologist on board. Therefore, the findings of the sub-study are attributable to the current real-world clinical setting of HF management.



In the WHARF trial, due to differential drop-outs between groups (including death and drop-outs), the baseline QOL scores were carried forward to compare scores for all patients at 6 months, which might have introduced statistical bias (Goldberg et al., 2002). In the current study, patients who dropped out were excluded from the analysis and were compared with participants who were included in the analysis. Patients who died from HF or cardiovascular causes were included in the clinically meaningful deterioration group and given extremely worsened scores of QOL and functional capacity to minimize bias.

Other HF monitoring studies have used traditional distribution-based statistical methods and placed importance on statistical significance rather than clinical significance (Goldberg et al., 2002). The current study addressed the importance of RTM and adjusted RTM while measuring true change from the baseline. Also, the importance of measuring clinically meaningful change from the baseline, instead of using traditional methods of assessing statistically significant difference, was addressed. Traditional statistics were translated into more refined methods to provide meaningful information to clinicians on the effect of technology-driven HF management vs. standard care on functional capacity and QOL of patients with HF, clinically meaningful change, and factors affecting the degree of change and direction of change at the end of 6 months.

#### Limitations of the Study

The study was conducted for Medicare beneficiaries over 65 who were hospitalized for HF less than 1 year ago (i.e., NYHA Class II–IV). Therefore, the findings may not be attributable to patients with HF who are younger than 65. Patients were randomized from four geographic and hospital settings. The findings may not be

attributable to other centers where HF care is provided primarily by cardiologists, especially those with special training in advanced HF management.

Because a substantial proportion of patients (~25%) did not have a cardiologist on board and primary care physicians provided all of their HF management, there was overutilization of services and less efficient utilization of HF monitoring alerts. Physicians and patients were not blinded to the technology-driven HF monitoring system; hence, there was a likelihood of introducing treatment bias. Drop-outs, lost-to-follow cases, and missing data were unexpectedly high, and the compliance rate was low. The substudy was limited to a 6-month follow-up. Hence, the long-term effects of technology-driven HF monitoring are not known.

#### Post-Hoc Power Analysis

Post-hoc power is directly proportional to the effect size and inversely proportional to the p-value. If the effect size is trivial or the p-value is large, then the post-hoc power will be low. If the p-value is larger than 0.05, then the post-hoc power is likely to be less than 50%. Therefore, calculation of the p-value is not meaningful in the case of a negative clinical trial. Several authors have mentioned the inappropriateness of post-hoc power for negative end points and have suggested the use of confidence intervals instead of post-hoc power (Detsky & Sackett, 1985; Goodman & Berlin, 1994; Smith & Bates, 1992). Post-hoc power can provide meaningful information for positive end points (Detsky & Sackett, 1985; Goodman & Berlin, 1994; Smith & Bates, 1992).

In this study of change in functional capacity and QOL, there was no statistically significant clinically meaningful difference between the telemonitored vs. standard care groups. Based on the direction of change in functional capacity and QOL, the sample

was grouped into three groups: clinically meaningful improvement, deterioration, and no change. Confidence intervals for  $\chi^2$  static for above frequency distributions were calculated. The 95% confidence interval for the Chi-squared distribution with two degrees of freedom was 0.05–7.38. Chi-square statistics for distributions for clinically meaningful change in 6MWD (0.256), 6MW (0.131), MLHF score (3.877), physical dimension score (0.309), and emotional domain score (0.762) for the treatment vs. standard care group fell inside the 95% confidence interval.

Post-hoc power analysis was performed for multinomial logistic regression models and other clinical outcomes with statistical significance. Post-hoc power of the prediction model to predict clinically meaningful change in functional capacity measured in terms of 6MWD ( $\alpha = 0.05$ , number of predictors in the model = 12,  $R^2 = 0.236$ , valid sample size = 181) and 6MW ( $\alpha = 0.05$ , number of predictors in the model = 12,  $R^2 = 0.240$ , valid sample size = 161) was 99.9%. Post-hoc power of the model to predict clinically meaningful change in QOL measured in terms of MLHF score ( $\alpha = 0.05$ , number of predictors in the model = 4,  $R^2 = 0.201$ , valid sample size = 201) was 92%. Post-hoc power to detect mean difference between the telemonitored vs. standard care groups for the number of clinic visits for all causes, for HF, and for other cardiovascular causes was 97%, 65%, and 95%, respectively. The post-hoc power to detect mean difference in cost of clinic claims (treatment group vs. standard care group) was 65%. Other well-powered studies are needed to further investigate underpowered clinical endpoints.

### Conclusions and Implications

1. Both the technology-driven HF management and standard care groups showed significant improvement in functional capacity and QOL from baseline after 6 months. However, the differences between groups were not statistically significant. Therefore, there is no meaningful benefit in terms of functional capacity and QOL from the addition of home telemonitoring systems to standard care. In fact, home telemonitoring systems may deteriorate QOL by interfering with patients' daily routine activities. In addition, there was a statistically significant increase in the total number of clinic visits, the number of clinic visits for HF, and the number of visits for other cardiovascular causes. In turn, there was a statistically significant increase in the cost of clinic (Medicare) claims.
2. In clinical trials, if outcomes are measured pre and post test, RTM must be estimated; otherwise, treatment effects will be overestimated and biased. If RTM is significant, then it should be adjusted to determine true change from baseline.
3. In clinical trials, investigators should focus on measuring clinically meaningful treatment effects rather than statistically significant difference. In this study, clinically meaningful change in functional capacity and QOL were defined according to often-used outcome measures in HF research. The change in functional capacity in terms of 6MWD by 42 meters (39 m–42 m) or 6MW by 3668 kg/m (3063 kg/m–3820 kg/m) and QOL in terms of MLHF total score by 6.0 (5.0–7.0), physical dimension score by 4.0 (3.0–4.0), emotional dimension score by 2.0 (2.0–3.0) from baseline were determined to be either clinically

meaningful improvement or deterioration based on the direction of change for elderly patients with chronic HF.

4. Change in functional capacity is a good predictor of change in emotional dimension score and a fair predictor of change in physical dimension score but not predictive of change in overall QOL. If clinicians target improvement in functional capacity, they can improve the emotional component and, to a fair extent, the physical component, but not vice versa. Striving for clinically meaningful improvement in functional capacity may be a clinically meaningful strategy for dealing with depression in patients with chronic HF.
5. BMI, LVEF, beta-blocker usage, and ICD implantation are the best clinically relevant predictors of change in functional capacity. Strategies should be targeted to decrease BMI and improve cardiac index, rather than focusing on LVEF to improve functional capacity. Beta-blockers stabilize functional capacity, and ICD implantation improves functional capacity. There are no best predictors of improvement in QOL.
6. There is a lack of agreement between objective assessment of NYHA functional class and objective measurement of functional capacity (6MWD). The correlation coefficient between these two outcome measures is very poor. Utilization of NYHA class change as an outcome measure in clinical trials for measuring treatment effect may lead to over- or under-estimation.

#### Future Directives for Research

Due to lack of consistency and a high degree of operator variability when classifying patients with HF using the NYHA method, there is an imminent need for

development of novel classification methods by designing norm-referenced equations comprising clinical parameters predictive of cardiac function or index. More explorative research should be performed to identify diabetic care processes that might show clinically meaningful improvement in functional capacity for the diabetic subset of patients with HF. More aggressive and frequent monitoring of patients appears to be counterproductive and to lead to noncompliance, frequent drop-outs, and loss to follow up.

