PREDICTORS OF HEALTH SERVICE USE IN PERSONS WITH HEART FAILURE

by

MARY ANN C. LAWLOR

Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Frances Payne Bolton School of Nursing

CASE WESTERN RESERVE UNIVERSITY

MAY, 2021

CASE WESTERN RESERVE UNIVERSITY

SCHOOL OF GRADUATE STUDIES

We hereby approve the thesis/dissertation of

Mary Ann C. Lawlor

candidate for the degree of Doctor of Philosophy

Committee Chair

Mary Dolansky, PhD, RN, FAAN

Committee Member

Nicholas Schiltz, PhD

Committee Member

Carolyn Harmon-Still, PhD, CMP, MSM, AGPCNP-BC, CRRP

Committee Member

Eiran Gorodeski, MD, MPH

Date of Defense: April 12, 2021

*We also certify that written approval has been obtained for any proprietary material contained therein.

Table of Contents

List of Tables

List of Figures

List of Figures Continued

Acknowledgements

I would first like to acknowledge the Robert Wood Johnson Foundation for the great honor in supporting my academic endeavors by funding my scholarship through their Robert Wood Johnson Future of Nursing Scholars program. I would further like to recognize the Frances Payne Bolton School of Nursing, as well as the members of my professional committee for their guidance and expertise. I would lastly like to thank my family for their continued love and support, especially my husband Sean, daughter Reagan, and my mother Mary Ann.

Predictors of Health Service Use in Persons with Heart Failure

Abstract

by

MARY ANN C. LAWLOR

Heart failure is a growing epidemic with more than 26 million individuals affected worldwide and an estimated 800,000 new cases diagnosed annually. Heart failure accounts for a disproportionately high amount of health service use, including more than 2% of all national hospitalizations. Increasing rates of heart failure diagnoses coupled with the rapidly aging population result in frequent and unnecessary use of health services such as visits to the emergency department, hospitalizations, and 30-day readmissions. The purpose of the study was to (1) identify the strength of the relationship between the community-based socioeconomic predisposing, enabling, perceived need, and environmental predictors and outcomes of health service use (2) build a model to predict health service use in the heart failure population, and (3) determine the time to event and relative risk of health service use for persons with heart failure. The longitudinal retrospective study used the Medicare Expenditure Panel Survey (MEPS) database (N=1714). Variables include age, gender, race, income, education, employment, marital status, insurance coverage, body mass index, smoking status, comorbidity, physical functioning, mental health, and hospital length of stay. Data were analyzed using logistic regression and Cox Proportional Hazards (time to event) analysis. Predictive models for each health service use outcome were built using the SPSS Modeler package. Risk factors and predictors of hospitalizations and emergency department visits include persons over the age of 65 who have self-reported poor physical functioning, and baseline

asthma/COPD, kidney, and cardiovascular disease. Risk factors and predictors of 30-day hospital readmissions include persons of Asian descent who were divorced, had seven or more comorbid conditions at baseline, and experienced a prior hospital length of stay longer than one week. Individuals experienced hospitalizations (M=5.5 months) and emergency department visits (M=5.3 months) at similar rates after study enrollment. The predictive models for hospitalizations and emergency department visits performed poorly (average AUC 0.589), however the 30-day hospital readmission models performed slightly better (average AUC 0.704). Many of the risk factors identified in the study are modifiable. Further research is necessary to examine implications of socioeconomic factors and their impact on health service use in a community-based heart failure population.

Chapter I

Introduction

Problem and Purpose

Heart failure (HF) is a dynamic disease process that requires a multifaceted and individualized approach to care. As a result of the challenges of disease trajectory and treatment complexity, persons with HF often experience frequent use of health services such as emergency department visits, hospitalizations, and 30-day readmissions (Ambrosy et al., 2004). Recent research has attempted to identify factors and build models that predict health service use in persons with HF, but the results of these studies remain inconsistent, and existing predictive models lack generalizability and reproducibility (Chamberlin et al., 2018 & Eapen et al., 2015). Furthermore, much of the current HF literature focuses on 30-day readmission risks for acute patients, rather than persons living in the community. Consequently, the rates of health service use in the national HF population continue to rise (Aranda et al., 2007), leading to poor patient outcomes. To address this problem, the purpose of the current study was to (1) identify factors that are most predictive of health service use (i.e., emergency department visits, hospitalizations, and 30-day readmissions) in persons with HF, (2) create a model that predicts health service use in persons with HF and (3) determine the time to event and relative risk of health service use for persons with HF.

Background

Heart Failure

Heart failure affects approximately 26 million individuals worldwide, almost six million of whom live in the United States alone (Ambrosy et al., 2014), and continues to be one of the most rapidly growing medical diagnoses with 800,000 new cases identified worldwide each year (Holley, Harvey, & John, 2014). Heart Failure can be an acute or chronic state of cardiac disease characterized by the clinical symptoms of dyspnea, fatigue, and fluid retention. The notable symptoms of HF occur because of a weakening of the cardiac muscle, increased cardiac filling pressures secondary to pulmonary hypertension, or a combination of the two etiologies (Hosenpud and Greenburg, 2000). Causes of HF can be the result of genetic predisposition, unhealthy lifestyle, postpartum complications, myocardial infarction, congenital abnormalities, substance abuse, hypertension, and a multitude of additional factors (American Heart Association [AHA], 2017). According to the New York Heart Association (NYHA), HF is classified by degree of severity, with some individuals experiencing no symptoms (Class I) and others experiencing extreme shortness of breath at rest (Class IV) (Raphael et al., 2006). More than half of persons with HF are over the age of 75 (Farre et al., 2017) and have multiple comorbidities further adding to treatment complexity. Diagnosis and treatment regimens of HF vary significantly among individuals and can be quite complex, especially in elderly patients. The complexity of this disease and lack of appropriate self-management skills results in higher mortality (17%) and frequent health service use such as emergency department visits (Birmingham et al., 2017), hospitalization, and 30-day readmission rates (30%) within a year of diagnosis (Ambrosy et al., 2014).

Hospitalization and 30-Day Readmission

Clarification of terms is needed to facilitate understanding of the HF population under study, where the phrase *persons with HF* is used to describe the population, and *patients with HF* is used to describe those who are currently or have recently experienced an inpatient hospitalization requiring medical treatment. *Health service use* will refer to the unplanned use of hospital-based services including, emergency department visits, hospitalizations, and 30-day readmissions.

National hospital admission rates are rapidly increasing with more than 2% of all hospitalizations attributed to HF (Evans et al., 2014) and its many complications. Once hospitalized, persons with HF experience abnormally long lengths of stay (M=3.4 days) compared with most diseases, with more than 20% of patients staying at least a week and the most critical patients experiencing lengths of stay from several weeks to months (Aranda, Johnson, and Conti, 2009). Of further concern, it is estimated that more than a third of the patients discharged from a HF-related hospitalization will experience an unplanned hospital readmission within 30-days (Allam et al., 2019) and up to half of those patients will be readmitted within six months (Aranda, Johnson, and Conti, 2009). Surprisingly, it is estimated that more than 65% of HF-related hospitalizations and readmissions are preventable. Recent studies indicate that a combination of factors including improper medical management, medication non-adherence, deficits in self-care maintenance, hospital or facility procedures and discharge disposition may impact unnecessary readmissions (Chen et al. 2016). Not only are hospitalizations and readmissions preventable in HF, but they pose an enormous financial burden on the health care systems, and result in poor patient outcomes. It is estimated that up to twenty-five percent of patients experiencing mortality within one year of an inpatient hospital admission (Farre` et al., 2017).

Emergency Department Visits in HF

There are a multitude of studies interested in identifying factors related to increased hospitalizations and 30-day readmissions in persons with HF, however very few studies exist that attempt to predict emergency department visits in the HF population. A recent study (Krieg et al., 2016) conducted as a scoping review noted a significant lack of data related to the prediction of emergency department visits in persons with chronic conditions such as HF. A few studies were found that discuss predictors of emergency department visits in the general population (Posada et al., 2018 & Birmingham et al., 2017 & Pines et al., 2011) as well as in specific groups such as Medicaid users (Capp et al., 2016), however these studies included diverse samples and were not specific to the HF population. Of the general population-based studies of emergency department visits, one study reported that a diagnosis of HF was among the most frequently processed diagnoses in the emergency department (Montoy et al., 2019). Another study of costs related to HF reported that in 2006, there were nearly one million HF-related emergency department visits nationally, from which 80% of patients were directly admitted to the hospital (Storrow et al., 2014). In one recently published study, the authors (Posada et al., 2019) used multiple regressions to identify predictors of clustered emergency department visits in patients with HF; however, the timeline of the study was limited to within six months of a prior hospitalization. After dichotomizing the outcome variable by number of emergency department visits, the study found that persons with HF who visit the emergency department three of more times within 6 months of a hospitalization are typically men, between the ages of 65 and 74 with multiple comorbidities (OR=2.0) (Posada et al., 2019).

Factors that Predict Health Service Use

Several factors have been postulated to predict health service use in persons with HF. The most cited factors in the HF literature include age, gender, race, clinical laboratory values (brain natriuretic peptide, sodium, hemoglobin, and potassium), echocardiogram results, and prior hospital admission data (McLaren et al., 2016 & Chen et al., 2017 & Eapen et al., 2015). However, the significance of these factors in study models is grossly inconsistent and inconclusive. As a result, the relationship among patient factors and health service use in persons with HF remains unclear. Nevertheless, current HF research continues to attempt to clarify these relationships. Factors that were often reported as significant predictors of health service use in current studies include, race, comorbidity, mental health, physical health, and insurance status. Belonging to a minority race of African American (Blecker et al., 2018) or Hispanic (Ponce et al., 2018) is repeatedly predictive of increased health service use in the HF population. Comorbidities including renal failure (McLaren et al, 2016), chronic obstructive pulmonary disease (COPD) and diabetes (Au et al., 2012) were often identified as predictors of frequent hospitalizations in persons with HF. Depression (Chamberlin et al., 2018 & McLaren et al., 2016) and physical limitations as measured by decreased ability to complete activities of daily living (Yamada et al., 2017) were also found to be significant predictors of frequent health service use. Persons with Medicare (OR=1.35) and Medicaid (OR=2.3) insurance coverage (Chamberlin et al., 2018) were found to be extremely high risk for unplanned health service use, specifically 30-day readmissions. However, the factors of age, gender, education, income, BMI, smoking status, marital status, and employment vary significantly in their reported predictability and significance. Age of the individual is heavily debated as a predictive factor, as some authors report those over the age of 65 were more likely to use health services (Hamner et al., 2005), while others report those under 65 are at higher risk (Ahmad et al., 2018). Of note, more than half of persons with HF are over the age of 65 (Ambrosy et al., 2004). Gender, defined as being biologically male or female, is often inconsistently reported across studies, as some studies report females (Ponce et al., 2018 & Mortazavi et al., 2016) are more likely to experience a 30-day readmission, while many other studies reported being male was more predictive (Eapen et al., 2015 & Ahmad et al., 2018). Of the few models that include educational attainment and income, those persons with less than a high school education (Golas et al., 2018) and those with income below the second quartile (Chamberlin et al., 2018) were more likely to experience an unplanned readmission than those with formal college education or of higher socioeconomic status. Marital status is not often included in predictive models and is described as data that are not routinely collected during inpatient hospital stays (Yu et al., 2015). However, marital status is a component of large administrative databases, and was included in the current study model to add depth to the HF literature. There is much debate on the most reliable source of data for such models, as well as the number of factors needed for an accurate predictive model. Many studies have built predictive models that utilize only clinical factors, such as laboratory values from a sample of electronic health records, while other studies focus on only sociodemographic data from administrative datasets. Additional studies have attempted to identify a combination of clinical and demographic factors that may predict health service use; however, the results continue to vary significantly (Bradford et al., 2016). Based on the available data as well as results from prior literature, the variables of age, gender, race, education, employment, income, marital status, insurance, physical function, mental health, smoking status, BMI, comorbidity, and hospital length of stay will be included in the proposed study model.

Individual Characteristic Clusters

A novel approach to predictive modeling has been applied in a small number of studies that attempt to cluster or group individuals based on common characteristics. The clustering approach allows researchers to identify "types" of patients, who are most likely to use health services. This approach is unique in that it may allow for clinicians to have a general picture, or phenotype, of what high risk patients look like so that they may identify and intervene earlier to prevent unnecessary health service use. A few studies (Go et al., 2019 & Son and Won, 2018 & Ahmad et al., 2018) utilized this approach to predict 30-day readmissions in the HF population. Only one was found to identify patient clusters that predict emergency department visits (Posada et al., 2019). Go et al., (2019) identified two clusters of patients, those admitted to the hospital with high frequency (two or more HF admissions in a year) and those admitted with low frequency (fewer than two admissions). The authors report that 25% of the study population was classified in the high frequency cluster and these individuals were likely male, current smokers, and had multiple comorbidities (Go et al., 2019). Another similar study conducted in Korea clustered patient factors by symptomatology, whereby those who reported "bodily pain and energy insufficiency" had the most frequent 30-day hospital readmissions (Son and Won, 2018). One additional study was found to examine clusters of characteristics relative to emergency department visits in persons with HF. The authors dichotomized outcomes as three or more emergency department visits per year and found that those in this category were often male, ages 65-74 with multiple comorbidities (Posada et al., 2019). While these patient clusters may significantly change the mode by which clinicians identify and treat patients with HF, current research is still inconsistent and does not consistently include emergency department use as an outcome.

Large Databases

Large administrative databases are often used in risk prediction and cost-related studies for their ability to provide large samples and robust medical data. The Medicare Expenditure Panel Survey (MEPS) is an example of a large database and was used for the current study. The MEPS database was founded in 1996 and is a large national claims database that collects information regarding both the use of health services and costs of such services across the general population. The MEPS database is managed by the Agency for Healthcare Research and Quality (AHRQ) under the United States Department of Health and Human Services. MEPS data are unique in that they are collected annually and followed longitudinally, allowing for continuous analysis and results dissemination. The MEPS recruits a new study sample annually, then follows each sample over a two-year study period, during which time participants provide information throughout five rounds, or waves, of surveys. Data collection includes two types of surveys, household and insurance, with all MEPS data collected via self-report. Data on demographics, health conditions, income, employment, use of health services and insurance coverage are collected throughout the subsequent waves. On average, MEPS collects data on approximately 30,000 individuals each year.

Conceptual Framework

The Health Service Use Model (Andersen, 1968) was used to guide the study. The Andersen model posits that population characteristics, specifically predisposing, enabling, and evaluated need factors have an impact on health service use. Andersen's original

Health Services Utilization Model (1968) was developed to explain the relationship between individual and community factors that lead to health service use in the general population. Since its development, the model has been applied to a variety of populations, including the HF population. Andersen's original model (Figure 1) and revised model (Figure 2) are displayed below, as components of both models were used in developing the current study model and theoretical substruction. The conceptual model for the proposed study (Figure 3) includes the concepts and variables of interest.

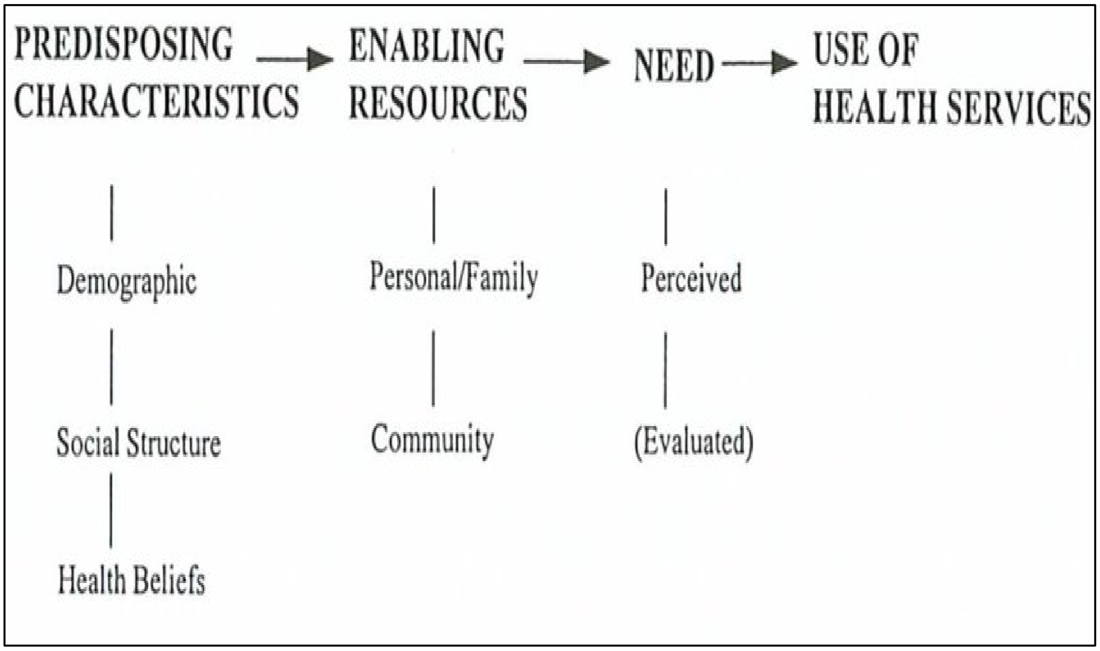


Figure 1- Andersen Health Service Use Model (1968)

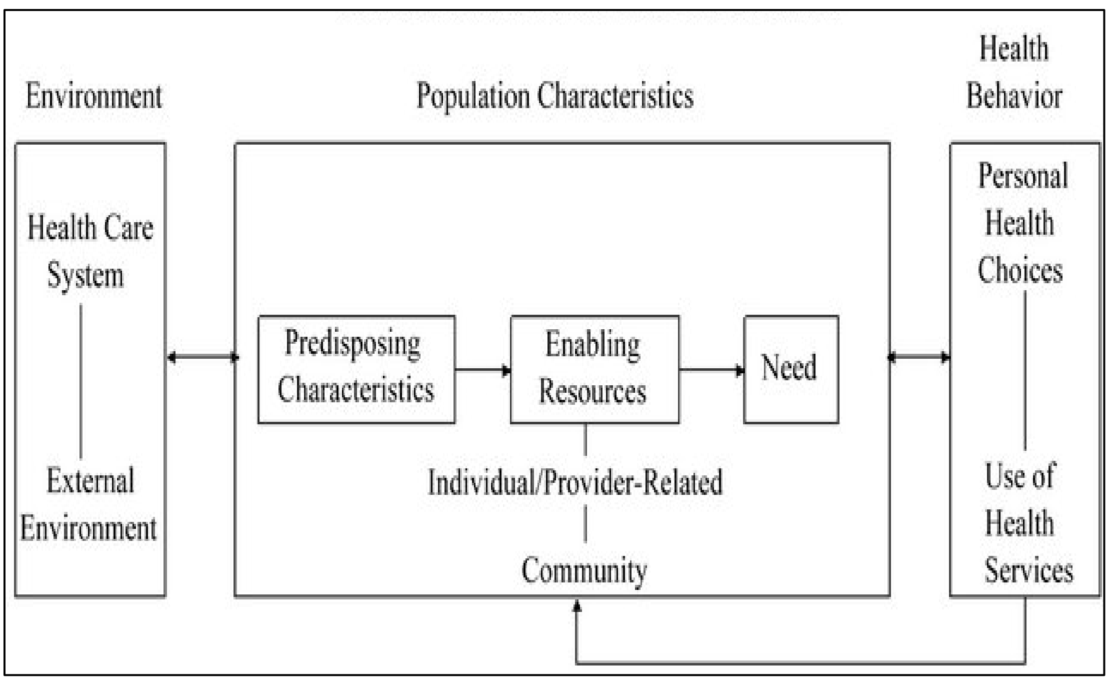


Figure 2 – Andersen's Revised Model (1995)

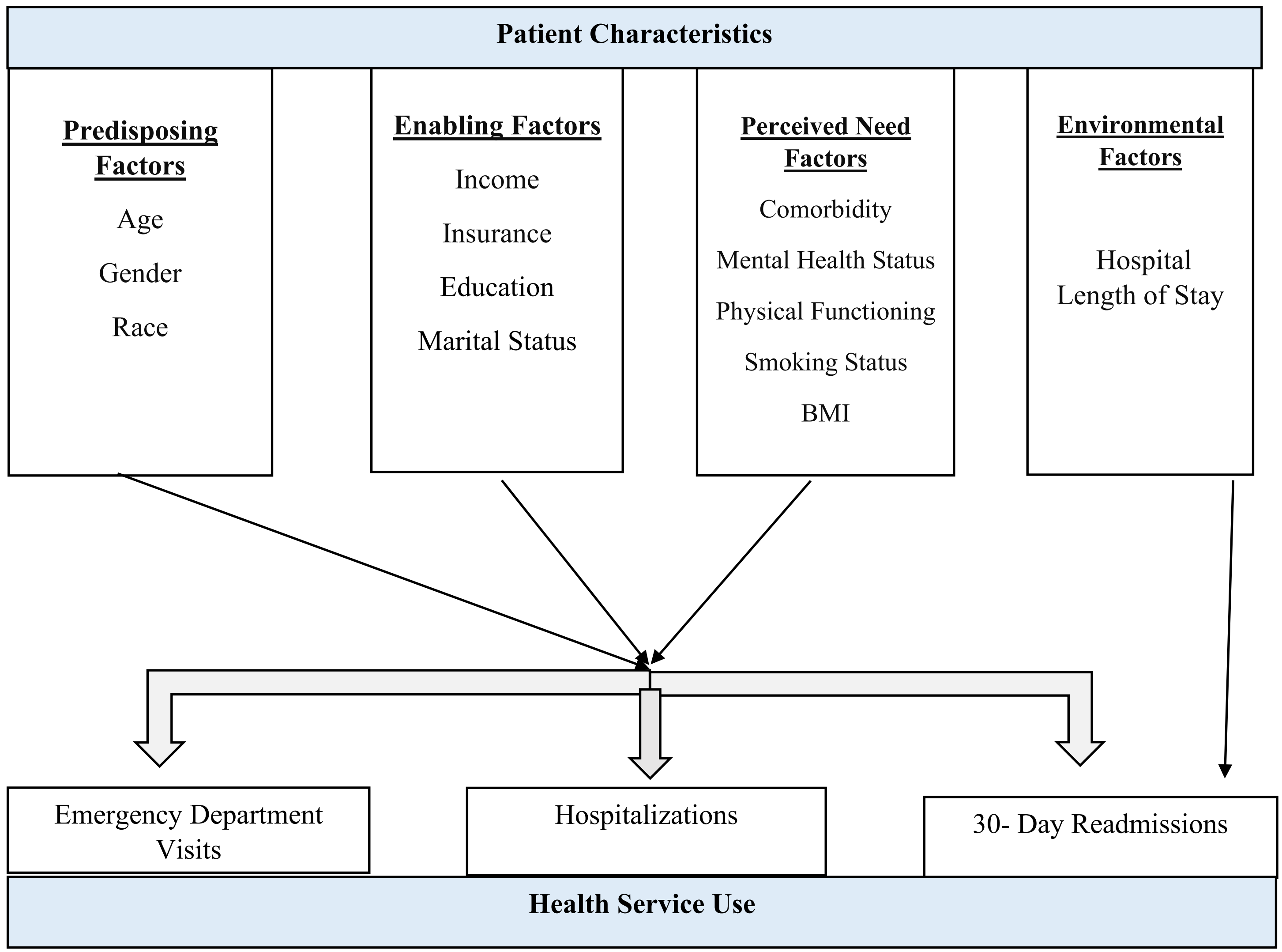


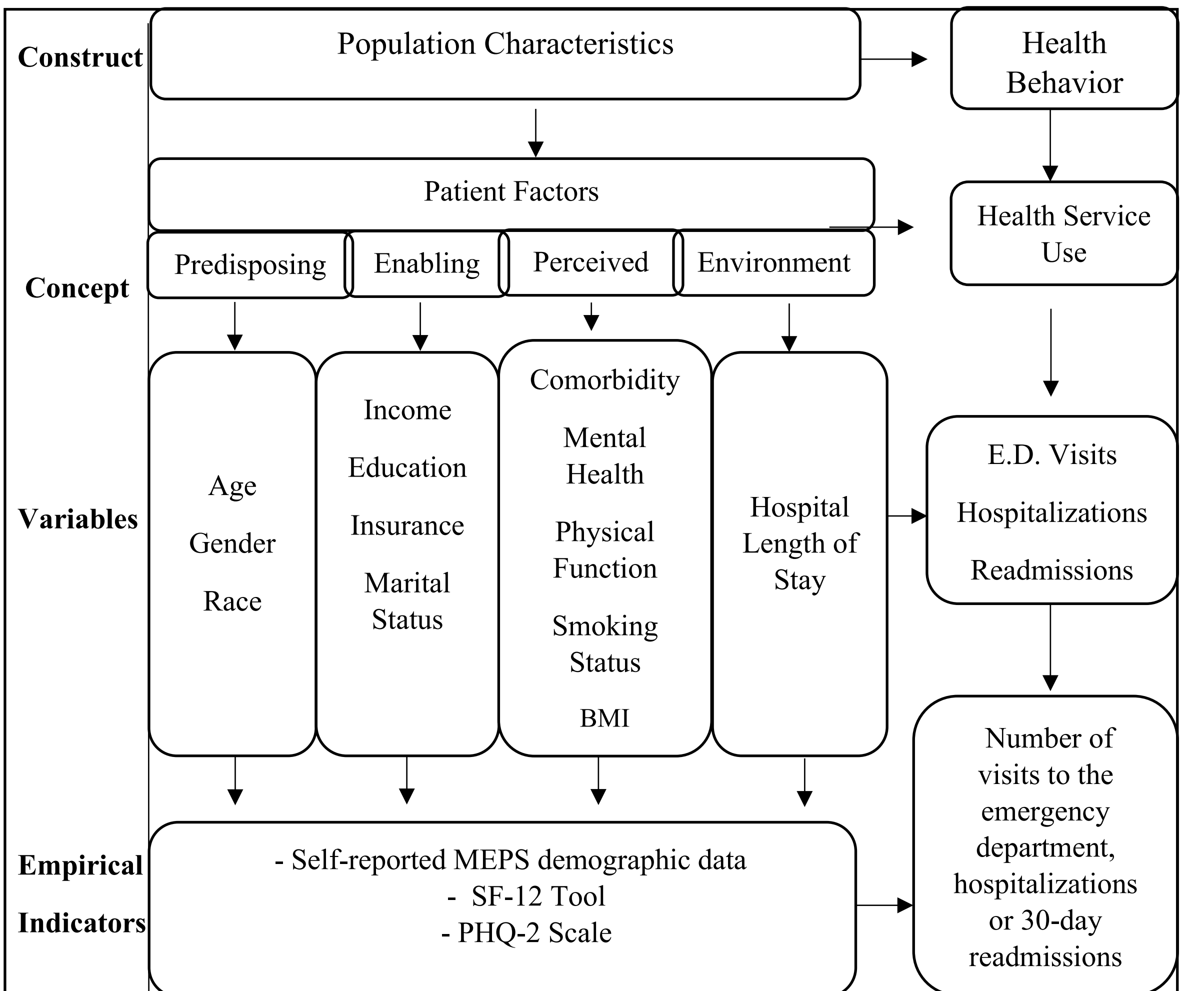
Figure 3 – The Conceptual Model for the Proposed Study

Application of Conceptual Model for the Current Study

The concepts and variables of the study were defined both theoretically and operationally to provide an accurate representation of the relationships among variables. Theoretical definitions are definitions of the concepts and variables of the study are based on previously published studies, reports, and reviews. Theoretical definitions were derived mainly from Andersen's models (1968, 1995). Operational definitions, also called empirical indicators are detailed descriptions of the context in which the variable is being measured. Operational definitions of the variables were derived from the MEPS database from which the data for the current study were originally obtained. Concepts that originated from the Andersen model (1968) that were in the current study model include: predisposing factors, enabling factors and perceived need factors. Concepts that originated from the revised Andersen model (1995) that were used in the current model include the environmental factors. Independent predictor variables for the current study include age, gender, race, income, employment, insurance status, marital status, BMI, smoking status, physical functioning, and mental health. The dependent outcome variable of health service use included emergency department visits, hospitalizations, and 30-day readmissions. Of note, Andersen reports that the enabling factors have the strongest impact on health service use and will be explored in this study (Andersen, 1995).

Theoretical Substruction

The theoretical substruction of the concepts of the study (Figure 4) are presented below. The template for the theoretical substruction was developed by Bekhet and Zausneiwski (2008), to explain the relationships among the constructs, concepts, variables, and empirical indicators of a study. Constructs, concepts, and variables presented in the substruction were derived from Andersen's original (1968) and revised (1995) Healthcare Utilization Models and are also reflected in the conceptual model for the study (Figure 3). A review of the existing HF literature supports a moderate relationship between the constructs and concepts presented in the substruction.



Note: MEPS; Medical Expenditure Panel Survey, SF-12; Short Form 12, PHQ-2; Personal Health Questionnaire

Figure 4. Graphic Representation of the Theoretical Substruction of the Concepts of the Study

Research Questions

The proposed study identified factors that are most predictive of health service use in persons living in the community with HF. Guided by Andersen's Healthcare Utilization Model and using the Medical Expenditure Panel Survey database, the research questions of the study were as follows:

RQ1: What are the relationships among the predisposing (age, gender, and race), enabling (income, employment, education, marital status, and insurance status), perceived need (comorbidity, BMI, physical functioning, mental health, and smoking status) and environmental (length of stay) factors, and:

RQ1a: emergency department visits in persons with HF?

RQ1b: hospitalization in persons with HF?

RQ1c: 30-day hospital readmission in persons with HF?

Exploratory RQ2: Using the significant predictors (predisposing, enabling, perceived need and environmental) how accurate is a model at predicting:

RQ2a: emergency department use in persons with HF?

RQ2b: hospitalizations in persons with HF?

RQ2c: 30-day readmissions in persons with HF

RQ3: Using a time-to-event analysis, what is the likelihood that a person with HF will experience a hospitalization or emergency department visit, and how long on average did it take for a participant to experience a hospitalization or emergency department use after enrollment in the study.

Definition of Concepts in the Study

Predisposing Factors

Predisposing factors are defined as personal characteristics of natural origin that are unable to be adjusted by the patient (Andersen, 1995). Variables of interest in this category include age, gender, and race. While there is some disagreement among current literature as to the true predictive significance of the variables of age, gender and race, they are consistently used in current predictive models. The debate among results may stem from the self-reported nature of data collection for demographic variables.

Age. Age was theoretically defined as the number of years a person has lived on Earth from 0-100 years. Age was operationally defined through self-report based on month and year of birth.

Sex. Sex was theoretically defined as the biological assignment of male or female based on genetic makeup (World Health Organization, 2019). Sex was operationally defined through self-report of male or female.

Race. Race was theoretically defined as a socially defined construct based on self-identification that is not biologically or genetically based (United States Census Bureau, 2020) Race was operationally defined through self-report where an individual can indicate, Caucasian, Hispanic, African American, Asian or other.

Enabling Factors

Enabling factors are defined as personal characteristics that are specific to the person throughout a lifetime that may or may not be adjusted by the individual (Andersen, 1995). Variables in this category include income level, employment status, marital status, level of education and status of insurance coverage.

Income Level. Income level was theoretically defined as the amount, in dollars, that an individual earns in a year. Income was operationally defined based on a self-reported scale of the amount of money an individual earns in a year starting from zero.

Employment Status. Employment status was defined as the established relationship between an individual and an organization, including the time spent earning a wage for services (Centers for Medicare and Medicaid Services, 2019). Employment status was operationally defined by self-report of being unemployed, working full time, or part-time, as well as being retired.

Marital Status. Marital status was theoretically defined as the legal agreement of marriage between two individuals (Medical Expenditure Panel Survey, 2008). Marital status was operationally based on a self-reported of being married, divorced, widowed/er or never married.