Call center nurses were not authorized to independently optimize medications, even in consultation with a cardiologist or primary care physician. Therefore, instead of twice-daily monitoring by call center nurses, less frequent home monitoring by healthcare personnel trained in cardiology who are authorized to perform optimization/adjustment of medications either independently or in consultation with a cardiologist might lead to better resource utilization and reduction of ER visits, outpatient clinic visits, and inpatient hospitalizations. On the other hand, if standard of care is provided while adhering to HF guidelines and closely monitoring HF clinic follow-ups, then there may not be room for further clinically meaningful benefit through home telemonitoring.

Analysis of combined clinical endpoints and costs should be performed to determine the benefits of technology-driven HF management and HF monitoring strategies. Further analysis should be conducted to evaluate efficient utilization of information gathered through technology-driven HF monitoring and subsequent resource utilization by primary care physicians vs. cardiologists with and without advanced training in HF in providing HF care.

## REFERENCES

- Abraham, W. T., Fisher, W. G., Smith, A. L., Delurgio, D. B., Leon, A. R., Loh, E., ...  
Packer, M. (2002). Cardiac resynchronization in chronic heart failure. *New England Journal of Medicine*, 346, 1845–1853.
- ACC/AHA guidelines for the evaluation and management of chronic heart failure in the adult: Executive summary. (2001). *Circulation*, 104, 2996–3007.
- ACC/AHA guidelines for the evaluation and management of heart failure in the adult: Focused Update. (2009). *Circulation*, 1977-2013.
- Ahring, K. K., Ahring, J. P., Joyce, C., & Farid, N. R. (1992). Telephone modem access improves diabetes control in those with insulin requiring diabetes. *Diabetes Care*, 15, 971–975.
- Altman, D. G., & Dore, C. L. (1991). Baseline comparisons in randomized clinical trials. *Statistics in Medicine*, 10, 797–802.
- American Heart Association. (2002). *Heart and stroke facts: 2002. Statistical supplement*. (2002). Dallas, TX: Author.
- American Heart Association. (2005). *Heart disease and stroke statistics—2005 update*. Dallas, TX: Author.
- American Heart Association. (2006). Heart disease and stroke statistics—2006 update: A report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation*, 113, e85–e151.

- American Heart Association. (2008). *Annual scientific sessions*. New Orleans, LA: Author.
- American Heart Association. (2009). Heart disease and stroke statistics—2009 update: A report from the American Heart Association Statistics Subcommittee. *Circulation*, 119, e21–e181.
- American Thoracic Society. (2002). ATS statement: Guidelines for 6-minute walk test. *American Journal of Respiratory and Critical Care Medicine*, 166, 111–117.
- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice-Hall.
- Barnett, A. G., Van Der Pol, J. C., & Dobson, A. J. (2005). Regression to the mean: What it is and how to deal with it. *International Journal of Epidemiology*, 34, 215–220.
- Bennett, S. J., Oldridge, N. B., Eckert, G. J., Embree, J. L., Browning, S., Hou, N., & Deer, M. (2002). Discriminant properties of commonly used quality of life measures in heart failure. *Quality of Life Research*, 11, 349–359.
- Billard, A., Rohmer, V., Roques, M. A., Joseph, M. G., Suraniti, S., & Giraud, P. (1991). Telematic transmission of computerized blood glucose profiles for IDDM patients. *Diabetes Care*, 14, 130–134.
- Bittner, V., Weigner, D. H., Yusuf, S., Rogers, W. J., McIntyre, K. M., Bangdiwala, S. I., ... Guillete, M. (1993). Prediction of mortality and morbidity with 6-minute walk test in patients with left ventricular dysfunction. *JAMA*, 270, 1702–1707.
- Blyth, F. M., Lazarus, R., Ross, D., Price, M., Cheuk, G., & Leeder, S. R. (1997). Burden and outcomes of hospitalization for congestive heart failure. *Medical Journal of Australia*, 167(2), 67–70.



- Bondmass, M., Benatar, J. D., Castro, G., & Avitall, B. (2001). A prospective randomized study comparing outcomes and outpatient care delivery methods for chronic heart failure. *Journal of the American College of Cardiology*, 37, 1A–648A.
- Bondmass, M., Malhotra, V., Castro, G., & Avitall, B. (1999). The long-term effect of telemedicine intervention on heart failure admissions and length of stay. *Journal of Cardiac Failure*, 5, 78.
- Bradburn, N. (1969). *The structure of psychological well-being*. Chicago, IL: Aldine.
- Brater, D. C. (1997). Diuretic therapy. *New England Journal of Medicine*, 336, 525–533.
- Bullinger, M. (1997). Health related quality of life and subjective health. Overview of the status of research for new evaluation criteria in medicine. *Psychother Psychosom Med Psychol*, 47, 76–91.
- Bulmer, M.G. (1965). *Classic principles of statistics*. New York: Dover Publications.
- Cafagna, D., Ponte, E., & Burri, R. (1997). The concept of quality of life in cardiac failure. *Minerva Medicine*, 88(4), 151–162.
- Cahlin, L. P., Mathier, M. A., Semigran, M. J., Dec, G. W., & Disavo, T. G. (1996). The six-minute walk test to predict peak oxygen uptake and survival in patients with advanced heart failure. *Chest*, 110, 325–332.
- Capomolla, S., Pinna, G. D., La Rovere, M. T., Maestri, R., Ceresa, M.,.....& Corebelli, F. (2004). Heart failure case disease management program: a pilot study of home telemonitoring versus usual care. *European Heart Journal*, 6, F91–F98.

- Carter, R., Holiday, D. B., Nwasuruba, C., Stocks, J., Grothues, C., & Tiep, B. (2003). 6-minute walk work for assessment of functional capacity in patients with COPD. *Chest, 123*, 1408–1415.
- Cella, D., Eton, D. T., & Lai, J.-S. (2002). Combining anchor and distribution-based methods to derive minimal clinically important differences on the functional assessment of cancer therapy (FACT) anemia and fatigue scales. *Journal of Pain Symptom Management, 24*, 547–561.
- Chrysegolos, E. T., Gemme, D., Coleman, K., Cheng, E. S., & Meyer, T. E. (1999). Telemonitoring devices further improve outcomes of a multidisciplinary heart failure outpatient program. *Journal of Cardiac Failure, 5*, 73.
- Chuang, M. L., Lin, I. F., & Wasserman, K. (2001). The body weight–walking distance product as related to lung function, anaerobic threshold and peak VO<sub>2</sub> in COPD patients. *Respiratory Medicine, 95*, 618–626.
- Claire, G., Martin, C. B., Joseph, C. C., Tamiza, P., & Stephen, P. M. (2005). Establishing a minimally important difference in 6-minute walk distance and SF-36 among patients with pulmonary arterial hypertension. *Chest, 128*, 365S.
- Claire, R., Cathy, B., Justin, D., Whinnett, Z., Manisty, C., Sutton, R., ... Francis, D. (2007). Limitations of the New York Heart Association functional classification system and self-reported walking distances in chronic heart failure. *Heart, 93*, 476–482.
- Cleland, J. G. F. (2000). Improving patient outcomes in heart failure: Evidence and barriers. *Heart, 84*, i8–i10.