Education. Education was theoretically defined as the number of years of completed formal education based on national requirements (Connelly, Gayle and Lambert, 2016). Education level was operationally measured based on self-report of attending high school, some college, or baccalaureate degree, or graduate degree.

Insurance Status. Insurance status was defined as the degree to which an individual is legally and financially eligible for coverage for medical services (Medical Expenditure Panel Survey, 2008). Insurance status was operationally defined by self-report where individuals indicate whether they have a form of insurance such as: Medicare, Medicaid, no insurance or private insurance

Evaluated Need Factors

Evaluated need factors are defined as personal characteristics of health that the patient can choose to change (Andersen, 1995). Variables in this category include the medical comorbidities of which a patient may be currently diagnosed at the time of hospitalization, body mass index (BMI), smoking status, physical functioning status and mental health status.

Comorbidity. Comorbidity was theoretically defined as the presence of any two or more medically diagnosed conditions occurring simultaneously (Valderas et al., 2009). The presence of a comorbidity was operationally defined by self-report regarding the presence of medically diagnosed conditions. For the purpose of the study, the number of comorbidities was determined by the presence of any one or more illnesses occurring simultaneously with a diagnosis of HF.

BMI. Body mass index was theoretically defined as the ratio of height to weight of an individual reported as a percentage. According to the Centers for Disease Control (CDC) BMI ranges are as follows; below 18.5 indicating underweight, 18.5-24.9 indicating ideal weight, 25-29.9 indicating overweight and more than 30 indicating obesity. BMI was operationally determined by a calculation of values given through self-report of an individual's height and weight.

Smoking Status. Smoking status was theoretically defined as the degree to which an individual smokes tobacco. Smoking status was operationally defined by self- report where an individual can indicate "yes" or "no" to currently smoking tobacco.

Physical Functioning. Physical functioning was theoretically defined as an individual's perceived ability to carry out socially defined activities of daily living that individuals are generally expected to be able to do (Medical Expenditure Panel Survey, 2008). Physical functioning was operationally measured through individual composite scores on the Short-Form12 (SF-12) survey (Ware et al., 1996).

Mental Health Status. Mental Health Status was theoretically defined as a dynamic state of internal equilibrium that enables individuals to use their abilities (Galderisi et al., 2015). Mental health status was operationally measured based on individual responses to the Personal Health Questionnaire (PHQ-2) scale (Kroenke, Spitzer and Williams, 2003).

Environmental Factors

Environmental factors were added to Andersen's revised model (1995) and are defined as external factors of influence that are not modifiable due to current health practice and policy. Andersen's revised model (1995) includes environmental factors such as the health care system, and external environment. For the purposes of this study the variable hospital length of stay (LOS) will be included as an environmental factor.

Hospital Length of Stay. Hospital LOS is defined as the number of days, that a patient spends at the hospital during an inpatient admission. Due to the nature of the variable, hospital LOS is often out of the control of the patient and are influenced by the health care system, specifically its current standards and practices. Hospital LOS was determined by self-report for those who have experienced at least one inpatient hospitalization. To ensure accuracy of study outcomes, hospital LOS data was only be used as a factor in the prediction of 30-day readmissions.

Health Service Use

The concept of health service use was originally defined by Andersen as the predisposition of the individual to use and secure services based on individual health behaviors and illness level (Andersen, 1968). Health services are any public medically based service provided to an individual and can include emergency department visits, hospitalizations and 30-day readmissions. For the purpose of this study, health service use was the outcome variable and was operationally defined as the number of times in one year that an individual a) visits the emergency department b) is hospitalized (c) is readmitted to the hospital within 30-days of a prior hospitalization. Many predictive models for 30-day readmission include all-cause readmissions rather than only HF-specific readmissions, therefore the current study reported all-cause health service use. All-cause readmission was defined as any unplanned 30-day readmission after a HF-related hospitalization for any reason. Of note, many studies including all-case readmissions report that only 30% of unplanned 30-day readmissions are related to complications of HF (Bradford et al., 2017), with other major causes including respiratory failure (57%), renal complications (20%), depression or mood disorders (15%) (Walsh and Hripcsak, 2014).

Significance of the Study

The current study presents several opportunities to improve care delivery and outcomes for persons with HF. The significance of the study is aligned with the objectives of the Triple Aim outlined by the Institute for Healthcare Improvement, which seeks to (1) improve the patient experience, (2) improve the health of populations, and (3) reduce the per capita cost of health care. Thus, the outcomes of the proposed study can (1) facilitate the ability of inpatient, outpatient, and homecare and long-term care nurses to make significant adaptations to the delivery of care by understanding who is at risk for increased health service use in order to tailor nursing interventions, (2) address several national reimbursement policies and population health initiatives outlined by the government, and (3) reduce the financial burden of unnecessary health service use on the nation and the health care system through early risk identification and prevention of unnecessary health services use.

Significance to Nursing

Under provision II of the nursing code of ethics, the nurse's primary responsibility is to the patient; whether an individual, family or population ([ANA], 2001). As a group, persons with HF comprise a large and vulnerable percent of the nation's population. As such, the nursing profession has a unique opportunity to contribute to the reduction of unnecessary emergency department visits, hospitalizations, and 30-day hospital readmissions for persons with HF. Not only is it the role of the nurse to promote health across populations but also the responsibility of the discipline to contribute to the process of health service use reduction. To achieve this goal, as related to the concepts and variables of the current study, the nurse can directly impact the perceived need factors that affect health service use. The perceived need factors include the presence of a comorbidity, BMI, smoking status, physical functioning and mental health. Defined by Andersen (1968), perceived need factors are modifiable components of an individual's health status. Nurses can impact patient health outcomes related to the perceived need variables through early recognition of high-risk patients, frequent and consistent patient education, as well as assisting patients with self-management. Specifically, nurses can present programs regarding smoking cessation, engage patients in early and frequent physical therapy and address mental health awareness and coping skills. It is important for nurses to recognize that health service use can be estimated by risk prediction models. In the current study, on average, persons living in the community with HF experienced an unplanned hospitalization or emergency department visit at around 5 months. As such, primary care providers, specifically nurses, should be able to determine the most likely period of time an individual is at greatest risk for health service use. With this information nurses can provide interventions prior to a health decline, especially during routine primary care visits. Further, it is essential that nurses become familiar with risk prediction models and the factors that identify individuals as high risk. While many of the current HF research and prediction models engage in a medically based focus, the current study integrates components of the nursing paradigm into HF research. Specifically, the socioeconomic variables used in the study capture many of the personal, environmental, and holistic health aspects of the nursing paradigm, which provides an opportunity for nurses to recognize and impact health behaviors that lead to unnecessary and repeated health service use.

Significance to Population Health

The current study presents medical providers with several opportunities to improve national health systems and patient outcomes. In addition, the study addressed many national health goals and initiatives set out by leading governmental organizations. The study addressed Healthy People 2020 objective HDS-24, under the topic of Heart Disease and Stroke; to reduce hospitalizations of adults with HF as the principal diagnosis, as well as subcategory HDS-24.1, to reduce hospitalizations of adults aged 65 to 74 years with HF as the principal diagnosis. The current study also addressed the new *Rise Above Heart Failure* initiative set out by the American Heart Association. The *Rise Above* initiative is aimed at increasing awareness of HF symptoms and treatments, as well as reducing the health system impact associated with the disease. The goals of the initiative are to reduce heart failure hospitalizations by ten percent and increase awareness and understanding of this potentially deadly condition by 10% by the year 2020 (Pamela et al., 2017). If modifiable or treatable, the identified factors from the current study may help to significantly decrease health service use, annual health care costs and burden on the country's health care system. Many publications have cited the need for further investigation of the use of data science techniques to develop an accurate model to aid in the prediction of health service use. In Andersen's 1995 revision of the Health Service Use Model, he commented about the model and stated, "This model recognizes personal health practices such as diet, exercise, and self-care as interacting with the use of health services to influence formal outcomes; a comprehensive and systematic perspective which will be relevant and important for the indefinite future".

Significance to Financial Aspects

The global financial burden of HF was estimated to be more than 110 billion dollars in 2012, and the United States HF expenditure alone accounted for more than 20% of this global cost (Cook et al., 2014). Disease related complications of HF amount to more than 17% of the national HF expenditure, often the result of unnecessary emergency department visits and hospital readmissions (Ambrosy et al., 2014). The excessive cost of HF has brought about initiatives to reduce health care expenditures, resulting in significant changes to Medicaid reimbursement protocols. The recent Hospital Readmission Reduction Program (HHRP) developed by Medicaid penalizes hospitals by withholding reimbursements for high numbers of readmissions. In total, there are just over 6000 hospitals in the United States, and in 2017 more than 2500 hospitals, almost half of the nation's hospitals, were penalized for high rates of readmissions (American Hospital Association, 2019). These national penalties resulted in a total withholding of 290 million dollars from U.S. hospitals (Pandey et al., 2016). Without these reimbursements, hospitals are unable to provide proper care to those in need. This is of great concern as it is estimated that the prevalence of heart diseases, including HF are expected to increase by 130% by the year 2030 (Cook et al., 2014). Health systems have set out to reduce health service use, specifically hospitalization and readmission rates for persons with HF. Early identification of these high-risk individuals may allow providers to prevent unnecessary hospitalizations and readmission by initiating early screenings and tailored comprehensive care regimens.

Study Assumptions

Assumptions are the unspoken beliefs about the world that seem so obvious they do not need to be stated explicitly (Brookfield, 1987). The assumptions of the study are as follows.

1. Persons with heart failure do not wish to be frequently readmitted to the hospital.

2. Health outcomes are impacted by a multitude of personal and environmental factors.

3. Reducing unnecessary health service use will improve patient outcomes and reduce financial and economic burdens.

Chapter II

Literature Review

The chapter will provide applications of the relevant concepts, a background of the current health service use trends among persons with HF, a summary of techniques used prior to data analytics in the prediction of health service use among persons with HF, a review of the current analytic techniques and models utilized in the prediction of health service use in persons with HF, and recommendations for future needs based on gaps identified in the literature. The aim of this literature review is twofold; (a) to identify and understand the current state of the science behind factors most predictive of health service use in persons with HF, and (b) to examine current trends surrounding the application of data analytic approaches and models used in the prediction of health service use among persons with HF. For the purposes of this review, it is important to differentiate between the search terms of interest. The term *health service use* will be defined as any unplanned use of health services by an individual diagnosed with HF, specifically visits to the emergency department, hospitalizations or 30-day readmissions. A *hospitalization* is defined as any inpatient hospital stay during a one-year period lasting more than 24 hours, while a *30-day readmission* is defined as an unplanned subsequent admission to a hospital within 30-days after a previous hospital stay.

Characteristics of Persons with Heart Failure

There are currently two recognized categories of HF (preserved or reduced) both differing in disease etiology and treatment recommendations. The two classifications are medically diagnosed by an echocardiogram (ECHO) and are based on cardiac ejection fraction (EF). EF is the percentage of blood pumped out of the left ventricle and into systemic circulation with each heartbeat (AHA, 2015). In the typical healthy adult, a normal EF can range from 60-75%, indicating adequate cardiac function. However, in a patient with HF the EF is typically abnormal, and in some cases can reach critically low levels. In the current HF literature, there is still some debate on standard ranges of EF with most sources citing three ranges: less than 40%, greater than 50% and a "grey" area between 41-49% (AHA, 2019). Heart failure with a reduced ejection fraction (HFrEF), formerly known as systolic HF, is categorized by an EF of less than 40% and results in a decreased flow of oxygenated blood from the heart due to cardiac muscle thinning. Persons with HFrEF are typically younger males with an ischemic coronary disease (Abebe et al., 2016). Conversely, heart failure with a preserved ejection fraction (HFpEF), formerly known as diastolic HF, is categorized by an EF of greater than 50% and is the result of increased filling pressures in the heart often due to hypertension or coronary artery disease (CAD). HFpEF accounts for more than 60% of cases and typically affects older women with hypertension (Abebe et al., 2016). Those individuals with an EF between 41-49% often require further diagnostics based on symptomology to determine the appropriate HF categorization. Until recently, studies only dichotomized the categories of HF as either an EF of less than 40% indicative of HFrEF, or an EF greater than 40% indicative of HFpEF (Solomon et al., 2007). While there are clear etiological differences between types of HF, a standard categorization based on EF has not yet been solidified in medical literature. Heart failure can further be divided into classifications of severity as determined by the New York Heart Association (NYHA). According to the NYHA, classifications of HF fall on a scale of Class I to IV, with no physical limitation seen in Class I, scaling to severe symptoms at rest seen in Class IV. Individuals experiencing class III, or IV HF symptoms are the most likely to seek medical attention (Solomon et al., 2007), however HF classifications are not consistently reported in predictive models.

Impact of Heart Failure on Society and Health Service Use

A paucity of studies examined the national and global burdens of HF on health care systems both procedurally and financially. Complications of HF alone account for an estimated 2% of all hospitalizations, amounting to more than 5 billion dollars in healthcare expenditure annually (Ambrosy et al., 2014). The financial burden of the disease is expected to increase drastically as the population continues to age, with an estimated 50% of individuals with HF being over the age of 65 and 11% being over the age of 80 (Angarall et al., 2019). The inconsistent trajectory of disease progression in HF coupled with the complexity of disease management and treatment regimens, results in a multitude of complications leading to frequent use of health services, especially 30-day readmissions. It is estimated that at least 25%, and in some studies up to 50% of patients hospitalized with a diagnosis of HF will experience an unplanned readmission to the hospital within 30 days of discharge (Allam et al., 2019). Another author reports that up to 45% of patients will be readmitted within six months of discharge (Ross et al., 2008). With the steadily increasing rate of HF diagnoses, it is imperative that health systems continue to work towards reducing the amount of unnecessary health services utilized by this population. Fortunately, unplanned readmissions within 30 days of a HF-related admission are thought to be preventable (Ambrosy et al., 2014) and are often related to poor quality care or discharge management (Chen et al., 2019). However, HF is a complex and multifaceted disease process that requires individualized treatment regimens and consistent self-care management. Often, for a variety of reasons such as inadequate disease knowledge, age, access to care and medication regimen complexity, the disease is mismanaged. Corbretti (2017) reports that approximately 75% of individuals over the age of 65 with HF are taking at least eleven medications per day. Other studies cite adherence to components of self-management in the HF treatment regimen such as daily weighing, medication adherence and symptom recognition as being imperative in preventing disease complications (Hawkins et al, 2012). It is well known that self-management in HF is crucial to preventing unnecessary health service use, however other components of care at the facility level are now thought to impact readmission rates in this population. Despite this knowledge hospitalizations and 30-day readmission rates remain high in the HF population. Hospital-based characteristics such as discharge disposition, medication reconciliation and nurse staffing levels are factors that may be modifiable to prevent unnecessary readmissions (Chen et al., 2019). Regarding readmission risk, one study identified that patients who were discharged to a nursing facility experienced unplanned 30-day readmissions at a higher rate than those who were discharged home (Jiang et al., 2017). As a result, there is a mandate for health care systems to drastically reduce rates of health service use, specifically 30-day readmission in the HF population.

In 2012 as an extension of the Affordable Care Act, the Centers for Medicare and Medicaid Services (CMS) enacted the Hospital Readmission Reduction Program (HRRP) which works to reduce 30-day readmissions by penalizing hospitals for having high rates of readmission. In 2017, as a result of the HRRP program, over 2500 hospitals across the United States were penalized for high readmission rates, resulting in a withholding of reimbursement funds in the amount of $500 million dollars (Kakarmath et al., 2018). Since the enactment of HRRP, there has been impetus for healthcare systems and researchers to shift from inpatient disease management to preventing 30-day readmissions, specifically HF-related readmissions. Initial studies found a slight reduction in readmission rates among persons with HF from 25% in 2009 prior to the HRRP initiative, to 23.5% in 2013 a year after its induction (Chamberlin et al., 2018). The reductions in readmissions were minimal and more recent studies report sustained or increased readmission rates in the HF population. Interestingly, the HRRP program does not include penalties specifically for excessive emergency department use, often a predecessor to a hospital admission. However, persons with HF tend to experience high rates of emergency department use because of mismanagement or misrecognition of worsening symptoms (Montoy et el., 2019). For this reason, it is important for future HF research to focus on models that examine predictors of emergency department use as well as hospitalizations and 30-day readmission.