- Cleland, J. G., Louis, A. A., Rigby, A. S., Janssens, U., & Balk, A. H. (2005). Noninvasive home telemonitoring for patients with heart failure at high risk of recurrent admission and death: the Trans-European Network-Home-Care Management System (TEN-HMS) study. *Journal of American College of Cardiology*, 45, 1654–65.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York, NY: Academic Press.
- Cohn, J. N., Archibald, D. G., Ziesche, S., Franciosa, J. A., Harston, W. E., Tristani, F. E., & Dunkman, W. B. (1986). Effect of vasodilator therapy on mortality in chronic congestive heart failure. Results of a Veterans Administration cooperative study. *New England Journal of Medicine*, 314, 1547–1552.
- Cohn, J. N., Johnson, G., Ziesche, S., Cobb, F., Francis, G., Tristani, F...Smith, R. (1991). A comparison of enalapril with hydralazine-isosorbide dinitrate in the treatment of chronic congestive heart failure. *New England Journal of Medicine*, 325, 303–310.
- Criteria Committee NYHA. (1964). *Diseases of the heart and blood vessels: Nomenclature and criteria for diagnosis*. Boston, MA: Little Brown.
- Cooper, C. B. (2001). Exercise in chronic pulmonary disease: Aerobic exercise prescription. *Medicine & Science in Sports & Exercise*, 33, S671–S679.
- Cordisco, M. E., Beniaminovitz, A., Hammond, K., & Mancini, D. (1999). Use of telemonitoring to decrease the rate of hospitalization in patients with severe congestive heart failure. *American Journal of Cardiology*, 84, 860–862.

- Crosby, R. D., Kolotkin, R. L., & Williams, R. G. (2003). Defining clinically meaningful change in health-related quality of life. *Journal of Clinical Epidemiology*, 56, 395–407.
- Crosby, R. D., Kolotkin, R. L., & Williams, R. G. (2004). An integrated method to determine meaningful changes in health related quality life of life. *Journal of Clinical Epidemiology*, 57, 1153–1160.
- Curran, D., Molenbergs, G., Fayers, P., & Machin, D. (1998). Incomplete quality of life data in randomized trials: Missing forms. *Statistics in Medicine*, 17, 697–709.
- Davis, R. C., Hobbs, R. F. D., Kenkre, J. E., Roalfe, A. K., McLeod, S., Hare, R., & Davies, M. K. (1999). Quality of life in heart failure and others conditions. *Journal of the American College of Cardiology*, 33, 211A.
- DeBusk, R. F., Miller, N. H., & Parker, K. M. (2004). Care management for low-risk patients with heart failure: a randomized, controlled trial. *Annals of Internal Medicine*, 141, 606–13.
- Deering, M., Baines, B., Christianson, C., & Milner, J. (2002). Patients and providers evaluate daily home weight and symptom monitoring for CHF management. *Journal of Cardiac Failure*, 8, S97.
- Delgado, D. H., Rao, V., Ross, H. J., Verma, S., & Smedira, N. (2002). Mechanical circulatory assistance: State of the art. *Circulation*, 106, 2046–2050.
- Demers, C., McKelvie, R. S., Naggassa, A., & Yusuf, S. (2001). Reliability, validity, and responsiveness of the six-minute walk test in patients with heart failure. *American Heart Journal*, 142, 698–703.

- Dempster, M., & Donnelly, M. (2000). Measuring health-related quality of life among people with ischaemic heart disease. *Heart*, 83, 641–644.
- Detsky, A. S., & Sackett, D. L. (1985). When was a negative clinical trial big enough?: How many patients you needed depends on what you found. *Archives of Internal Medicine*, 145, 709–712.
- de Lusignan, S., Althans, A., Wells, S., Johnson, P., Vandenberg, M., & Robinson, J. (2000). A pilot study of radiotelemetry for continuous cardiopulmonary monitoring of patients at home. *Journal of Telemedicine and Telecare*, 6(Suppl. 1), 119–122.
- de Lusignan, S. (2001). Compliance and effectiveness of 1 year's home telemonitoring. The report of a pilot study of patients with chronic heart failure. *European Journal of Heart Failure*, 3, 723–30.
- de Lusignan, S., Meredith, K., Wells, S., Leatham, E., & Johnson, P. (1999). A controlled pilot study in the use of telemedicine in the community on the management of heart failure. *Student Health Technology and Information*, 64, 126–137.
- Deyo, R. A., Inui, T. S., Leininger, J., et al. (1982). Physical and psychosocial function in rheumatoid arthritis: Clinical use of a self-administered health status instrument. *Archives of Internal Medicine*, 142, 879.
- Digitalis Investigation Group. (1997). The effect of digoxin on mortality and morbidity in patients with heart failure. *New England Journal of Medicine*, 336, 525–533.

- Dupuy, H. J. (1984). The Psychological General Well-Being (PGWB) Index. In N. K. Wenger, M. E. Mattson, C. D. Furburg, & J. Elinson (Eds.), *Assessment of quality of life in clinical trials of cardiovascular therapies* (170–183). New York, NY: Le Jacq.
- Edwards, D. W., Yarvis, R. M., Mueller, D. P., et al. (1978). Test-taking and the stability of adjustment scales: Can we assess patient deterioration? *Eval Q*, 2, 275–292.
- Ertle, D., & Litman, G. I. (2002). Hospital outpatient disease management initiative achieves marked reduction in inpatient admissions and costs for congestive heart failure. *Journal of Cardiac Failure*, 8, S90.
- Felker, G. M., O’Conner, C. M., & Ghali, J. K., (2001). Inotropic therapy for heart failure: An evidence-based approach. *American Heart Journal*, 142, 393–401.
- Fleg, J. L., & Lakatta, E. G. (1988). Role of muscle loss in the age-associated reduction in  $\dot{V}O_2$  max. *Journal of Applied Physiology*, 65, 1147–1151.
- Freedland, K. E., Carney, R., Rich, M. W., Caracciolo, A., Krotenberg, J. A., & Smith, L. J. (1991). Depression in elderly patients with congestive heart failure patients. *Journal of Geriatric Psychiatry*, 24(1), 59–71.
- GESICA Investigators (2005). Randomized trial of telephone intervention in chronic heart failure: DIAL trial. *British Medical Journal*, 331, 425.
- Gibbons, R. J., Balady, G. J., Beasley, J. W., Beasley, B. T, Wolf, F., Duvernoy. C., ... Winters, L., Jr. (1997). ACC/AHA guidelines for exercise testing: A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee on Exercise Testing). *Journal of the American College of Cardiology*, 30, 260–311.

- Gill, T. M., & Feinstein, A. R. (1994). A critical appraisal of the quality of quality-of-life measurements. *JAMA*, 272, 619–626.
- Gilson, B. S., Gilson, J. S., Bergner, M., Bobbit, R. A., Kressel, S., Pollard, W. E., & Vesselago, M. (1975). The sickness impact profile. Development of an outcome measure of health care. *American Journal of Public Health*, 65, 1304–1310.
- Goldberg, L. R., Piette, J. D., Walsh, M. N., Frank, T. A., & Jaski, B., (2002). A daily electronic home monitoring system in patients with advanced heart failure improves survival: The WHARF (weight monitoring in heart failure) trial. *Journal of Cardiac Failure*, 8, S54
- Gollob, M. H., & Seger, J. J. (2002). Current status of the implantable cardioverter-defibrillator. *Chest*, 119, 1210–1221.
- Goodman, S. N., & Berlin, J. A. (1994). The use of predicted confidence intervals when planning experiments and the misuse of power when interpreting results. *Annals of Internal Medicine*, 121, 200–206.
- Gorkin, L., Norvell, N. K., Rosen, R. C., Charles, E., Shumaker, S. A., McIntyre, K. M., ... Woods, P. (1993). Assessment of quality of life as observed from the baseline data of the studies of left ventricular dysfunction (SOL VD) trial of quality of life substudy. *American Journal of Cardiology*, 71, 1069–1073.
- Gott, M., Barnes, S., Parker, C., Payne, S., Seamark, D., Garibella, S., & Small, N. (2006). Predictors of the quality of life of older people with heart failure recruited from primary care. *Age and Ageing*, 35(2), 172–177.

- Green, P. C., Porter, C. B., Bresnahan, D. R., & Spertus, J. A. (2000). Development and evaluation of the Kansas City Cardiomyopathy Questionnaire: A new health status measure for heart failure. *Journal of the American College of Cardiology*, 35, 1245–1255.
- Guyatt, G. H. (1993). Measurement of health-related quality of life in heart failure. *Journal of the American College of Cardiology*, 22, 185A–191A.
- Guyatt, G. H., Bombardier, C., & Tugwell, P. X. (1986). Measuring disease-specific quality of life in clinical trials. *Canadian Medical Association Journal*, 134, 889–895.
- Guyatt, G. H., & Jaeschke, R. J. (1997). Reassessing quality of life instruments in the evaluation of new drugs. *Pharmacoeconomics*, 12, 616–626.
- Guyatt, G. H., Nogradi, S., Halcrow, S., Halcrow, S., Singer, J., Sullivan, M. J., & Fallen, E. L. (1989). Development and testing of a new measure of health status for clinical trials in heart failure. *Journal of General Internal Medicine*, 4(2), 101–107.
- Guyatt, G. H., Sullivan, M. J., Thompson, P. J., Fallen, E. L., Pugsley, S. O., Taylor, D. W., & Berman, L. B. (1985). The 6-minute walk: a new measure of exercise capacity in patients with chronic heart failure. *Canadian Medical Association Journal*, 132, 919–923.
- Guyatt, G. H., Thompson, P. J., Berman, L. B., Sullivan, M. J., Townsend, M., Jones, N. L., & Pugsley, S. O. (1985). How should we measure function in patients with chronic heart and lung disease? *Journal of Chronic Disease*, 38, 517–524.



- Guyatt, G. H., Townsend, M., Keller, J., Singer, J., & Nogradi, S. (1991). Measuring functional status in chronic lung disease: conclusions from a random control trial. *Respiratory Medicine*, 85(Suppl. B), 17–21.
- Hauck, W. W., Anderson, S., & Marcus, S. M. (1998). Should we adjust for covariates in nonlinear regression analyses of randomized trials. *Controlled Clinical Trials*, 19, 249–256.
- Havranek, E. P., Ware, M. G., & Lowes, B. D. (1999). Prevalence of depression in congestive heart failure. *American Journal of Cardiology*, 84, 348–350.
- Hays, R. D., & Woolley, J. M. (2000). The concept of clinically meaningful difference in health-related quality-of-life research: How meaningful is it? *Pharmacoeconomics*, 18, 419–423.
- Heidenrich, P. A., Ruggerio, C. M., & Massie, B. M. (1999). Effect of a home monitoring system on hospitalization and resource use for patients with heart failure. *American Heart Journal*, 138, 633–640.
- Heyting, A., Tolboom, J., & Essers, J. (1992). Statistical handling of drop-outs in longitudinal clinical trials. *Statistics in Medicine*, 11, 2043–2061.
- Hjalmarson, A., Goldstein, S., Fagerberg, B., Wedel, F. Waagstein, & J. Kjekshus. (2000). Effects of controlled-release metoprolol on total mortality, hospitalizations, and well-being in patients with heart failure: The Metoprolol CR/XL randomized intervention trial in congestive heart failure (MERIT-HF). *JAMA*, 283, 1295–1302.
- Hunt, S. A. (1998). Current status of cardiac transplantation. *JAMA*, 280, 1692–1698.

- Hunt, S. A., Baker, D. W., Chin, M. H., M. P. Cinquegrani, A. M. & Feldman. (2001).  
ACC/AHA guidelines for the management of chronic heart failure in the adult.  
*Journal of the American College of Cardiology*, 38, 2101–2113.
- Hunt, S. M., McKenna, S. P., McEwen, J., Backett, E. M., Williams, J., & Papp, E.  
(1980). A quantitative approach to perceived health status: A validation study.  
*Journal of Epidemiology and Community Health*, 34, 281–286.
- Husted, J. A., Cook, R. J., Farewell, V. T., & Gladman, D. D. (2000). Methods for  
assessing responsiveness: A critical review and recommendations. *Journal of  
Clinical Epidemiology*, 53, 459–468.
- International Conference on Harmonization of Technical Requirements for Registration  
of Pharmaceuticals for Human Use. (1999). ICH harmonized tripartite guideline.  
Statistical principles for clinical trials. *Statistics in Medicine*, 18, 1905–1942.
- Jaarsma, T., Halfens, R., Abu-Saad, H. H., Dracup, K., Stappers, J., & Van Ree, J.  
(1999). Quality of life in older patients with systolic and diastolic heart failure.  
*European Journal of Heart Failure*, 1(2), 151–160.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to  
defining meaningful change in psychotherapy research. *Journal of Consulting and  
Clinical Psychology*, 59, 12–19.
- Jaeschke, R., Singer, J., & Guyatt, G. H. (1989). Measurement of health status.  
Ascertaining the minimally important difference. *Controlled Clinical Trials*, 10,  
407–415.

- Jerant, A. F., Azari, R., & Nesbitt, T. S. (2001). Reducing the costs of frequent hospital admissions for congestive heart failure: A randomized trial of a home telecare intervention. *Med Care*, 39, 1234–1245.
- Jerome, L., Fleg, I. L., Pina, G. J., Balady, B. R., Chaitman, B. F., Lavie, C., ... Bazzarre, T. (2000). Assessment of functional capacity in clinical and research applications: An advisory from the Committee on Exercise, Rehabilitation, and Prevention, Council on Clinical Cardiology, American Heart Association. *Circulation*, 102, 1591–1597.
- Jiang, W., Alexander, J., Christopher, E., Kuchibhatla, M., Gaulden, L. H., Cuffe, M. S., ... O'Connor, C. M. (2001). Relationship of depression to increased risk of mortality and rehospitalization in patients with congestive heart failure. *Archives of Internal Medicine*, 161, 1849–1856.
- Johnson, J. A., Parker, R. B., & Patterson, J. H. (2002). Heart failure. In J. T. DiPiro, R. L. Talbert, & G. C. Yee, (Eds), *Pharmacotherapy: A pathophysiologic approach* (5th ed.; pp. 185–218). New York, NY: McGraw-Hill.
- Johnson, P. A., Goldman, L., Orav, E. J., Garcia, T., Pearson, S. D., & Lee, T. H. (1995). Comparison of the medical outcomes study short-form 36-item health survey in Black patients and White patients with acute chest pain. *Med Care*, 33, 145–160.
- Johnson, B., Wheeler, L., Deuser, J., & Sousa, K. H. (2000). Outcomes of the Kaiser Permanente tele home health research project. *Archives of Family Medicine*, 2000, 9(1), 40–45.

- Juniper, E. F., Guyatt, G. H., Willan, A., & Griffith, L. (1994). Determining a minimal important change in a disease specific quality of life questionnaire. *Journal of Clinical Epidemiology*, 47, 81–87.
- Kadish, A., Dyer, A., Daubert, J. P., Quigg, R., Estes, N. A., & Anderson, K. P. (2004). Prophylactic defibrillator implantation in patients with nonischemic dilated cardiomyopathy. *New England Journal of Medicine*, 350, 2151–2158.
- Kannel, W. B., & Belanger, A. J. (1991). Epidemiology of heart failure. *American Heart Journal*, 121, 951–957.
- Kaplan, R. M., & Anderson, J. P. (1988). The quality of well-being scale: Rationale for a single quality of life index. In S. R. Walker (Ed.), (51–77). London, England: MTP Press.
- Katschnig, H. (1997). How useful is the concept of quality of life in psychiatry? In H. Katschnig, H. Freeman, & N. Sartorius (Eds.), *Quality of life in mental disorders* (pp. 3–16). Chichester: Wiley.
- Kazis, L., Anderson, J., & Meenan, R. (1989). Effect sizes for interpreting changes in health status. *Med Care*, 27, S178–S189.
- Kesinger, T., Gilani, S., & Jennison, S. H. (2002). Electronic home monitoring reduces hospital admission, length of stay and readmission frequency in a selected heart failure population. *Journal of Cardiac Failure*, 8, S94.
- Knox, D. A., Mueller, T. M., Vuckovic, K. M., & Acker, K. (2002). Remote titration of beta blocker therapy for heart failure by advanced practice nurses, titration protocols and daily patient telemanagement. *Journal of Cardiac Failure*, 8, S83.