Data Science and Health Service Use

Before the introduction of the HRRP program in 2012, persons with HF were not routinely screened or followed during an inpatient hospital stay for factors that could lead to increased readmission risk. Prior to the boom of medical technologies including electronic charting and recording, patient data were inconsistently recorded, stored and tracked due to an archaic paper charting system. The national adoption of the Electronic Health Record (EHR) has allowed researchers to access vast amounts of medical data that were previously unavailable or impossible to access. During this time, several risk prediction tools including the LACE screener, described later in this review (van

Walraven et al., 2010) did exist, however until recently, these tools were not studied or validated for use in the HF population. Further advances in statistical analytics have encouraged a new age of methodologies that are being used in large databases to identify unique characteristics of people who may be at greatest risk for health service use. Falling under the realm of data science, predictive modeling techniques including multivariate regression and machine learning, rely on computerized algorithms and logistics to arrive at a best fitting model. While several methods of building predictive models through data analytics have been tested, this review will define and focus on the most cited approaches found in the HF literature including linear, logistic and multivariate regression, random forest, boosting, neural networks and support vector machine (SMV). Many modeling techniques utilize the area-under-the-curve (AUC) or c-statistic as well as the positive predictive value (PPV) to report the effectiveness of the model. For consistency, the AUC statistic will be used in this review to compare the predictive ability and effectiveness of the models. For reference, an AUC statistic of 0.50 would represent a model that could correctly predict an outcome half of the time, no more than random chance, while an AUC statistic of 1.0 could accurately predict an outcome every time. Of note, the mean AUC statistic across models included in this review is 0.69 representing a modest predictive ability in the currently available literature used to predict 30-day readmissions in persons with HF.

Literature Search Methods

A PubMed and CIHNAL search was conducted using the following MESH terms; *heart-failure, health service use OR emergency department OR readmission OR hospitalization, and prediction OR predictive model.* The search was limited to articles published after 1999 and written in English. For the purposes of this review, there was no distinction made between types (HFrEF and HFpEF) or classifications (NYHA I –V) of HF. The initial search yielded 157 related articles. After removing duplicates across databases and all articles that were not specific to the use of health services including emergency department visits or readmissions related to the disease of HF, 48 articles remained for review. Articles that presented models predicting mortality and readmission concurrently were excluded for the purposes of this study. The articles were reviewed with a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist. PRISMA is an evidence-based minimum set of items for reporting in systematic reviews and meta-analyses that focuses on randomized trials but can also be used as a basis for reporting systematic reviews of other types of research, particularly evaluations of interventions (Prisma, 2015). Articles included in the final review were published between 2000 and 2019. Eight studies were conducted and published outside of the United States (China, Korea, Australia, Sweden, Italy and the Netherlands), and two additional studies were conducted in collaboration between the United States and other countries. Of the articles reviewed, twenty-eight were comparative retrospective or prospective cohort studies based on analyses of large claims databases, electronic health records or a combination of the two data sources. Additionally, two articles were conducted as literature reviews of existing data analytic methods used in the prediction of hospital readmissions in persons with HF. Only one article was found to discuss the prediction of emergency department use in persons with HF, however, was limited to 6 months after a hospital discharge. There were no articles published as clinical trials. The language used in the current HF literature regarding health service use prediction is grossly inconsistent, with no studies including the term *health service use* in the title, but various combinations of the term(s*) readmission, rehospitalization, hospitalization or hospitalizations and unplanned admissions*. None of the articles under review used the Medicare Expenditure Panel Survey (MEPS) as a sample for the study. Based on the review of the current literature, there are no studies conducted in the United States to date that have used the MEPS database to (a) produce a model that consistently and effectively predicts health service use in persons with HF or (b) include emergency department use as an independent outcome in health service use prediction.

Concepts in the Literature

The concepts of the study are heavily influenced by the Health Service Use Model (Andersen, 1968) which posits that patient characteristics; specifically predisposing, enabling, evaluated need and environmental factors have an impact on health service use. The original model was developed in 1968 and revised in 1995 to explain health service use in the general population, however for the purposes of this study, the concepts will be applied to the HF population. While none of the studies included in this review specifically cited Andersen's Health Service Use model, many of the variables chosen for the current study are included in the original Andersen model and are consistently utilized throughout the current HF literature. The definitions of Andersen's concepts are presented below and follow with the variables chosen from each concept for use in the current study. The concepts and variables are presented in Chapter 1, Figure 1.

Predisposing Factors

Based on Andersen's model, predisposing factors are defined as personal characteristics of natural origin that are unable to be adjusted by the patient (Andersen, 1995). Variables of interest in this category include age, gender and race. While there is some disagreement among current literature as to the true predictive significance of the variables of age, gender and race, they are consistently used in current predictive models. One study reported that the female gender is more predictive of 30-day hospital readmission (Ponce et al., 2018), however two similar studies (Awan et al., 2019 & Ashfaq et al., 2019) reported men are more likely to be readmitted (OR 1.1, $p<0.01$). Regarding race, many studies report African Americans are more likely to be readmitted (Eapen et al., 2015 & Chamberlin et al., 2018) when compared to other races. Research centered on socioeconomic predictors of health service use in persons with HF tend to report a more impactful significance of age, gender and race than studies focusing on only clinical predictors. While not consistently reported as a significant predictor, the variable of age is included in all study models. One study reported a significantly increased readmission risk for persons with HF over the age of 65 (Hamner and Ellison, 2004), however a more recent study reported that those under 65 (OR=1.0) are more likely to be readmitted (Chamberlin et al., 2018). Much is known about the distinction in types of HF with age, as more than half of the individuals over the age of 65 experience HFpEF (Schopfer and Foreman, 2016). The difference in etiologies of HFrEF and HEpEF based on age may be important for future research to consider when building predictive models. When added to a predictive model, demographic variables such as age, gender and race can improve the performance of the model by up to four percent (Walsh and Hripsak, 2014).

Enabling Factors

Enabling factors are defined as personal characteristics that are assigned to the patient throughout a lifetime that may or may not be adjusted by the individual (Andersen, 1995). Variables in this category include income level, employment status, marital status, level of education, and status of insurance coverage. Like the predisposing factors presented above there is some discrepancy as to their significance and presence in predictive models. Two studies that included level of education in the model found that educational attainment was a significant ($P<0.001$) predictor of readmission (Golas et al., 2018) and individuals with less than high school diploma were likely to experience a hospital readmission than those with formal college education (Eapen et al., 2015). One study reported that employment status and marital status impact the risk of hospital readmission. The authors found more than 65% of retired and 40% of married individuals experienced unplanned a 30-day readmission, while only 4% of employed individuals and 16% of single individuals experiencing a 30-day readmission (Bradford et al., 2017). A similar study confirmed the impact of unemployment (OR 2.59, $p=0.01$) on readmissions (Tsuchihashi et al., 2001). In terms of insurance status, nearly all studies include the variable of insurance coverage in the predictive model. Individuals covered by Medicare and Medicaid make up a vulnerable and misunderstood population who tend to experience high levels of health service use (Chen et al., 2019). One study found nearly 75% of patients who experienced a hospital readmission utilized Medicare, while only 19% used a commercial insurance (Jiang et al., 2019). This was confirmed by another study that reported 80% of patients who experienced a hospital readmission for HF utilized Medicare insurance (Hamner et al., 2005). For this reason, the predisposing characteristic variables of age, gender and race and the enabling factor variables of education, income, employment, marital status and insurance coverage will be included in the model for the proposed study.

Evaluated Need Factors

Evaluated need factors are defined as personal characteristics of health that the patient can change that impact current health-related needs (Andersen, 1995). Such personal characteristics include the medical comorbidities that may be diagnosed at the time of hospitalization, body mass index (BMI), smoking status, physical functioning, and mental health status. One study reported that the number of comorbidities a patient has is an independent predictor of negative outcomes in persons with HF (Hamner and Ellison, 2004). Another study demonstrated that when added to a predictive model, clinical variables and comorbidities can increase the performance of the model by five percent each (Walsh and Hripcsak, 2014). For example, commonly identified comorbidities include end stage renal disease (Keenen at al., 2008), hypertension (McLaren et al. 2016), and diabetes (Au et al., 2012). Over the last decade, mental heal was added to HF predictive models. Johnson and colleagues (2012) reported that persons with diagnosed depression are almost 1.5 times more likely to experience a hospital readmission than those who are not depressed. Similarly, researchers validated these results and identified major depression as a factor highly predictive of hospital readmission (Freeland et al., 2016). Studies have also analyzed physical functioning in HF predictive models. A study conducted in Japan specifically examined the impact of activities of daily living (ADL) on hospital readmissions in persons with HF. This study used the Functional Index Measure (FIM) (Granger et al.,1986) which scores individuals from 18-126 on a total of eighteen functional daily tasks, such as bathing and chores.

Results of the study indicate that for those individuals who scored > 75 on the FIM were more likely to experience a hospital readmission compared to those who scores less than 75 (Kitamura et al., 2019). A similar study (Yamada et al., 2012) using the Performance Measure for Activities of Daily Living (PMADL-8) assessed functional limitation as a predictor of readmission in persons with HF. The PMADL-8 (Shimizu et al., 2010) is an eight-item assessment that scores individuals on a functional scale of 8-32, with higher scores indicative of greater physical limitation. Using a risk cut off score of 20, the authors report that individuals with scores greater than 20 at one month after a hospital discharge were twice as likely to experience a readmission (Yamada et al., 2012). Of note, those who scored above 20 were likely men, over the age of 72 with ischemic HF (Yamada et al). The current study model will include the variables of perceived physical function and psychological health based on the inclusion of these variables in several recent studies examining hospital readmission risk prediction in persons with HF.

Environmental Factors

Environmental factors are defined as external forces that are specific to the healthcare system that may impact the use or attainment of health services (Andersen, 1995). For the current study, hospital length of stay (LOS) is identified as an environmental factor that may impact future health service use. Hospital LOS has been included in several predictive models in the HF literature. Using logistic regression, Allam et al, (2012) found that a previous hospitalization within a year, specifically related to hospital LOS were among the top predictors in the risk prediction model, with an AUC of 0.60. LOS during an inpatient hospitalization is often at the discretion of the medical provider and usually cannot be changed by the patient. One study found that on average, the LOS for patients with HF is 7.9 days, and 8.8 days for patients who experienced a recent 30-day hospital readmission (McLaren et al., 2015). Another author found that increased LOS (11days, p<.001) was significantly associated with unplanned 30-day hospital readmissions (Au et al., 2012).

Health Service Use

The concept of health service use was originally defined by Andersen as the use of any health-related service, however for the purpose of this study, is defined by the outcome variables as the number of times an individual uses (a) the emergency department or (b) experiences a hospitalization or (c) is readmitted to the hospital within a 30-day period. Based on the current HF literature, the prediction of health service use has generally been limited to a 30-day timeframe based on the restrictions set by the HRRP program, and this will remain consistent for the current study. In the identified HF literature, studies related to health service use were limited only to hospital readmissions.

Emergency Department Use in Heart Failure

Most current literature examines the predictors of emergency department visits in the general population and are not specific to HF, though the implications are notable and may be significant when considering future research. Findings from these studies will be presented first to facilitate understanding of emergency department visits in the general population. One study using a robust sample size of over six million emergency department visits, which included multiple ICD-9 codes indicative of HF, found that three percent of patients who visited the emergency department with a diagnosis of HF (0.8% of the total sample) were subsequently admitted to an inpatient unit. The study is not clear on the cause of emergency department use and therefore cannot distinguish between a HF related visit versus an all-cause visit. However, the strongest predictor of emergency department use in the general population was two or more emergency visits in the last six months (OR=3.0) followed by having Medicare based insurance (OR=1.6) and being male (OR=1.12) (Montoy et al., 2019), While this information is helpful, it does not specifically focus on persons with HF and therefore may not be generalizable to the HF population. An additional report was published as a scoping review of national emergency department use. The authors (Kreig et al., 2016) reported that across the twenty articles included in the review, the most frequently cited predictors of chronic emergency department use in the general population are less than a high school education (Sun et al., 2003), presence of mental illness (Bieler et al., 2012), having Medicaid insurance (Griswold et al., 2005) and being from a low-income group (Friedman et al., 2009). A recent study was found to specifically examine emergency department visits in the HF population. Researchers (Posada et al., 2019) used logistic regression to identify common characteristics of patients with HF who experienced clustered emergency department visits (three or more emergency department visits) within six-months after an acute HF hospitalization. Patient characteristics of those who experienced three or more emergency department visits include men (OR=1.0), individuals 65 to 74 years old (OR=0.8) with a Charlson Comorbidity score of >5 chronic conditions (OR=2.1). This study closely resembles components of the proposed study, however, does not include a predictive model for emergency department visits and only examines emergency department visits within a six-month period. To date, there are no articles that have built a HF predictive model including emergency department use as an outcome variable. Further research on emergency department use in the HF population is necessary to identify factors that can then be utilized to build models. It may be possible that many of the predictors of emergency department use in persons with HF may also be predictive of readmission.

Instruments Previously used to Predict Health Service Use

Prior to the use of data analytics to predict health service use in persons with HF, several alternative tools and methods were utilized. However, until recently none were validated for use in the HF population. The LACE index was developed in 2010 in Canada as a tool to predict readmissions in currently hospitalized medical/surgical patients (van Walraven et al., 2010). The tool evaluates a patient's relative readmission risk and produces a risk score from 0-22, where an increased score represents increased readmission risk. The LACE tool takes into consideration several factors of an inpatient hospital stay thought to affect unplanned readmissions including (L) length of stay, (A) acuity of the admission, (C) comorbidities and (E) emergency department visits. Per the reports produced from the original tool, medical-surgical patients scoring above 12 were considered especially high risk for readmission. Since its development, the LACE tool has been applied to several patient populations including surgical, medical and critically ill patients. There has been little consistency in effectively using the tool to accurately predict readmission risk across populations. While not developed for use in the HF population, Wang et al (2014), attempted to validate the LACE tool for readmission prediction in the HF population, but was unsuccessful. Using a cutoff risk score of greater than 10, the authors identified patient characteristics using the LACE tool via retrospective chart review. Results indicated there was no statistical significance in unplanned readmissions based on the LACE score between groups of individuals who were readmitted within 30 days compared those individuals who were not. The insignificant results may have been related to the low cutoff scored chosen for the study, which was lower than the previous score of 12 in the general patient population. Interestingly, the study did show that an increased LACE score was highly correlated with an increased number of emergency department visits within 30 days of discharge (Wang et al.). The authors did not however indicate whether the emergency department visits resulted in a hospitalization. Furthermore, the results suggest that the LACE tool could be useful in initially identifying patients who may be at high risk of emergency department use, often a predecessor to a hospital admission. Another study conducted in Canada compared LACE, LaCE+ and Charlson Comorbidity Index for readmission prediction in persons with HF. In contrast to previous research by Wang et al (2012), Au et al (2012) note that for persons with HF, the LACE cutoff score should be 14, much higher than the previously cited cutoff of 10. The LaCE+ tool includes the variable of age and specific components of the Canadian administrative databases and therefore may not be generalizable to United States databases. The study found that 60% of patients who were readmitted within 30-days experienced a LACE score of at least 14. However, the LACE tool alone only produced an AUC statistic of 0.58, while the LaCE+ produced an AUC of 0.60, and the Charlson Comorbidity Index produced an AUC of 0.55 (Au et al., 2012) indicating a modest predictive ability across tools. The LACE tool alone is not recommended for use in readmission prediction in the HF population (Wang et al., 2014) and further studies of the LACE tool are required to validate and determine the appropriate risk score cutoff for use in the HF population. While potentially effective in readmission risk prediction based on inpatient factors, the components included in the original LACE tool are not a comprehensive representation of many of the socioeconomic characteristics of patients with HF in the community and therefore may not accurately represent the needs of the population. Chamberlin et al., (2018) developed a similar tool to predict readmissions specifically in the HF population. Through a retrospective review and analysis of a large claims database, the Readmission After HF Scale (RAHF) was developed to include the top predictors of HF readmissions and included: age, income, race and comorbidity. Based on the identified predictors, patients were given a risk score where >15 indicated high readmission risk. Characteristics of individuals who scored > 15 includes those who tended to be less than 65 years of age, African American, from low-income households and had many comorbidities (Chamberlin et al., 2018). The RAHF scale reported a r^2 value of 0.9588 indicating a robust ability to explain a large percentage of readmission variability (Chamberlin et al., 2018). The authors report that of the patients who were readmitted within 30 days, only 35% of the readmissions were related to HF exacerbations, with other causes related to renal disorders and pneumonia. Like the LACE tool, the RAHF scale only predicts readmissions for currently hospitalized patients with HF and does not consider outpatient data or sociodemographic variables. For this reason, the LACE and RAHF are not appropriate for use on administrative claims data and therefore will not be considered for use in the current study. There were no additional studies identified validating the use of the RAHF scale in the prediction of readmissions in the HF population, thereby limiting its generalizability. There are only a small number of risk screening tools citing the ability to predict readmissions in patients with HF without the use of data analytic techniques. These tools remain limited in their applicability to the general HF population. The recent changes to electronic health databases have prompted the use of large claims data repositories in nursing research. Large claims databases not only provide an adequate sample population but also include diverse and unique variables often not found in small data files. With the addition of large available data sources and processing abilities of analytic software programs, several recent studies have attempted to incorporate big data repositories with advanced analytics to create models that are able predict readmissions in persons with HF.