- Kolotkin, R. L., Crosby, R. D., & Williams, G. R. (2002). Integrating anchor-based and distribution-based methods to determine clinically meaningful change in obesity-specific quality of life. *Quality of Life Research*, 11, 670.
- Lapworth, D. J., & Dibiase, A. (2000). Decreased hospitalizations using a home-based electronic weight monitoring system as part of a comprehensive heart failure program. *Journal of Cardiac Failure*, 6, 69.
- Laramée, A. S., & Levinsky, S. K. (2003). Case management in a heterogeneous congestive heart failure population. *Archives of Internal Medicine*, 163, 809–17.
- Leidy, N. K. (1994). Using functional status to assess treatment outcomes. *Chest*, 106, 1645–1646.
- Lipkin, D. P., Scriven, A. J., & Poole-Wilson, P. A. (1986). Six-minute walk test for assessing exercise capacity in chronic heart failure. *British Medical Journal*, 292, 653–655.
- Llewellyn-Thomas, H. A., Williams, J. I., Levy, L., & Naylor, C. D. (1996). Using trade-off techniques to assess patients' treatment preferences for benign prostatic hyperplasia. *Medical Decision Making*, 16, 262–272.
- Louis, A. A., Turner, T., Gretton, M., Baksh, A., Mortel, J. G.F. H., Sack, S., ... Husser, D. (2003). A systematic review of telemonitoring for the management of heart failure. *European Journal of Heart Failure*, 5, 583–590.
- Macropoulos, L. R., & Selna, M. J. (2002). CHF hospital admissions reduced by 48% (to 0.272 PPPY at 360 days) in large Medicare population using advanced home monitoring program. *Journal Cardiac Failure*, 8, S97.

Majania, G., Pierobon, A., Giardini, A., Callegaria, S., Opasichb, C., Cobellic, F., ...

Fisher, M. L. (1999). Relationship between psychological profile and  
cardiological variables in chronic heart failure: The role of patient subjectivity.  
*European Heart Journal*, 20, 1579–1586.

Massie, B. M., & Shah, N. B. (1997). Evolving trends in the epidemiologic factors of  
heart failure: Rationale for preventive strategies and comprehensive disease  
management. *American Heart Journal*, 133, 703–712.

Massie, B. M., West, J., Van Ostaeyen, D., & Salbalvaro, A. (2001). A controlled trial of  
heart failure management programs. *Journal of the American College of  
Cardiology*, 37, 1A–648A.

McAlister, F. A., Stewart, S., Ferrua, J., & McMurray, J. J. V. (2004). Multidisciplinary  
strategies for the management of heart failure patients at high risk for admission:  
A systematic review of randomized trials. *Journal of the American College of  
Cardiology*, 44, 810–819.

McHorney, C. A., & Tarlov, A. R. (1995). Individual patient monitoring in clinical  
practice: Are available health status surveys adequate? *Quality of Life Research*,  
4, 293.

Miyamoto, S., Nagaya, N., Satoh, T., Kyotani, S., Sakamaki, F., Fujita, M., ... Miyatake,  
K. (2000). Clinical correlates and prognostic significance of six-minute walk test  
in patients with primary pulmonary hypertension. *American Journal of  
Respiratory and Critical Care Medicine*, 161, 487–492.

- Morales, F. J., Martinez, A., Mendez, M., Agarrado, A., Ortega, F., FernandezGuerra, J., ... Burgos, J. (1999). A shuttle walk test assessment of functional capacity in chronic heart failure. *American Heart Journal*, 138, 292–298.
- Moss, A. J., Zareba, W., Hall, W. J., Klein, H., & Wilber, D. (2002). Prophylactic implantation of a defibrillator in patients with myocardial infarction and reduced ejection fraction. *New England Journal of Medicine*, 346, 877–883.
- Mossey, J. M., & Shapiro, E. (1982). Self-rated health: A predictor of mortality among the elderly. *American Journal of Public Health*, 72, 800–808.
- National Heart, Lung, and Blood Institute. (2004). *Facts about heart failure*. Retrieved from <http://www.nhlbi.nih.gov/health/public/heart/other/hrtfail.htm>
- Ni, H., Toy, W., Burgess, D., & Wise, K. (2000). Comparative responsiveness of Short-Form 12 and Minnesota Living with Heart Failure Questionnaire in patients with heart failure. *Journal of Cardiac Failure*, 6, 83–91.
- O’Connell, J. B., & Bristow, M. R. (1994). Economic impact of heart failure in the United States: Time for a different approach. *Journal of Heart Lung Transplant*, 13, S107–S112.
- O’Keffe, S. T., Lye, M., Donnellan, C., & Carmichael, D. N. (1998). Reproducibility and responsiveness of quality of life assessment and six minute walk test in elderly heart failure patients. *Heart*, 80, 377–382.
- Osoba, D., Rodrigues, G., Myles, J., Zee, B., & Pater, J. (1998). Interpreting the significance of changes in health-related quality-of-life scores. *Journal of Clinical Oncology*, 16, 139-144.

- Packer, M., Bristow, M. R., Cohn, J. N., Colucci, W. S., Fowler, M. B., Edward, M.,....Shursterman, N. H. (1996). The effect of carvedilol on morbidity and mortality in patients with chronic heart failure. *New England Journal of Medicine*, 334, 1349–1355.
- Packer, M., Cohn, J. N., Abraham, W. T., et al. (1999). Consensus recommendations for the management of chronic heart failure. On behalf of the membership of the advisory council to improve outcomes nationwide in heart failure. *American Journal of Cardiology*, 83(2A), 1A–38A.
- Packer, M., Gheorghiade, M., Young, J. B., Constantini, D. O., Adams, K. F., Cody, R.J., ... Jolly, M. K. (1993). Withdrawal of digoxin from patients with chronic heart failure treated with angiotensin-converting-enzyme inhibitors. *New England Journal of Medicine*, 329, 1–7.
- Paul, L. E. (2003). The six-minute walk test. *Respiratory Care*, 48, 783–785.
- Perera, S., Mody, S. H., Woodman, R. C., & Studenski, S. A. (2006). Meaningful change and responsiveness in common physical performance measures in older adults. *Journal of the American Geriatric Society*, 54, 743–749.
- Pina, I. L., Apstein, C. S., Balady, G. J., Belardinelli, R., Chaitman, B.R., Duscha, B.D., ... Sullivan, M. J. (2003). Exercise and heart failure: A statement from the American Heart Association Committee on Exercise, Rehabilitation, and Prevention. *Circulation*, 107, 1210–1225.
- Pitt, B., & Cohn, J. N., (1991). Effect of enalapril on survival in patients with reduced left ventricular ejection fractions and congestive heart failure (SOLVD). *New England Journal of Medicine*, 325, 293–302.