Studies Comparing Data Science Techniques

A variety of data science and analytic techniques have been used to determine the effectiveness of each model in the prediction of readmissions in persons with HF. Methods used include traditional logistic regression, LASSO regression, random forest, classification and regression tree (CART), support vector machine, neural networks, as well as gradient boosting and generalized linear model net (GLMN). A study conducted by Lorenzoni et al., (2012) focused solely on the prediction of hospitalization and reported that of the eight analytic techniques tested, GLMN was the most effective method with an AUC statistic of 0.80. Of importance, the authors admit that the sample size was incredibly small *(N=380*) which may have had a significant negative impact on the power and generalizability of the study. Additionally, due to the sample size, there was no external validation or training dataset available, limiting the reliability of the models used for analysis. The remaining seven analytic techniques tested experienced little to no difference in predictive ability, with an average AUC statistic of 0.651. While many existing predictive models focus on both HFrEF and HFpEF, the study by Angraal et al., (2019) is specific to persons diagnosed with HFpEF. This is a valuable contribution to the field, as the etiologies and complications associated with HFpEF and HFrEF differ significantly. It may be necessary in future studies to create predictive models that are able to differentiate between classifications and types of HF. Although current predictive models report similarities in top predictors of readmission across all types of HF, it may be possible that the variables that best predict hospitalization in HFpEF differ from the variables that predict hospitalizations in HFrEF (Angrall et al., 2019). It is difficult to fully compare the results of these studies as the outcomes, methods and patient populations differed significantly. Additionally, one of the two studies cannot be fully reliable based on sample size and approach.

Two of the articles compared advanced machine learning methods to a traditional logistic regression approach. Allam et al., (2018) concluded that a basic logistic regression (AUC 0.64) performed comparably to a variety of neural network approaches (AUC 0.63) in an analysis of a large administrative claims database. A similar article comparing machine learning techniques against a simple logistic regression produced conflicting results. Mortazavi et al., (2016) concluded that compared to logistic regression a random forest approach resulted in 17% more discriminate predictability while a support vector machine (SVM) approach resulted in a 13% increase in the predictive ability of the model. Mahajan, King and Meghban (2019) also note that the use of some predictive model is better than the use of no predictive model. The inconsistency and inconclusive results of these comparative studies offers little guidance for building better models. However, consistent with numerous publications, the authors of these articles agree that a variety of data inputs, specifically the addition of clinical information regarding past hospitalizations (Allam et al., 2019), may greatly improve the predictability of machine learning models. Further supporting the need for various data sources, Mahajan and Ghani (2019) compared three machine learning models whereby structured clinical data (AUC 0.64), unstructured language data (AUC 0.52) and a combination of the two (AUC 0.64) were analyzed. Consistent with previous studies, the model that included both structured and unstructured data slightly outperformed the others. Several authors cite the need for clinical, administrative, social and psychological data when building a predictive model. While the agreement in data sourcing is promising, the average AUC statistic of 0.64 found across predictive models does not hold a great deal of predictive power. Despite the consistently low AUC statistic reported across all articles included in this review, there is not a consistent number of reported predictor variables identified with a model, with some studies reporting as few as eight (Awan et al., 2019) and others report as many as 3500 (Golas et al., 2018). Interestingly, one recent study reported that the use of a feature selection approach with principal component analysis to reduce the number of predictor variables from 47 to 8 (AUC 0.62) without compromising the predictive ability of the model (Awan et al., 2019). The eight predictors reported in this study were age, type of admission, lack of physician visit in 6 months, length of hospital stay, use of antineoplastic drugs, history of HF, chronic kidney disease and depression (Awan et al., 2019). While the predictive power of this model is low, these findings are significant as they may allow for the use of more streamlined models requiring fewer predictor variables to produce consistent AUC statistics. Based on the review of current data analytic methods used to predict readmissions in persons with HF, there is no established or validated method that results in a consistent AUC statistic

New Models Developed to Predict Health Service Use

Several additional models were identified as using a variety of methods in combination with traditional logistic regression. Using a large dataset from the Get with the Guidelines HF Registry, Frizzell et al., (2014) developed a model that included 250 variables and produced an AUC statistic of 0.618. Consistent with other studies (Allam et al.,2019) the authors conclude that a traditional logistic regression performs comparably to advanced machine learning techniques. The authors acknowledge that their model is among many failed attempts in the literature to produce an effective predictive model and suggest that an unidentified number of covariates influencing readmissions are likely present that have not yet been included in a predictive model (Frizzell et al., 2016). Keenan et al., 2008 produced a similar model using a traditional logistic regression based on a large Medicaid database that included 37 variables to produce an AUC statistic of 0.60. The readmission rate identified in this study is consistent with the national average of twenty percent. The authors note that if readmissions could be reduced by even 25% there would likely be 50,000 fewer readmissions per year (Keenen et al., 2008), a major cost-saving initiative. A third study by Evans et al., 2016 used a traditional logistic regression approach with the addition of natural language processing to develop a predictive model for all-cause readmissions based on inpatient hospital data. This study used a relatively small sample size and did not report an AUC statistic; however, the study did report a positive predictive value of 97% rendering the model quite accurate. The remaining four studies developed models through a variety of machine learning techniques and will be reviewed as follows. A unique study by Yu et al., (2015) developed a model using a support vector machine approach focusing specifically on persons over the age of 65. The authors focused on developing predictive models that were specific to each medical institution citing demographic factors as especially influential of readmissions. The model produced an AUC statistic of 0.72, for all-cause readmission, but only 0.63 for HF specific readmissions in people over the age of 65. Because of the complex nature associated with changes to disease processes in the elderly, the authors of this study report that predictive models are likely impossible to develop for use in an elderly HF population (Yu et al., 2015). The authors also suggest that predictive models should be developed that analyze a shorter time post discharge (i.e., one week) if the goal is to focus on HF specific readmissions. This is concurrent with the suggestions of other authors who note that most 30-day readmissions are not related to HF (Frizzell et al., 2016) and that the increased length of time between discharge and an unplanned readmission may allow for additional variables to impact the predictability of a model. Another study used a large database of persons with HF over the age of 65 to produce an all-cause predictive model for readmission. Data taken from the Western Australia Morbidity Registry included 47 variables that produced an AUC statistic of 0.62 respectively. Interestingly, the readmission rate in Australia closely resembles that of the United States, with approximately 24% of the HF population experiencing an unplanned 30-day readmission. The authors note that the AUC statistic alone should not be used to judge the effectiveness of a model, and other classifiers such as sensitivity and specificity should be taken into consideration (Awan et al., 2019). Yet another model was developed by Shameer et al. using a naïve Bayesian algorithm to predict 30-day all-cause readmissions. The model used a retrospective review of a large Mount Saini HF cohort that included 4215 variables extracted from the EHR. The model produced the highest AUC statistic identified in the literature at 0.78. Interestingly, this study reported that of the variables identified for use in the model, the number of medications prescribed was most predictive of 30-day readmissions, with some patients taking upwards of 28 medications (Shameer et al. 2018). Of note, the authors of this study report that outside diagnostic codes other than ICD-9 codes were used in the analysis, potentially causing multiple combinations of missed diagnoses, leading to an overfitting of the model. The final study model included in the review was conducted by Trevethan et al., (2017) to predict HF specific readmissions in a small cohort of patients. The model was developed using a Cox hazards approach and an AUC statistic was not reported. This study was unique as it focused specifically on the inclusion of echocardiogram (EHCO) data as a predictor variable. The authors note that while the results of an ECHO are not independently predictive of a readmission, the right atrial filling pressures increased the predictive ability of the model by 29%. For this reason, the authors recommend an ECHO upon a HF related admission to be included in the clinical data. Another study by Ashfaq et al., used a neural networks model to predict all-cause readmissions in persons with HF. Based on only structured data from the Swedish HF Registry, the model used 8 variables and produced an AUC statistic of 0.71 (Ashfaq et al., 2019). The sample in this study experienced a 27% readmission rate. Of note, the authors found that an increased length of stay (5 days) resulted in decreased risk of 30-day readmissions for sicker patients yet resulted in an increased risk for less sick patients (Ashfaq et al., 2019). Therefore, readmissions may be prevented by reducing unnecessarily long hospital stays for certain classes of patients. A third model using deep neural networks was developed by Golas et al., (2018) to predict all-cause readmissions.

Results of this study were consistent with those from Kakarmath et al., with a slight lower AUC statistic of 0.705. Additionally, the authors note that models predicting readmissions in persons with HF should be inclusive of all types of readmission rather than HF specific, as often there are multiple causes of 30-day readmissions in this population. An additional study, one of the first of its kind, examined patient risk trajectory throughout the course of an inpatient hospitalization to predict readmission risk. This study identified several groupings of patient characteristics from low to high risk based on a multitude of clinical variables. The authors found that dynamic clinical predictors, those that change over time, were most predictive of readmission risk. For example, those who experienced an abnormally low sodium level or high potassium level close to discharge were at greater risk for readmission (Jiang et al., 2017). Results of this study are valuable as they confirm the notion that static variables alone may not be able to effectively predict health service use in the HF population due to the complexity of the disease trajectory. For models that use solely inpatient data to predict heath service use, a dynamic model may be most effective, however further studies of this nature would be needed to confirm the results of the previous study by Jiang et al., 2017. Kakarmath et al., (2018) produced a neural networks model to predict all-cause readmissions in persons with HF. Based on more than 27,000 hospital admissions in an east-coast health system, the model included 3512 variables, utilizing both structured and unstructured data to produce an AUC statistic of 0.71. Twenty-three percent of the sample experienced a readmission within 30 days. Of note, the authors point out that an estimated 25% of the sample may be readmitted to an outside health system and are therefore excluded from readmissions data (Kakarmath et al., 2018). Of the studies identified where predictive models were developed, two did not report an AUC statistic and the others reported inconsistent results. The average AUC static reported among the articles was 0.691, an overall modest predictive ability among models. Across studies, seven different machine learning methods were used with the most popular being logistic regression and neural networking. Data for analysis was obtained from both EHR systems as well as administrative claims data, with some studies using a varied combination of structured and unstructured data. Based on the above review, it is clear there are several gaps and inconsistencies among predictive models in the HF literature. It is important to note that across all models reviewed, there are no two that consistently share methods, predictors, data sources or results.

Models Using Only Socioeconomic Data

While the use of clinical variables is important in the prediction of health service use, few studies have fully examined the effects of testing only socioeconomic data on the predictive ability of a model. Seven articles were found to test factors related to socioeconomic status of a patient and include variables such as income, education, race, and insurance. These articles will be reviewed together as they present a unique perspective into the need for a variety of diverse data sources when building predictive models. Three of the articles in this grouping implicitly reviewed socioeconomic status as a predictor of readmission in persons with HF, and one article cited regional differences between 30-day readmission risk. Another study conducted in Japan considered variables such as financial resources, family caregivers and professional support as predictors for 30-day readmission. This study sample experienced a slightly higher readmission rate (40%) than the average, and results showed that the leading socioeconomic predictors of readmission were both having no occupation and poor attendance to follow up visits (Tsuchihashi et al., 2001). While valuable, the Japanese culture is extremely different from the American culture and socioeconomic factors may not be applicable or generalizable among persons with HF in the United States. An additional study explored the impact of socioeconomic predictors on readmissions in a Hispanic cohort of persons with HF. Of note, data shows that Hispanics are more likely to experience HFrEF, and within that population Hispanic women experience a twofold increase in 30-day readmissions related to complications of HF (Pounce et al., 2018). The results of this study are critically important as they indicate the need for inclusion of cultural and gender variables when analyzing readmission risk in the HF population. The authors note that socioeconomic factors alone may not produce a meaningful impact on readmission in the Hispanic population. Further studies between cultures and ethnicities must be conducted to determine the presence of unique socioeconomic predictors among persons with HF and the impact these predictors have on health service use. Another recent study reported that hospitals who care for vulnerable populations may be disproportionately penalized due to socioeconomic variables that are not considered (Eapen et al., 2015). In accordance with several studies presented in this review, the authors agree that clinical data alone may not be sufficient in the prediction of health service use, as there are distinct health differences across races and populations. Results of these studies indicate that age, race, income and education are all associated with 30-day readmission (Eapen et al., 2015). Another study of EHR data specifically from the Veteran's Health Administration examined regional differences in 30-day readmission rates in persons with HF. While this study utilized clinical rather than socioeconomic data, its results suggest a relationship between geographic or social components relative to readmission prediction. The authors report that across the four regions of the United States, individuals in the southern and eastern regions experienced positive correlations with 30-day readmission risk, while the northern and western regions experienced slightly more negative correlations (Mahajan, Mahajan and Megahban, 2018). The results further support the notion that socioeconomic variables may play a critical role in the predictive ability of a model. Another recent study examined the differences in the predictive ability of clinical and patient characteristics when added to a model. The authors report that among patient-related variables, a distinct group emerged and included those who were disabled, retired or unemployed and over the age of 85 having higher rates of readmission, while those who were using commercial insurance had lower rates (Bradford et al., 2017). Interestingly, the authors note that the highest recurrence of unplanned readmission happened on day 4 post-discharge, yet only 35% of the readmissions were related to complications of HF. A final study in this grouping specifically examined the impact of prior hospital admission on 30-day readmissions in persons with HF. The author reports that individuals who were not recently hospitalized experienced a decreased rate of readmission (14%) compared to those who recently experienced a hospitalization (26%) (McLaren et al., 2015). When these data were used in a predictive model, the AUC statistic improved from 0.57 to 0.62, indicating that prior hospitalizations have a significant impact on 30-day readmissions in persons with HF. The authors also report that persons with HFrEF were less likely to be readmitted than persons with HFpEF. The results of this study continue to support results from previous studies that patient characteristics are necessary in addition to clinical variables to build an effective predictive modes.

Gaps in the HF Literature

A detailed review of the current literature reveals significant gaps across studies related to the use of data science in the prediction of health service use among persons with HF. After appraisal of the articles used in this review it was determined that inconsistencies and gaps fell into three distinct categories including discrepancies in the type of data analytic utilized, inconsistencies in study sample and source of data predictors, and ineffective model generalizability or reliability.