- Pitt, B., Zannad, F., & Remme, W. J., (1999). The effect of spironolactone on morbidity and mortality in patients with severe heart failure. *New England Journal of Medicine*, 341, 709–717.
- Pollock, M. L., Gaesser, G. A., Butcher, J. D., Despres, J., Dishman, R. K., & Franklin, B. A. (1998). American College of Sports Medicine position stand: The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness and flexibility in healthy adults. *Medicine & Science in Sports & Exercise*, 30, 975–991.
- Poole-Wilson, P. A. (2000). Six minute walk predicts prognosis in patients with heart failure. *European Heart Journal*, 21, 507–508.
- Rector, T. S. (2005). *Overview of the Minnesota Living with Heart Failure Questionnaire*. Retrieved from [http://www.mlhfq.org/\\_dnld/mlhfq\\_overview](http://www.mlhfq.org/_dnld/mlhfq_overview)  
1/1/2005
- Rector, T. S., & Cohn, J. N. (1992). Assessment of patient outcome with the Minnesota Living with Heart Failure Questionnaire: Reliability and validity during a randomized double-blind, placebo-controlled trial of pimobendan. *American Heart Journal*, 124, 1017–1025.
- Rector, T. S., Kubo, S. H., & Cohn, J. N. (1992). Patients' self-assessment of their congestive heart failure: Part 2: Content reliability and validity of a new measure, The Minnesota Living with Heart Failure Questionnaire. *Heart Failure*, 3, 198–209.

- Rector, T. S., Kubo, S. H., & Cohn, J. N. (1993). Validity of the Minnesota Living with Heart Failure Questionnaire as a measure of therapeutic response to enalapril or placebo. *American Journal of Cardiology*, 71, 1106–1107.
- Rector, T. S., Tschumperlin, L. K., Kubo, S. H., Bank, A. J., Francis, G. S., McDonald, K. M., ... Silver, M. A. (1995). Use of the Living with Heart Failure Questionnaire to ascertain patients' perspectives on improvement in quality of life versus risk of drug-induced death. *Journal of Cardiac Failure*, 1, 201–206.
- Redelmeier, D. A., Bayoumi, A. M., Goldstein, R. S., & Guyatt, G. H. (1997). Interpreting small differences in functional status: The six-minute walk test in chronic lung disease patients. *American Journal of Respiratory and Critical Care Medicine*, 155, 1278–1282.
- Riegel, B., Carlson, B., Kopp, Z., LePetri, B., Glaser, D., & Unger, A. (2002) Effect of a standardized nurse case management telephone intervention on resource use in patients with chronic heart failure. *Archives of Internal Medicine*, 162, 705–12.
- Riegel, B., Carlson, B., Glaser, D., & Romero, T. (2006). Randomized controlled trial of telephone case management in Hispanics of Mexican origin with heart failure. *Journal of Cardiac Failure*, 12, 211–19.
- Revicki, D. A., Osoba, D., Fairclough, D., Barofsky, I., Berzon, R., Leidy, N. K., & Rothman, M. (2000). Recommendations on health-related quality of life research to support labeling and promotional claims in the United States. *Quality of Life Research*, 9, 887–900.

- Roglieri, J. L., Futterman, R., McDonough, K. L., Malya, G., Karwath, K. R., Bowman, D., ... Warburton, S.W., Jr. (1997). Disease management interventions to improve outcomes in congestive heart failure. *American Journal of Managed Care*, 3, 1831–1839.
- Rose, E. A., Gelijns, A. C., Moskowitz, A. J., Heitjan, D. F., Stevenson, L. W., Meier P., & Dembitsky, W, (2001). Long-term use of a left ventricular assist device for end-stage heart failure. *New England Journal of Medicine*, 345, 1435–1443.
- Roul, G., Germain, P., & Bareiss, P. (1998). Does the 6-minute walk test predict the prognosis in patients with NYHA Class II or III chronic heart failure? *American Heart Journal*, 136, 449–457,
- Rowell, L. B. (1988). Muscle blood flow in humans: How high can it go? *Medicine & Science in Sports & Exercise*, 20, S97–S103.
- Rui, C., Sonia, R., Joana, P., Paulo, B., Antonio, F., & Mario, C. (2005). Heart failure and health-related quality of life. *Clinical Practice and Epidemiology in Mental Health*, 1(19), 1–7.
- Samsa, G., Edelman, D., Rothman, M. L., Williams, G. R., Lipscomb, J., & Matchar, D. (1999). Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. *Pharmacoeconomics*, 15, 141–155.
- Scalvini, S., Zanelli, E., Volterrani, M., & Benigno, M. (2002). Effect of a home-based telecardiology on chronic heart failure outcomes: A case control pilot study. *European Journal of Heart Failure*, 1, 72.

- Senn, S. (1994). Testing for baseline balance in clinical trials. *Statistics in Medicine*, 13, 1715–1726.
- Shah, N. B., Der, E., Ruggerio, C., Heidenrich, P. A., & Massie, B. M. (1998). Prevention of hospitalizations for heart failure with an interactive home monitoring program. *American Heart Journal*, 135, 373–378.
- Skotzko, C. E., Krichten, C., Zietowski, G., et al. Depression is common and precludes accurate assessment of functional status in elderly patients with congestive heart failure. *Journal of Cardiac Failure*, 2000, 300–305.
- Smith, A. H., & Bates, M. N. (1992). Confidence limit analyses should replace power calculations in the interpretation of epidemiologic studies. *Epidemiology*, 3, 449–452.
- Solway, S., Brooks, D., Lacasse, Y., & Thomas, S. (2001). A qualitative systematic overview of the measurement properties of functional walk tests used in the cardiorespiratory domain. *Chest*, 119, 256–270.
- Spertus, J. A., Winder, J. A., Dewhurst, T. A., Deyo, R. A., & Stephan, D. F. (1994). Monitoring the quality of life in patients with coronary artery disease. *American Journal of Cardiology*, 74, 1240–1244.
- Speer, D. C. (1992). Clinical significant change: Jacobson and Truax (1991) revisited. *Journal of Consulting and Clinical Psychology*, 60, 402–408.
- Speer, D. C., & Greenbaum, P. D. (1995). Five methods for computing significant individual client change and improvement rates: support for an individual growth curve approach. *Journal of Consulting and Clinical Psychology*, 63, 1044–1048.

- Steptoe, A., Mohabir, A., Mahon, N. G., & McKenna, W. J. (2000). Health-related quality of life and psychological well being in patients with dilated cardiomyopathy. *Heart*; 83, 645–650.
- Stevens, D., Elpern, E., Sharma, K., Szidon, P., Ankin, M., & Ketsen, S. (1999). Comparison of hallway and thread-mill six-minute walk tests. *American Journal of Respiratory and Critical Care Medicine*, 160, 1540–543.
- Stewart, A. L., Greenfield, S., Hays, R. D., Wells, K., Rogers, W. H., Berry, S. D., ... Ware, J. E., Jr. (1989). Functional status and well-being of patients with chronic conditions. Results from the Medical Outcomes Study. *JAMA*, 262, 907–913.
- Stott, D. J. (2002). Chronic heart failure and cognitive impairment: Co-existence of conditions or true association? *European Journal of Heart Failure*, 4, 7–9.
- Stucki, G., Liang, M. H., & Fossel, A. H., (1995). Relative responsiveness of condition specific and health status measures in degenerative lumbar spinal stenosis. *Journal of Clinical Epidemiology*, 48, 1369–1378.
- Testa, M. (1987). Interpreting quality of life clinical trial data for use in the clinical practice of antihypertensive therapy. *Journal of Hypertension*, 5, S9–S13.
- Testa, M. A., & Lenderking, W. R. (1992). Interpreting pharmacoeconomic and quality-of-life clinical trial data for use in therapeutics. *Pharmacoeconomics*, 2, 107.
- Testa, M. A., & Simonson, D. C. (1996). Assessment of quality-of-life outcomes. *New England Journal of Medicine*, 28, 835–840.
- Tsuyuki, R. T., Fradette, M., & Johnson, J.A. (2004). Multicentre disease management program for hospitalized patients with heart failure: the Review of Education on