Discrepancies in Data Analytics

In total, more than ten different individual data analytic techniques and twenty various combinations of techniques were utilized to identify factors predictive of health service use in persons with HF. Nearly forty different models were identified in this review. While attempts were made to repeat results across analytic methods, not one method was consistently tested and validated successfully. Notably, several discrepancies exist between the perceived need for complex analytic techniques such as machine learning over simpler approaches like traditional regression. For example, many studies cited better predictive ability with a simple regression (Jiang et al., 2017), while others insisted on more complex techniques such as Random Forest or GLMN analysis (Lorenzoni et al., 2004). To date, no single data analytic technique or combination of techniques has been repeatedly tested and identified as the most effective in the prediction of health service use in persons with HF.

Inconsistencies in Study Sample and Data Source

Discrepancies exist relative to source of sample, data and category of predictors used. The literature has identified many data sources as potentially impactful on health service use prediction including clinical data, Medicare data, demographic data, psychological data and social data. However, across studies included in this review, the identified data sources are inconsistently represented. Many studies included only clinical data, while others included only Medicare data and others included a random combination of data sources. As a result, there is a lack of a consistent list of predictors produced that includes variables from all data sources. In addition, there are discrepancies among studies related to the factors most predictive of health service use within each data source. For example, several studies cite BNP as the most significant clinical laboratory predictor, while other studies cite sodium levels as the most significant. There are additional inconsistencies reported in the number of predictors needed to create an effective model. One study was able to use backwards regression to reduce the number of predictors to eight without compromising the effectiveness of the model (Yu et al., 2015), while other studies utilized as many as 3500 variables. A simple and more parsimonious model would likely be the most generalizable. In addition, the studies included a variety of sample types and size, with samples as small as thirty-five and others as large as 50,000. Many study samples utilized retrospective chart reviews of cohorts of patients previously hospitalized for HF, while other study samples used large Medicare databases. Of note, the criteria used to select the study cohort can have large effects on model performance (Walsh and Hripcsak, 2014). Further, only two study samples focused on the differences among types of HF, including HFrEF versus HFpEF. This may be valuable inclusion criteria for future studies as the predictors of HFrEF may differ from the predictors of HFpEF due to disease etiology. Interestingly, and of great importance, more recent studies cite the significance of including a variety of data sources in an analysis to produce comprehensive results. Unfortunately, there is not a national database available that includes all sources of variables. This is the result of gaps in data sourcing related to inconsistencies in the collection of variables within databases. For example, many authors claim that lab values, imaging files, demographic data, and psychological data are not routinely or consistently collected on all patients during medical visits. Further, because there is not a universal medical records system, patient data were often reported as missing due lack of transference of data across health care systems. To address the inconsistencies across data sources and availability of variables, a universal data system will need to be created that can not only track an individual across all health care systems but also prompts providers to routinely collect uniform variables at each health visit.

Ineffective Model Generalizability and Reliability

Very few studies were able to produce a model with an AUC statistic greater than 0.78, and those that were, have not been validated by repeated studies. The average AUC statistic across articles used in this review was 0.69 representing a modest predictive ability. One study suggested that an effective model will be heavily dependent on individual characteristics of the hospital system, and therefore it may not be possible to develop a single predictive model for health service use in persons with HF, but rather a single model for each health system (Chen et al., 2017). Further discrepancies exist regarding the use of emergency department data as a predictor or outcome in current models. Only one study was identified that examined the use of the emergency department in persons with HF. Further studies should examine the relationship between emergency department use and HF related readmissions.

Conclusion

The many inconsistencies across studies present several opportunities for future research to address the gaps identified in this review. Specifically, future research should focus on identifying and validating the most significant factors that predict of health service use across data sources and samples. Further, models should include emergency department use as an independent outcome. There is also an opportunity to further validate clusters of individuals in the HF population who are at the greatest risk for unplanned health service use. It will also be important to distinguish between types of HF, as the etiologies of HFrEF and HFpEF may require different predictive models and clusters. Finally, it will become necessary for an accurate representation of health service use, to create universal health records that can track patient data across health care systems to prevent missing or incomplete data.

Chapter III

Methods

Purpose

The purpose of the current study was to (a) identify predictors of health service use, (b) develop a model that predicts health service use in persons with HF, and (c) determine the relative risk and time to event of persons with HF. This chapter describes the design, setting, measures, data collection, data analysis methods and protection of human subjects. The research questions for the study are restated below for reference. Plans for data analysis are presented by research question.

RQ1: What is the strength of the relationship between the predisposing (age, gender, race), enabling (income, education, employment, marital status, insurance status), perceived need (comorbidities, BMI, physical functioning, mental health, smoking status), and environmental factors (hospital LOS) and:

RQ1a: emergency department use in persons with HF?

RQ1b: hospitalizations in persons with HF?

RQ1c: 30-day readmissions in persons with HF?

Exploratory RQ2: using the identified predictors from RQ1, how accurate is a model at predicting:

RQ2a: emergency department use in persons with HF?

RQ2b: hospitalizations in persons with HF?

RQ2c: 30-day readmissions in persons with HF.

RQ3: Using a time to event analysis, what is the likelihood that a person with HF will experience a hospitalization or emergency department visit and how long on average did it take for a participant to experience a hospitalization or emergency department use after enrollment in the study?

Design

The current study was conducted as a longitudinal, retrospective, secondary correlation analysis of a cohort of persons with HF from the Medical Expenditure Panel Survey (MEPS) database.

Sample

The sample for the current study was a subsample of persons with HF from the entire MEPS database and was characterized as a convenience sample, based on the availability of data to meet the needs of the researcher. The MEPS collects data from approximately 30,000 individuals each year. From this sample, 1,714 persons with HF were available from the MEPS database spanning the 2003 to 2016 calendar years. The sample included only those individuals with an International Classification of Disease (ICD-9) code of HF during the study period. Heart failure codes range from 428 to 428.9 and include systolic, diastolic, and unspecified HF. Of note, the MEPS data only include up to three digits of the available ICD-9 code, so all individuals with HF under code 428 were included in the study sample.

Inclusion and Exclusion Criteria

The inclusion criteria for the study sample were as follows, (1) an ICD-9 confirmed diagnosis of HF during the study period, and (2) adults above the age of 18. Inclusion criteria were selected based on parameters and results of previous studies. Exclusion criteria for the study were as follows, any cases with missing or incomplete health service use data. Exclusion criteria were selected based on the notion that if included, they would have a significant negative impact on the study outcomes.

Sample Size

The sample size for the current study (N=1,714) was determined by the available data, however a power analysis was conducted to ensure appropriate power parameters. First, a reverse power analysis calculation was conducted using the set sample size to determine the ability of the study to detect a small effect size of 0.01. Then, a traditional power analysis was conducted using the G-Power 3.0 software (Faul et al., 2007) to confirm the sample size required for highly powered study results. Per the G-Power calculation, given a set alpha of 0.01, a power of 0.90, 30 predictors and a medium effect size of 0.33 a sample size of 158 was required. Given the known sample size of 1,714, the study had an adequate sample size to produce the desired effect.

Effect Size

Effect size is defined as the strength and magnitude of the relationship among variables (Sullivan and Feinn, 2012). Cohen's f^2 is used as a measure of effect size when running a multiple logistic regression analysis. A large sample size will likely produce significant results, even if they are not clinically significant (Sullivan and Feinn, 2012). For the current study, an alpha of 0.01 was chosen to reduce the risk of committing a type I error and a power of 0.9 was chosen to reduce the risk of committing a type II error. In the current HF literature, effect sizes among factors of interest and health service use were reported as both Pearson r coefficients and odds ratios. Odds ratios (OR) are used in logistic regression outputs to represent the odds of an outcome occurring given the presence of a particular event (Szumilas, 2010). For example, in a study where a readmission risk score is predicted based on the presence of a given disease such as HF, an odds ratio result of 1.2 would indicate that persons with HF are 1.2 times more likely to experience a readmission outcome than those without the condition. An odds ratio can be converted to an effect size by first using the formula $d = logOR \times \frac{\sqrt{3}}{\pi}$ to obtain a Cohen's D, then the formula $r = \frac{d}{\sqrt{d^2+4}}$ to obtain the correlation coefficient (Borenstein et al., 2009). Finally, an effect size represented as Cohen's f^2 is then calculated as $f^2 = \frac{r^2}{1-r^2}$ where r^2 represents the amount of variance explained by the variable, expressed as a percentage (Cohen, 1992). According to Cohen (1992), an effect size of 0.2 indicates a small effect, 0.3 indicates a medium effect and 0.5 indicates a large effect.

Table 1. Effect Size Parameters for Variables in the Study

Article	**Variables**	**Effect Size (as r)**
Hamner and Ellison, 2005	Living status, length of hospital stay, insurance	r=0.40 ($p<0.05$)
Huyhn et al., 2017	Living/married status	r=0.72 ($p<0.05$)
		Mean r=.55
$F^2 = \frac{r^2}{1-r^2}$		Mean effect size f^2=0.33

Table 2. Odds Ratio Conversion for Variables in the Study

Article	**Variables**	**Results as Odds Ratios**	**Conversion to Cohen's f Effect Size**
Eapen et al., 2015	Education level	OR=0.94 (p=0.002)	f^2=0.06
Ponce et al., 2018	Hispanic Race	OR=2.0 (p<.001)	f^2=0.3
McLaren et al., 2015	Two or more hospitalizations in 1 year	OR=2.93 (p<0.0001)	f^2=0.65
Bradford et al., 2017	Employment Status, Retired	OR=2.3 (p=0.03)	f^2=0.4
Bradford et al., 2017	Hospital LOS >5days	OR=1.56 (p=0.006)	f^2=0.18
Chamberlin et al., 2018	Insurance, Medicare	OR=1.26 (p<0.01)	f^2=0.12
Chamberlin et al., 2018	Comorbidity, COPD	OR=1.15 (p<0.01)	f^2=0.1
1) $d = logOR \times \frac{\sqrt{3}}{\pi}$ 2) $r = \frac{d}{\sqrt{d^2+4}}$ 3) $F^2 = \frac{r^2}{1-r^2}$			Mean f^2=0.25

The current study used a medium effect size of 0.3 based on the reported effect sizes in the literature. The mean effect size determination (Table 1) was based on studies that reported an r coefficient. The mean odds ratios were converted to an effect size and are reported by variable in Table 2**.**

Setting

The MEPS data were originally collected on diverse groups of families and individuals across the United States. Individuals were eligible for participation in the MEPS survey if they had recently (within one year) participated in the National Health Interview Survey (NHIS) under the National Center for Health Statistics.

Measures

The MEPS demographic data were originally collected via self-report using a series of interviews and questionnaires. Physical functioning and mental health data were collected using previously tested and validated measures. To facilitate understanding of the study data, the measures used in the original data collection are presented by study concept below.

Predisposing Factors

Data on the predisposing factors of age, gender and race were collected through a series of surveys broken into two segments, A and B. Survey questions in section A determined the eligibility of each member of the family for participation in the survey. Survey questions in section B measured race, ethnicity, educational attainment, income and marital status.

Age. For the current study, age was measured as the number of years reported in the data file, or birthdate subtracted by current year. Age was measured as continuous data starting from 18 years and 0 days.

Gender. Gender was measured as nominal dichotomous data and will be dummy coded as (0) male and (1) female.

Race. Race was measured as categorical data and coded as (1) white, (2) black,

(3) American Indian, (4) Asian/Pacific Islander (including Chinese), (5) Filipino or (6) Hispanic.

Enabling Factors

Education. For the current study, education level was measured as categorical data and was coded as (1) high school only, (2) some college, (3) bachelor's degree or (4) graduate degree.

Marital Status. Marital status was measured as a categorical variable and coded as (1) married, (2) divorced, (3) widowed/er or (4) never married.

Employment. Original data on employment status were collected through a separate survey which included (a) questions about type of business or industry, firm size, how long the person has worked at each job, whether health insurance was offered, hours worked, and job titles or main duties, and (b) for those not currently working, questions about previous jobs and the reasons for not working (Medical Expenditure Panel Survey, 2008). For the current study, employment status was categorized as (1) unemployed, (2) employed, or (3) retired.

Income. Data on family income were originally collected through a series of questions that assessed information about the household member's federal income tax filing status, specifically about itemized deductions for health insurance premiums, tax credits, wages, other private income sources, and public assistance income (Medical Expenditure Panel Survey, 2008). For the current study, income was considered categorical data and was coded as (1) $0-5000, (2) $5001 – $15,000, (3) $15,001 - $25,000, (4) $25,001-50,000, (5) $50,001- 100,000, (6) $100,001-1250,000 and (7) more than $250,000.

Insurance. Data regarding individual and family insurance coverage were collected through a lengthy series of interviews and surveys and were supported by insurance company reports. Surveys collected information about (a) private health insurance obtained through an employer, direct purchase private insurance plans, and (b) public health insurance and identified the household members covered by health insurance, type of plan, name of each plan, nature of coverage under each plan, duration of coverage, and who pays various costs for the policy premiums (Medical Expenditure Panel Survey, 2008). Insurance surveys also took into consideration those who were uninsured as well as those who receive government support such as Medicare, Medicaid, and TRICARE plans. For the current study, insurance coverage was categorized as (1) uninsured, (2) private employer-sponsored, (3) Medicare and (4) Medicaid or (5) other.

Perceived Need Factors

The data on the perceived need factors of comorbidity, physical functioning, mental health, smoking status and body mass index (BMI) were collected using the Self-Administered Questionnaire (SAQ). The SAQ includes several measures that were previously tested and validated across populations, as described below. Data on smoking status and BMI were collected through a series of general questions within the original household surveys.

BMI. For the current study BMI was calculated as weight in kilograms divided by meters squared. BMI was measured as categorical data and coded as (1) below 18.5 for underweight, (2) 18.5-24.9 for average weight, (3) 25-29.9 overweight, and (4) above 30 as obese.

Smoking. For the current study smoking status was measured as dichotomous data where (0) non-smoker and (1) current smoker.

Physical Functioning. Physical functioning data were assessed using the Short-Form 12 (SF-12) tool. The SF-12 is a twelve-item tool that was derived from the SF-36 scale (Ware, Kosinski and Keller, 1996). The SF-12 is an assessment of an individual's perceived ability to perform activities of daily living (ADL's) and instrumental activities of daily living (IADL's). The SF-12 produces an age-adjusted mean score from 0-100, where a mean score of 50 (SD=10) is used as the indicator for average physical health. Participants who score less than 50, would be in suboptimal physical health, while those who score greater than 50 would be considered in optimal/above average physical health. Initial validation studies produced strong concurrent validity with the original SF-36 (r=0.90) and a high test-retest reliability (r=0.94). The tool has since been tested and validated across several populations including in diabetes, arthritis and heart failure, with all studies producing a Cronbach's alpha > .8 (Ware, Kronske and Keller, 1996 & Failde et al., 2010 & Cheek-Zamora, 2009 & Wee, Davis and Hamel, 2008). For the current study, physical functioning was measured as categorical data and was determined by the group mean and standard deviation with the following groups: (1) below average physical functioning, (2) average physical functioning, and (3) above average physical functioning.

Mental Health Status. Mental health data were assessed using the MCS component of the Short-Form 12 (SF-12) tool. Like the physical functioning component, participants who score less than 50 on the mental health component, would be in suboptimal mental health, while those who score greater than 50 would be considered in optimal/above average mental health. Initial validation studies produced strong concurrent validity with the original SF-36 (r=0.90) and a high test-retest reliability (r=0.94). The tool has since been tested and validated across several populations including in diabetes, arthritis and heart failure, with all studies producing a Cronbach's alpha > .8 (Ware, Kronske and Keller, 1996 & Failde et al., 2010 & Cheek-Zamora, 2009 & Wee, Davis and Hamel, 2008). For the current study, mental health status was measured as categorical data and was determined by the group mean and standard deviation with the following groups: (1) below average mental health, (2) average mental health, and (3) above average mental health.

Comorbidity. In 2000, the MEPS database began collecting specific information on common chronic health conditions, those that are most frequently diagnosed and reported in the general population. The presence of comorbidity was originally determined through self-report. For the current study, the conditions such as asthma, hypertension, coronary heart disease, emphysema, arthritis, cancer, colon cancer, hyperlipidemia, liver disease, thyroid disease, dementia, osteoporosis, kidney disease, hypertension, diabetes, stroke and prior heart attack are considered priority conditions will be included as possible comorbidities. These priority conditions were established by the Canadian Community-Based Primary Healthcare Signature Initiative (CBPHSI), a team of researchers who conducted a scoping review to determine the most studied chronic conditions (Fortin, Almirall &Nicholson, 2017). The conditions were identified in the medical literature based on their (a) relevance to primary care, (2) impact on affected individuals, (3) prevalence among primary care clientele, and (4) frequency of presence in the literature (Fortin, Almirall and Nicholson, 2017). For the current study, data were first measured as dichotomous and coded as (0) condition not present and (1) condition present for each priority condition. Data were then measured as continuous data to incorporate the sum number of comorbid chronic health conditions.

Environmental Factors

The MEPS data on the environmental factor of hospital LOS were collected through self-report for those who had experienced at least one hospitalization during the study period. For the current study, hospital LOS was measured as continuous interval data as the number of nights, starting from 0. Hospital LOS was only included as a factor in the 30-day readmission model.

Health Service Use

Data regarding the outcome variables included in health service use were recorded as the number of emergency department visits, hospitalizations, and 30-day readmissions in a year. Original survey data included the health conditions requiring emergency department care, medical services provided, any surgical procedures performed, prescribed medicines, and the physicians and surgeons providing emergency department care (Medical Expenditure Panel Survey, 2008). Similarly, data on hospital stays include details on the length of stay, reasons or conditions requiring hospitalization, surgical procedures performed, medicines prescribed at discharge, and the physicians and surgeons providing hospital care (Medical Expenditure Panel Survey, 2008). For the current study health service use was measured as dichotomous data for each outcome and coded as follows; for emergency department visits (0) no visits to the emergency department and (1) one or more visits to the emergency department; for hospitalizations (0) no hospitalizations and (1) one or more hospitalizations; and for 30-day readmissions (0) for no 30-day readmissions and (1) for one or more 30-day readmissions.

Procedures

Data Collection and Management

The MEPS data are publicly available and can be downloaded from the database as needed. The sample for the current study (N=1,714) was be extracted from the MEPS database file repository and stored in a secure BOX-drive under the Case Western Reserve University platform**.** MEPS data are already de-identified to ensure human subjects protection. For analysis, data was moved from the BOX-drive to a secure file within the IBM SPSS v26 software program. Data were cleaned prior to analysis through descriptive and frequency statistics. Cleaning ensures that data are normally distributed through measures of central tendency, measures of dispersion, and measures of distribution. Measures of central tendency, including mean, median, and mode, are used to determine averages of variables within a dataset. Measures of dispersion, including standard deviation and range, are used to determine the spread of the data, and posit that 99% of the data are within 3 standard deviations of the mean (Field, 2013). Measures of distribution, including skewness and kurtosis, are used to determine the symmetry and peaked value of the data, where a range of -3 to 3 will for skewness, and -20 to 20 for kurtosis will define normally distributed data. Data found to be non-normally distributed were log transformed to conform data to the assumptions of normality. Due to the large sample size powering the study, cases of data missing at random and cases with extreme outliers were removed from the sample to promote a complete and consistent sample.

Data Analysis

The data were analyzed with several logistic regressions. To assure consistency in the data, assumptions of logistic regression were tested. The primary assumptions of regression affect descriptive statistics and are as follows: (1) variables must vary, that is 90% of the data do not fall in one category which was determined through data frequencies, (2) no outliers or extreme cases, which was confirmed by a Cook's D value of <1 and (3) no multicollinearity among variables which was confirmed by a tolerance of <15.

The sample used in this study (N=1,714), were not only a subsample of the entire MEPS database, but also of the national HF population. For this reason, data were analyzed in SPSS using the complex samples feature. The complex samples feature is traditionally used in large survey data to account for a sampling style that is non-random. Data for the current study was obtained by cluster sampling from a larger population and includes only those persons with a HF diagnosis. Complex samples analyses ensure that correct weights are applied to variables within the sample because they are not occurring independently. This is because the data were obtained in five waves over a two-year period, and respondents are asked to give multiple responses during the study period. Further, each two-year period produces a different composition of a sample of persons with HF. In other words, the samples differ in composition and demographic features and therefore are not equal and should be weighted differently. If run under traditional analysis assuming random sampling, standard errors would be falsely deflated, and the results would not be valid. The complex samples feature in SPSS was activated prior to analysis to ensure appropriate weights are applied to the subsample, ensuring valid results. Plans for data analysis are discussed systematically and guided by the following research questions below.

RQ1a: What is the strength of the relationship among the predisposing (age, gender, race), enabling (income, employment, education, marital status, insurance status), perceived need (comorbidity, BMI, physical function, mental health, smoking status), and environmental factors (hospital LOS), and emergency department use in persons with heart failure? To answer this research question, a logistic regression approach was used. Logistic regression is used to determine the strength of the relationship between an independent predictor variable (IPV) and dichotomous dependent outcome variable (DOV). A separate regression was run for each variable in the study model, including each of the nineteen priority conditions, totaling 31 regressions. Of note, hospital LOS was included as a factor for 30-day readmission, due to the predisposing need for individuals to have experienced a hospitalization to experience a 30-day readmission. Results were interpreted using SPSS outputs as odds ratios and 95% confidence intervals, and factors will be presented sequentially from most to least predictive.

RQ1b: What is the strength of the relationship between the predisposing (age, gender, race), enabling (income, employment, education, marital status, insurance status), perceived need (comorbidity, BMI, physical function, mental health, smoking status), environmental factors (hospital LOS), and hospitalizations in persons with heart failure? This question was answered using the same statistical analysis as in RQ1a, with a logistic regression and interpreted with odds ratios and 95% confidence intervals.

RQ1c: What is the strength of the relationship between the predisposing (age, gender, race), enabling (income, employment, education, marital status, insurance status), perceived need (comorbidity, BMI, physical function, mental health, smoking status), environmental factors (hospital LOS), and 30-day readmissions in persons with heart failure? This question was answered in the same manner as RQ1a-b, with a logistic regression and interpreted with odds ratios and 95% confidence intervals. Hospital LOS was included as a factor.

Exploratory RQ2a: Using the significant factors identified in RQ1 to build a model, how accurate is the model at predicting emergency department visits in persons with HF? To answer this question, a multivariate regression was used to build a model in SPSS v26 with the add-on Modeling package. To improve rigor of the study, the data were split using a testing holdout method, where 80% of the data (N=1,375) were used to build the model, and the remaining 20% (N=340) were used to test the model's predictive ability. The significant factors were simultaneously added to the regression. The effectiveness of the model was determined through the area under the curve (AUC) statistic, as well as the positive and negative predictive values and sensitivity and specificity. The AUC statistic is frequently used in the literature to gauge the ability of the model to correctly predict those who will use health services. For reference, an AUC statistic of 0.5 indicates a model that has no more than random chance for correctly predicting an outcome, while an AUC of 1.0 could correctly predict an outcome 100% of the time. The average AUC static across the current HF literature was roughly 0.7, indicating a modest predictive ability. Positive Predictive Value (PPV) is a representation of a predictive model's ability to accurately categorize individuals who are truly at risk for experiencing the outcome. In other words, PPV of a model is a representation of true positives in the sample. PPV is determined through the formula $PPV = \frac{a}{a+b}(100)$ and is reported as a percentage where a PPV of 1.0 would indicate a model that correctly predicts an outcome 100% of the time (Trevethan, 2017).

Exploratory RQ2b: Using the significant factors identified in RQ1 to build a model, how accurate is the model at predicting hospitalizations in persons with HF? This question was answered in a similar manner to RQ2a and was analyzed using multivariate regression to develop and test a model.

Exploratory RQ2c: Using the significant factors identified in RQ1 to build a model, how accurate is the model at predicting 30-day readmissions in persons with HF? This question was answered in a similar manner to RQ2a-b and used multivariate regression to develop and test a model. Hospital LOS was included in this model.

RQ3: What is the likelihood that a person with HF will experience a hospitalization or emergency department visit and how long on average did it take for a participant to experience a hospitalization or emergency department use after enrollment in the study? To answer this question, a time-to-event survival analysis using a Cox hazard regression (Cox, 1972) was utilized. Survival analysis focuses on the time between two events, specifically, the likelihood that the outcome event will occur in a proposed timeframe (Wright, 2000). The Cox hazards analysis is a set of regressions used to assess survival based on predictive covariates (Kraisangka and Druzdzel, 2018) and an outcome event. The survival analysis is typically described through a hazards ratio, often denoted as h(t), and is the estimated risk that an individual under observation experiences the outcome of interest (Clark et al., 2003). The Cox regression model is a semiparametric approach which assumes (1) that hazard functions are proportional for all levels of time (Gillespie, 2006), meaning that between any set of observations at a given time the hazard risk will remain consistent and (2) that covariates in the model are linear and not time dependent (Schober and Vetter, 2018). Further, it has been estimated that each covariate should experience ten occurrences to produce a non-biased hazard ratio (Peduzzi et al., 1995) In longitudinal studies of time-to-event occurrences, subjects often do not experience the outcome event during the duration of the study. Cox hazards models are unique in that they consider both those who experience the event, and those who do not (Wright, 2000). To achieve an accurate time-to-event estimation, Cox hazards data are often censored. Right censoring occurs when the participant does not experience the outcome event during the study period (Wright, 2000), and therefore it is only known that their time-to-event was longer than the study. Left censoring occurs either when a participant's entry is not known, or the outcome event occurred prior to the start of the study (Wright, 2000). The survival analysis model uses the parameters of right censored data to estimate the time-to-event risk for those individuals who did not experience the event during the study period. The hazard risk ratio is measured on a scale where a score greater than 1 indicates greater risk, and a score less than 1 indicates less risk (George, Seals and Aban, 2014) of experiencing the outcome event. For the current study, the outcome event was either a hospitalization or emergency department visit during the two-year study period. Results are presented as hazards ratios within a 95% confidence interval, as well as graphically in a Kaplan-Meier curve (Kaplan and Meier, 1958).

Human Subjects Protection

Ensuring human subjects protection is an essential component of conducting ethical and fair research. The three principals of human subject's protection are determined by the Belmont Report (1979) and are respect for persons, beneficence and justice. Respect for persons maintains that individuals should be treated as autonomous, and those with limited autonomy should be protected (U.S. Department of Health and Human Services, 1987). Beneficence is defined as the ethical treatment of peoples and is guided by the principles of (a) do no harm and (b) maximize benefits to the individual (U.S. Department of Health and Human Services, 1987). Justice is concerned with the equitable distribution of resources and benefits, as determined by individual need, merit, contribution, and effort (U.S. Department of Health and Human Services, 1987). Though the current study is being conducted as a secondary analysis of a large de-identified dataset, the researcher will maintain these components of ethical research. During the original MEPS data collection participants were instructed on the components of the survey and were asked to sign an informed consent. The MEPS informed consent included: statements about the research, an explanation of the research purpose, a description of the procedures, risks and benefits, the voluntary nature of participation, a statement about compensation, disclosure of participant confidentiality and contact information. Participants were informed of the provisions of data confidentiality under sections 944(c) and 308(d) of the Public Health Service Act [42 U.S.C. 299c-3(c) and 42 U.S.C. 242m(d) (MEPS, 2008). Participants also received a $50 gift card for each completed survey section. Though data are de-identified, it was the responsibility of the researcher to ensure that files were stored properly, analysis was conducted appropriately, and results were correctly disseminated to ensure respect for persons involved in the study.

Study Timeline

The timeline for the current study (Figure 5) is displayed below. The study timeline included an IRB submission in the summer of 2020, followed by data extraction, cleaning and coding over a period of approximately four months. Data analysis took place in the fall of 2020 using the SPSS v26 software program. Once analysis was completed the results and discussion were composed in the spring of 2021. Once completed, the final dissertation will be submitted to the review board at Case Western Reserve University and the study will close on, or before, June 30, 2021.

	May 2020	June 2020	July 2020	Aug 2020	Sept 2020	Oct 2020	Nov 2020	Dec 2020	Jan 2021	Feb 2021	March 2021	April 2021	May 2021	June 2021
IRB	X	X												
Data Collection and Cleaning		X	X	X	X									
Data Analysis					X	X	X	X	X					
Dissertation									X	X	X	X	X	
Dissemination														X

Figure 5 – Study Timeline

Chapter IV

Results

The purpose of the study was to (1) identify the strength of the relationship among the predisposing, enabling, perceived need, and environmental predictors and outcomes of health service use; including hospitalization, 30-day readmission and emergency department use in persons with HF, (2) build a model to predict health service use in the HF population, and (3) determine the time to event and likelihood of health service use for persons with HF.

Prior to analysis, data were cleaned and coded in SPSS v26. Frequencies were run prior to analysis to ensure all assumptions of regression were met. Parameters of skewness (-3, 3) and kurtosis (-20, 20) were met by nearly all variables. Most of the data were normally distributed, however there were four comorbidity variables which violated the assumptions of skewness and kurtosis due to unequal distribution of data. These include the presence of liver disease, kidney disease, obesity, and stroke at baseline. Due to its insignificance in other HF related health service use studies, the variable of liver disease was omitted from the analysis. Additionally, the variable of obesity at baseline was repeatedly measured in the current study as the variable BMI, and therefore was not included in analysis. However, the presence of kidney disease and stroke have been found to be significant predictors of health service use in previous HF studies, and for this reason, were included in analysis despite slight violations in skewness. To counteract the unequal distribution of data and due to the complex multi-stage sampling design and longitudinal nature of the Medical Expenditure Panel Survey (MEPS), the complex samples analysis tool in SPSS was used to ensure accurate estimation of the standard errors and statistical tests based on appropriate weights within the sample. For analysis, the independent predictor variables were compared individually to each outcome variable of health service use.

The sub-sample for the current study (n=1714) was derived from the larger national MEPS database (N=148,747). All subjects in the sub-sample were identified as having heart failure at the time of data collection. Though not utilized for data analysis, it is interesting to note the geographic distribution of the original MEPS HF sub-sample; 14 percent (n=241) of the sample was from the Northeastern part of the United States, 25 percent (n=433) from the Midwest, 44 percent (n=759) from the southern United States, and 16 percent (n=281) from the western United States. Descriptive statistics related to the independent predictor variables and dependent outcome variables of the sample are displayed in Table 3**.**

Predisposing factors. Participants in the sample ranged in age from 21-85 years (M=66.5). Much of the sample was Caucasian (n=1058, 61.7%), with an equal distribution of female (n=967, 56.4%) and male (n=747, 43.6%) participants.

Enabling factors. Nearly half of the sample reported being married (n=764, 44.5%) at the time of enrollment. Almost a third of participants had completed at least a high school education (n=475, 31.6%), were retired (n=532, 31%), making between $10000 and $50000 (N=676, 39.4%) per year, and had some form of public medical insurance (Medicare 33.3%, Medicaid 20.5%) at the time of enrollment.

Perceived need factors. Much of the sample included individuals who were not current smokers (n=1306, 76.2%), did not have a self-reported cognitive impairment (n=1311, 76.4%). More than half of the sample (n=893, 52%) reported being obese at the time of enrollment. Of note, categorizations of BMI within the sample were guided by the

Centers for Disease Control recommendations where a BMI of <18.5% indicate underweight, 18.6-24.5% indicated normal weight, 25-29.9% indicates overweight, and >30% indicates obese (CDC, 2020). Subject's physical functioning and mental health were assessed using the Short-Form 12 (SF-12) questionnaire. The SF-12 questionnaire is comprised of two sections including a physical component (PCS) and a mental component (MCS). Possible scores on the SF-12 range from 0-100 with an average national mean score of 50, indicating average physical functioning and/or mental health. Among the study sample, mean scores for both the physical and mental components were significantly lower than the national average. Among the study sample, scores on the physical component ranged from 6-65 (M=33.01) and on the mental component ranged from 14-77 (M=47.07). Regarding the presence of comorbidity among the sample, analysis was conducted both by number and type of comorbidity. Nearly half of the sample (n=774, 45%) reported having between 4 and 6 comorbidities in addition to their HF diagnosis. Common comorbidities in the sample (i.e., those with more than 20% of the sample affected) included hyperlipidemia (n=876, 51.5%), hypertension (n=1228, 71.6%), arthritis (n=371 (21.6%), asthma/COPD (n=433, 25. %), cardiovascular disease (n=398, 23.2%), and diabetes (n=704, 41%).