- ACE Inhibitors in Congestive Heart Failure Treatment (REACT) Study. *Journal of Cardiac Failure*, 10, 473–80.
- Vaccarino, V., Kasl, S. V., Abramson, J., & Krumholz, H. M. (2001). Depressive symptoms and risk of functional decline and death in patients with heart failure. *Journal of the American College of Cardiology*, 38, 199–205.
- Wang, L., Yu, C.-M., Chau, E., & Lamm W.-F. (2002). Feasibility of predicting CHF hospitalization using pacemaker-based impedance sensor in CHF patients. *Journal of Cardiac Failure*, 8, S81.
- Ware, J. E., Snow, K., Kosinski, M., & Gandek, B. (1993). *SF-36 Health Survey: Manual and interpretation guide*. Boston, MA: The Health Institute.
- Wasserman, K., Hansen, J. E., Sue, D. Y., Casaburi, R., & Whipp, B. J. (1999). *Principles of exercise testing and interpretation* (3rd ed.). Philadelphia, PA: Lippincott, Williams & Wilkins.
- Weisman, I. M., & Zeballos, R. J. (1994). An integrated approach to the interpretation of cardiopulmonary exercise testing. *Clinics in Chest Medicine*, 15, 421–445.
- Wenger, N. K. (1989). Quality of life: Can it and should it be assessed in patients with heart failure? *Cardiology*, 76, 391–398.
- Wielenga, R. P., Erdman, R. A. M., Huisveld, I. A., Bol, E., Dunselman, P. H., Baselier, M. R., & Mosterd, W. L. (1997). Effect of exercise training on quality of life in patients with chronic heart failure. *Journal of Psychosomatic Research*, 45, 459–464.

- Williams, R. E., Keiler, L., Sprang, M., & Mehan, C. (1998, Month). *Telemanagement of congestive heart failure: Results of daily weight and symptom tracking*. Paper presented at the American College of Cardiology Annual Scientific Session, Location, Location.
- Woodend, A. K., Fraser, M., Sherrard, H., & Stueve, L. (2002). Readmission and quality of life: The impact of telehome care in heart failure. *Journal of Cardiac Failure*, 8, S98.
- World Health Organization. (1946). *Constitution of the World Health Organization*. Geneva, Switzerland: Author.
- Woodend, A.K., Sherrard, H., Fraser, M., Stuewe, L., Cheung, T., & Struthers, C. (2008). Telehome monitoring in patients with cardiac diseases who are at high risk of readmission. *Heart & Lung*, 37, 36–45.
- Wyrwich, K. W., Nienaber, N. A., Tierney, W. M., & Wolinsky, F. D. (1999). Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. *Med Care*, 37, 469–478.
- Wyrwich, K. W., Tierney, W. M., & Wolinsky, F. D. (1999). Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. *Journal of Clinical Epidemiology*, 52, 861–873.
- Yusuf, S., Pfeffer, M. A., Swedberg, K., Granger, C. B., Held, P., McMurray, J. J., ... Ostergren, J. (2003). Effects of candesartan in patients with chronic heart failure and preserved left-ventricular ejection fraction: The CHARM-Preserved Trial. *Lancet*, 362, 777–781.

- Yusuf, S., Sleight, P., Pogue, J., Bosch, J., Davies, R., & Dagenais, G. (2000). Effects of an angiotensin-converting-enzyme inhibitor, ramipril, on cardiovascular events in high-risk patients. *New England Journal of Medicine*, 342, 145–153.
- Zugck, C., Kruger, C., Durr, S., et al. (2000). Is the 6-minute walk test a reliable substitute for peak oxygen uptake in patients with dilated cardiomyopathy? *European Heart Journal*, 21, 540–549.



## APPENDIX

### GLOSSARY<sup>1</sup>

**ACE-I:** Angiotensin converting enzyme inhibitors, a class of medications that inhibit angiotensin converting enzyme

**Angina:** Angina is chest pain that is due to an inadequate supply of oxygen to the heart muscle

**Beta-blockers:** A class of medications that block beta-adrenergic substances such as adrenaline (epinephrine), a key agent in the “sympathetic” portion of the autonomic nervous system and activation of the heart muscle

**Blood Pressure:** Pressure of the blood within the arteries

**BUN:** Blood urea nitrogen, measures primarily the urea level in blood

**Clinic visit:** Refers to outpatient care at physician’s office

**Cohort:** In a clinical trial, a group of study subjects or patients

**Congestive Heart Failure:** A condition in which the heart’s function as a pump is inadequate to meet the body’s needs

**Coronary Angiography:** Accurate method for evaluating and defining coronary artery disease (CAD)

---

1. Adapted from HCFA Heart Failure Demonstration Project protocol, pp. 71–77.

**Creatinine:** Creatinine is produced from creatine, a molecule of major importance for energy production in muscles

**Death due to Non-cardiovascular causes:** Death not due to any cardiac or vascular events

**Death due to Progressive Heart Failure:** Death preceded by worsening signs and symptoms of heart failure and/or circulatory collapse due to pump failure in the absence of acute myocardial infarction

**Death due to other Cardiac Causes including Vascular Causes:** Death due to other cardiac causes, such as non-sudden arrhythmic death, due to vascular events such as stroke, pulmonary embolus, ruptured aneurysm, etc.

**Diastolic Dysfunction:** Abnormal relaxation pattern of the left ventricle

**Echocardiography:** A diagnostic test, which uses ultrasound waves to make images of heart muscles, valves and other internal structures. It measures the function of heart and its structures

**Ejection Fraction:** The proportion of blood that is pumped out of a filled ventricle with each heart beat

**Emergency room visit (ER):** Refers to outpatient care at emergency room, not considered as hospitalization

**Fatal Myocardial Infarction:** Death as a result of an autopsy – verified myocardial infarction or death before hospital discharge from a hospital – verified acute myocardial infarction

**HCFA:** The health care finance administration, the part of the US Department of Health and Human Services (HHS) that is responsible for administering Medicare and Medicaid

**Hospitalization:** Refers to inpatient admission to hospital, either overnight stay or prolonged stay

**Functional Capacity:** Functional capacity refers to the performance of, or the capacity to perform, regular daily activities

**Morbidity:** Illness or disease

**Mortality:** It refers to a fatal outcome, death

**Quality of Life:** A state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity

**Six Minute Walk Distance:** Distance that a patient can quickly walk on a flat, hard surface in a period of 6 minutes

**Six Minute Work:** The product of 6-minute walk distance multiplied by the weight of the patient: 6-min walk distance (in meters) x body weight (in kilograms)

**Sudden Cardiac Death:** Death, or irreversible deterioration leading to death, occurring without warning or within one hour of symptoms (unless it is the result of circulatory collapse due to pump failure or fatal myocardial infarction or other well defines cardiac, vascular or non-cardiac causes. Sudden cardiac death will be further categorized as due to brady-arrhythmia, ventricular tachycardia, ventricular fibrillation, electromechanical dissociation, or unknown cause of death

**Systolic Dysfunction:** Impairment of forward pumping function of the heart

## CURRICULUM VITAE

Venkata S Chilakapati MD., MPH.  
5485, Michelles Oak Ct, Unit K  
Cincinnati, OH 45248  
Email: chilakapati2005@yahoo.com

- Level -II certification in ECHO Cardiography, Nuclear Cardiology, Cardiac CT angiography and Cardiac catheterization
- Board eligible for 2010 cardiology boards: General Cardiology, ECHO Cardiography, Nuclear Cardiology and Cardiac CT angiography boards

## EDUCATION

08 / 2008 – 06 / 2010

Cardiology Fellowship, Chief Fellow (2009-2010), University of Cincinnati, OH

07 / 2007 - 07 / 2008

Cardiology Fellowship (1<sup>st</sup> yr), Michael Reese Hospital, Chicago, IL, USA

06 / 2004 - 06 / 2007

Internal Medicine Residency, Michael Reese Hospital, Chicago, IL, USA

05 / 2002 – 05/2010

PhD Degree, Epidemiology and Clinical Investigation Sciences, University of  
Louisville, Louisville, KY

8 / 2001-05 / 2003

MSPH Degree, Epidemiology and Clinical Investigation Sciences, University of  
Louisville, Louisville, KY

1 / 2000 - 05 / 2001

MPH Degree, Master of Public Health, Western Kentucky University, KY, USA.

11 / 1993 - 08 / 1999

Medical Degree: Andhra Medical College, Andhrapradesh, India.

#### CERTIFICATION/LICENSURE(S)

ABIM certification

ECFMG certification

BLS and ACLS certification

Permanent Licensure, IL, USA

Controlled Substance Licensure, IL, USA

Permanent Licensure, OH, USA

Permanent Licensure, IN, USA

Human Subjects Research CITI training certification

Good Clinical Practice (GCP) CITI training certification

Authorized User of Radioactivity Education and Training

#### PROFESSIONAL MEMBERSHIPS

American Heart Association, USA (06/2007 - present)

American College of Cardiology (06/2007- present)

American Society for Nuclear Cardiology (09/2009 – present)

American Society for ECHO Cardiography (09/2009 – present)

American Medical Association, USA (12/2006 - present)

Golden Key Honor Society, USA (02/2007 - present)

Phi-Kappa-Phi Honor Society, USA (April, 2004 - present)

American Public Health Association, USA (May, 2001 - present)

Andhrapradesh Medical Council, India (June, 1998 - present)

#### AWARDS FROM MICHAEL REESE HOSPITAL, CHICAGO, IL

Robert A Weinstein Award: Resident Teacher of the Year, 2006 – 2007

#### PRIZES FROM ANDHRA MEDICAL SCHOOL, VISAKHAPATNAM, INDIA

First prizes for seminar presentations:

on HIV/AIDS: from Dept of Microbiology (1996)

on Hepatitis : from Dept of Medicine (1998)

on Intestinal Obstruction: from Dept of surgery (1998)

on child health in India: from Dept of Pediatrics (1998)

Quiz: Second prize in Microbiology (1996)

Best speaker award from Indian Medical Association, VSP chapter for seminar presentation on Voluntary Blood Donation (1998)

#### AWARDS FROM WESTERN KENTUCKY UNIVERSITY, BOWLING GREEN, KY

WKU scholarship/ Dean's list: GPA 4.0/4.0 (2000-2001)

John D. Minton's Best Outgoing Student Award from Graduate School, WKU  
2001

Outstanding Graduate Student award from Dept. of Public Health, WKU 2001

Best International Student Award from Office of International programs, WKU  
2001

Annual Globe (extra-curricular) Award from International center, WKU 2001

#### AWARDS FROM UNIVERSITY OF LOUISVILLE, LOUISVILLE, KY

Dean's Citation Award, 2010

University Fellowship award, 2001-2004

Travel Grant from American Federation for Medical Research (AFMR), 2004

#### STATISTICAL PACKAGE PROFICIENCY

SPSS, SAS, STAT-exact, Micro Array Analysis

#### ON-GOING RESEARCH ACTIVITIES IN CARDIOLOGY

Effect of Technology Driven Heart failure monitoring system on Functional capacity and Quality of Life in chronic heart failure patients (A multi-centered, randomized, controlled clinical trial).

Sub-Investigator: at the University of Cincinnati, Cincinnati

RADAR trial – Phase II multicenter clinical trial

ASCEND trial – Phase III multicenter clinical trial

TOPCAP trial – Phase II single center clinical trial

#### ABSTRACT/POSTER PRESENTATION(S)

Chilakapati V, Laber D et al, Phase II study - evaluation of safety and efficacy of regimen cyclophosphamide, etoposide and estramustine in treating androgen independent prostate cancer, Abstract- American Federation for Medical Research (AFMR: Clinical Research -2004 conf.).

#### PUBLICATIONS

Cilingiroglu, M., Chilakapati, V., Leesar, M., & Dunlap, S. (2009). Sirolimus-eluting stent fractures leading to restenosis in cardiac allograft vasculopathy, *Journal of Cardiovascular Revascularization Medicine*, 11(2), 120-123.

Choudry, M.A., Chilakapati, V., & Crump C. (2004) Phase II study of cyclophosphamide, etoposide and estramustine (CEE) in patients with androgen

independent prostate cancer (AIPC), *Journal of Clinical Oncology*, ASCO Annual Meeting Proceedings (Post-Meeting Edition), 22, No 14S (July 15 Supp.) 4708.

Chilakapati, V., Duncan, D., Nicholson, T., White, J., & Lindley, L. (2003). A case-control analysis of felony convictions among recreational drug users, *Journal of Psychological Reports*, 93, 365 - 370.

#### PROFESSIONAL EXPERIENCE

06/2008 - 06/2010

Cardiology Fellowship, Dept of Cardiology, University of Cincinnati, OH

06/2007 – 06/2008

Cardiology Fellowship (1<sup>st</sup> year), Dept of Cardiology, Michael Reese Hospital, Chicago, IL

06/2004 - 06/2007

Internal Medicine Residency, Department of Medicine, Michael Reese Hospital, Chicago, IL

05/2003 - 06/2004

Teaching associate, Dept. of Epidemiology and Clinical Investigation Sciences, University of Louisville, Louisville, KY

01 / 2001 - 05 / 2003

Research Associate, Dept. of Epidemiology and Clinical Investigation Sciences, University of Louisville, Louisville, KY

01 / 2000 - 05 / 2001

Graduate Assistant (teaching/research), School of Public Health, Western Kentucky University, Bowling Green, KY



09 / 1999 - 12 / 2000

Junior Medical Officer, CDR Hospitals, Visakhapatnam, Andhrapradesh, India

09 / 1999 - 12 / 2000

Medicine Internship, King George Hospital, Visakhapatnam, Andhrapradesh,  
India

#### INSTITUTIONAL LEADERSHIP ACTIVITIES

04 / 2009 - 04 / 2010

Chief Fellow, Division of Cardiovascular Diseases, University of Cincinnati, OH

07 / 2006 - 06 / 2007

Residency Recruitment Committee Member, Dept. of Medicine, Michael Reese  
Hospital, Chicago

07 / 2005 - 06 / 2007

Resident Council Member, Dept. of Medicine, Michael Reese Hospital, Chicago

01 / 2003 - 05 / 2004

Advisor, Student-Affairs Committee, Indian Students Association, University of  
Louisville, Louisville, KY

09 / 2000 - 08 / 2001

Graduate Council Member, Western Kentucky University, Bowling Green, KY

09 / 2000 - 08 / 2001

Graduate Students Representative, Western Kentucky University, Bowling Green,  
KY

09 / 2000 - 08 / 2001

Graduate Programs task group member, Western Kentucky University, Bowling  
Green, KY

01 / 2000 - 08 / 2001

International Student Mentor, Western Kentucky University, Bowling Green, KY

#### HOBBIES/INTERESTS

06 / 2003 – 05/2004

Wrote a column for Indian Students Association monthly news letter, University  
of Louisville (Column title: Courage to Confront a Killer: Doctors Without  
Borders)

05/1996-10/1999

Television Host and Anchor, City Channel Network, Andhrapradesh, India

Playing Shuttle, Badminton and Chess

Listening music, watching movies and drawing designs