Environmental Factors. Hospital LOS was utilized as an independent predictor for only the outcome of 30-day hospital readmission based on the premise of a previous hospitalization being required for a readmission. For those who experienced a hospitalization during the study period (n=691), hospital LOS ranged from 1-173 nights (M=4.4), with the most common duration being between 2 and 6 days (16%).

Outcome Variables. Forty percent (n=691) of the sample experienced a hospitalization during the study period, and of these nearly one-third experienced a 30-day hospital readmission (n=203, 29%). Similarly, forty-one percent of the sample (n=713) experienced at least one emergency department visit during the study period.

Table 3 - Descriptive Statistics of Variables in Study (N=1714)

Main Variable	N	%	Range/Mean	Skewness	Kurtosis
Predisposing Factors					
Age					
>65 (0)	726	42.3	21-85 (M=66.5)	.31	-1.91
<65 (1)	988	57.6			
Gender					
Male (0)	747	43.6		-0.26	-1.94
Female (1)	967	56.4			
Race					
Asian/Pacific Islander (1)	92	5.4			
African American (2)	440	25.7			
Hispanic (3)	124	7.2			
White (4)	1058	61.7			
Enabling Factors					
Education					
<8th grade (1)	221	14.7			
< High School (2)	277	18.4			
Graduate degree (3)	77	5.1			

Main Variable	N	%	Range/Mean	Skewness	Kurtosis
Bachelor's degree (4)	118	7.8			
High School/GED (5)	475	31.6			
Employment					
Not working d/t disability (1)	480	28			
Retired (2)	532	31			
Unemployed (3)	401	23.4			
Employed (4)	298	17.4			
Marital Status					
Never Married (1)	158	9.2			
Widowed/er (2)	449	26.2			
Divorced (3)	343	20			
Married (4)	764	44.5			
Insurance					
Private (1)	612	35.7			
Tricare (2)	53	3.1			
Uninsured (3)	122	7.1			
Medicaid (4)	352	20.5			
Medicare (5)	571	33.3			

Main Variable	N	%	Range/Mean	Skewness	Kurtosis
Family Income					
$0 (1)	530	30.9			
<$10000 (2)	139	8.1			
$10000-$50000 (3)	676	39.4			
>100000 (4)	118	6.9			
$50000-$100000 (5)	215	14.6			
Perceived Need Factors					
Comorbidity					
More than 7 comorbidities (1)	196	11.4			
4-6 comorbidities (2)	774	45.1			
2-3 comorbidities (3)	570	33.2			
0-1 comorbidities (4)	174	10.1			
Hyperlipidemia					
No (0)	838	48.9		-0.04	-2.0
Yes (1)	876	51.5			
Hypertension					
No (0)	486	28.3		-0.96	-1.08
Yes (1)	1228	71.6			

Main Variable	N	%	Range/Mean	Skewness	Kurtosis
Depression				2.21	2.92
No (0)	1493	87.1			
Yes (1)	221	12.9			
Musculoskeletal Disorder				1.45	0.9
No (0)	1359	79.2			
Yes (1)	355	20.7			
Arthritis				1.38	-0.1
No (0)	1343	78.3			
Yes (1)	371	21.6			
Osteoporosis				4.67	17.08
No (0)	1636	95.4			
Yes (1)	78	4.5			
Asthma/COPD				1.14	-0.7
No (0)	1281	74.7			
Yes (1)	433	25.2			
Cardiovascular Disease				1.27	-0.39
No (0)	1316	76.8			
Yes (1)	398	23.2			

Main Variable	N	%	Range/Mean	Skewness	Kurtosis
Stroke					
No (0)	1677	97.8		6.59	41.47
Yes (1)	37	2.2			
Stomach/GI					
No (0)	1472	85.8		2.06	2.26
Yes (1)	242	14.1			
Colon Cancer					
No (0)	1633	95.2		4.27	16.26
Yes (1)	81	4.7			
Diabetes					
No (0)	1010	58.9		0.36	-1.87
Yes (1)	704	41			
Thyroid Disease					
No (0)	1466	85.5		2.02	2.09
Yes (1)	248	14.5			
Cancer (any)					
No (0)	1476	86.1		2.09	2.37
Yes (1)	238	13.9			

Main Variable	N	%	Range/Mean	Skewness	Kurtosis
Kidney Disease					
No (0)	1697	99		9.9	96.12
Yes (1)	17	1			
Urinary/Renal Disease					
No (0)	1534	89.4		2.66	4.66
Yes (1)	180	10.5			
Dementia					
No (0)	1678	97.8		6.6	42.76
Yes (1)	36	2.1			
Physical Functioning (SF-12/PCS)					
25%ile (1) (M=23.79)	430	25.1	0-100 (M=33.01)		
50%ile (2) (M=31.72)	856	49.9			
75%ile (3) (M=41.46)	429	25.0			
Mental Functioning (SF-12/MCS)					
25%ile (1) (M=39.2)	429	25.0			
50%ile (2) (M=48.2)	857	50.0	0-100 (M=47.07)		
75%ile (3) (M=56.59)	429	25.0			

Main Variable	N	%	Range/Mean	Skewness	Kurtosis
Cognitive Impairment					
No (0)	1311	76.4			
Yes (1)	398	23.2			
Body Mass Index (BMI)					
Underweight <18% (1)	35	2.0			
Overweight >25% (2)	428	25.0	9-106 (M=31.1)		
Obese >30% (3)	893	52.1			
Normal weight <24.9% (4)	359	20.9			
Smoking Status					
Non-smoker (1)	1306	76.2		1.84	1.4
Current smoker (2)	251	14.6			
Environmental Factors					
Hospital Length of Stay (N=691)					
>30 days (1)	50	7.5	0-173 (M=4.4)		
1 day (2)	65	9.7			
2-6 days (3)	272	40.8			
7-13 days (4)	172	25.8			
14-20 days (5)	66	9.9			
21-29 days (6)	42	63			

Main Variable	N	%	Range/Mean	Skewness	Kurtosis
Outcome Variables					
Hospitalization in 12 months					
No (0)	1023	59.7	0-10 (M=0.73)	0.4	-1.85
Yes (1)	691	40.3			
30-Day Readmission (N=691)					
No (0)	488	70.6		0.4	-1.85
Yes (1)	203	29.4			
Emergency Department Visit in 12 months					
No (0)	1001	58.4	0-11 (M=0.73)	0.34	-1.89
Yes (1)	713	41.6			

Note: Mean/range only shown for continuous variables, skewness and kurtosis only shown for dichotomous variables.

Results of the Analysis of the Research Questions

Research Question 1a-c. What are the strengths of the relationships among the predisposing (age, gender, race), enabling (income, education, employment, marital status, insurance status), perceived need (comorbidities, BMI, physical functioning, mental health, smoking status), and environmental factors (hospital LOS) and:

RQ1a: hospitalizations in persons with HF?

RQ1b: 30-day hospital readmissions in persons with HF?

RQ1c: emergency department use in persons with HF? Dependent outcome variables for the study were dichotomized as either (0) did not experience or (1) did experience. Results of the logistic regressions (Table 4) are presented by outcome variable and reported as odds ratios, 95% confidence intervals, and significance (OR, 95%CI, p). For the purposes of the study, only variables with significant p-values were reported in this section, as they were used to build the model for research question 2.

Hospitalization. Forty percent of the sample (n=691) experienced a hospitalization within the first 12 months of enrollment in the study. Significant predictors of hospitalization included age, education, baseline comorbidity, SF-12/PCS score, and BMI. Of the predisposing factors, only age, specifically those over the age of 65, was a significant predictor of a hospitalization for persons with HF (OR=1.37, 95%CI=1.01-1.88, $p<0.05$). When compared with individuals under 65, persons over the age of 65 were 1.37 times more likely to experience a hospitalization. When compared to individuals who completed high school/GED, those with a graduate degree were less likely (OR=0.48, 95%CI=0.29-0.80, $p<0.005$) to experience a hospitalization in the first

12 months of the study. While the number of comorbidities present at baseline was not significant, the presence of specific baseline comorbidities was identified as a significant predictor of hospitalization including asthma/COPD (OR=1.26, 95% CI=1.01-1.57, $p<0.04$), cardiovascular disease (OR=1.50, 95% CI=1.18-1.89, $p<0.001$), and the most significant predictor was kidney disease (OR=3.6, 95% CI=1.13-11.44, $p<0.03$). When compared to those scoring in the 75th percentile on the SF-12/PCS component, individuals who scored in the 25th (OR=1.71, 95% CI=1.25-2.35, $p<0.001$) and 50th percentile (OR=1.75, 95% CI=1.33-2.31, $p<0.001$) were at least 1.7 times more likely to experience a hospitalization. Finally, regarding BMI, when compared to those who were of normal weight, those who were obese (BMI > 30%) were less likely to experience a hospitalization (OR=0.77, 95% CI=0.50-0.98, $p<0.03$). The environmental factor of hospital LOS was not used as an independent predictor for the hospitalization outcome. Non-significant predictors of hospitalization include gender, race, marital status, smoking status, insurance, income, employment, mental functioning, cognitive impairment, number of comorbidities, and the presence of baseline hypertension, musculoskeletal disorder, arthritis, stroke, stomach/GI disorder, colon cancer, diabetes, thyroid disease, cancer(any), osteoporosis, dementia. **30-Day Hospital Readmission.** The variable of 30-day hospital readmission was of great importance to the study, as approximately twenty percent of persons with HF nationally are readmitted to the hospital within 30-days of discharge (Keenan et al., 2008). For this study, only persons who experienced a hospitalization within the first 12 months (n=691) were used as part of the sample for this outcome. Of this subgroup that was hospitalized at least once, 29.3% (n=203) experienced an unplanned 30-day hospital

readmission. Significant predictors of a 30-day hospital readmission in persons with HF included Asian race, more than seven comorbidities, underweight BMI, divorced marital status, low family income, and hospital LOS. Of note, the 30-day hospital readmission outcome was the only outcome that included race and family income as significant predictors. When compared with Caucasians, individuals who reported being Asian/Pacific Islander (OR=3.59, 95% CI=1.60-8.08, $p<0.002$) were 3.59 times more likely to experience a 30-day readmission. Pertaining to the perceived need factor of BMI, when compared to those of normal weight, individuals who were underweight were less likely (OR=0.30, 95% CI=0.13-0.70, $p<0.006$) to experience a 30-day readmission. Families who reported incomes between \$10000-\$50000 (OR=1.89, 95% CI=1.11-3.25, $p<0.02$) and those who reported \$0 per year (OR=1.89, 95% CI=1.08-3.32, $p<0.03$) were more likely to experience a 30-day readmission than those who reported between \$50000-\$100000 per year. Individuals who reported being divorced (OR=1.58, 95% CI=1.04-2.41, $p<0.03$) at baseline were more likely than married individuals to experience a 30-day readmission. The number of baseline comorbidities rather than the presence of specific comorbidities, was found to be a significant predictor of 30-day readmission. When compared to individuals who reported having only one comorbidity at baseline, those who reported seven or more comorbidities (OR=3.98, 95% CI=1.85-8.55, $p>0.001$) were 3.98 times more likely to experience an unplanned 30-day readmission. The environmental factor of hospital LOS was added to the regression model for the 30-day readmission outcome. Individuals who were hospitalized for more than two weeks (OR=3.93, 95% CI=2.11-7.34, $p<0.000$), more than three weeks (OR=5.2, 95% CI=2.55-10.62, $p<0.000$) and more than one month (OR=2.27, 95% CI=1.17-4.42, $p<0.02$) were more likely to experience an unplanned 30-day readmission than persons hospitalized for less than one day. However, those hospitalized for less than one week (OR=0.36, 95% CI=0.21-0.61, $p<0.000$) were less likely to experience a 30-day readmission. Variables that were non-significant include gender, age, smoking status, insurance, employment, education, physical functioning, mental functioning, cognitive impairment, and the presence of all specific baseline conditions. **Emergency Department Use.** The outcome variable of interest was emergency department visits within a 12-month period. This outcome is important as there is little research related to predictors of emergency department visits in persons with HF. Forty-one percent (n=713) of the study sample experienced an emergency department visit within the first 12 months of the study. There was a great deal of overlap in the significant predictors of hospitalization and emergency department visits in persons with HF. Significant predictors included age, education, type of comorbidity, SF-12 scores, BMI, and insurance. When compared to those younger than 65, individuals who were older than 65 were 1.3 times more likely (OR=1.34, 95% CI=1.01-1.78, $p<0.04$) to visit the emergency department within the first 12 months of enrollment in the study. When compared to those who were married, individuals who were never married (OR=1.75, 95% CI=1.13-2.67, $p<0.01$) were 1.75 times more likely to visit the emergency department. When compared to individuals with Medicare, those who had Medicaid at baseline (OR=1.49, 95% CI=1.12-1.98, $p<0.006$) were 1.5 times more likely to visit the emergency department. Like the model for hospitalizations, individuals who had a graduate degree (OR=0.60, 95% CI=0.40-0.91, $p<0.02$) were less likely than individuals with a high-school diploma to visit the emergency department. While the number of baseline comorbidities were not found to be significant, several specific baseline comorbidities were found to be individually significant predictors of emergency department use. These include depression (OR=1.64, 95% CI=1.09-1.88, $p<0.01$), asthma/COPD (OR=1.51, 95% CI=1.22-1.89, $p<0.001$), osteoporosis (OR=1.69, 95% CI=1.09-2.63, $p<0.02$) and cardiovascular disease (OR=1.31, 95% CI=1.01-1.68, $p<0.05$). Compared to those who scored in the 75th percentile, persons who scored in the 25th (OR=1.68, 95% CI=1.21-2.33, $p<0.002$) and 50th percentile (OR=1.60, 95% CI=1.22-2.08, $p<0.001$) on the physical functioning component of the SF-12 were at least 1.6 times more likely to visit the emergency department. Individuals who reported being current smokers (OR=1.52, 95% CI=1.15-20.2, $p<0.004$) were 1.5 times more likely to use the emergency department. Finally, regarding BMI, when compared to those with normal weight, individuals who were obese (OR=0.71, 95% CI=0.54-0.93, $p<0.01$) were less likely to use the emergency department. The non-significant predictors of emergency department include gender, race, income, employment, mental functioning, cognitive impairment, number of comorbidities, and the presence of baseline hypertension, musculoskeletal disorder, arthritis, stroke, stomach/GI disorder, colon cancer, diabetes, thyroid disease, cancer(any), kidney disease, dementia.

Table 4 - Significant Independent Predictors by Outcome Variable – Logistic Regression Results

Main Variable	**OR**	**95%CI**	**p**
Hospitalization			
Age			
>65 (reference group: <65)	1.37	1.01-1.88	0.05
Education			
Graduate degree (reference group: high school)	0.48	0.29-0.80	0.005
Comorbidity			
Asthma/COPD (reference group: no disease)	1.26	1.01-1.57	0.04
CVD (reference group: no disease)	1.50	1.18-1.89	<0.001
Kidney Disease (reference group: no disease)	3.60	1.13-11.44	0.03
Physical Functioning (SF-12/PCS)			
25%ile (M=23.79) (reference group: 75^{th}%ile)	1.71	1.25-2.35	<0.001
50%ile (M=31.72) (reference group: 75^{th}%ile)	1.75	1.33-2.31	
Body Mass Index (BMI)			
Obese (reference group: normal weight)	0.77	0.60-0.98	0.04

Note: Only significant predictors of hospitalizations are included in the table. Non-significant predictors include gender, race, marital status, smoking status, insurance, income, employment, mental functioning, cognitive impairment, number of comorbidities, hypertension, musculoskeletal disorder, arthritis, stroke, stomach/GI disorder, colon cancer, diabetes, thyroid disease, cancer(any), osteoporosis, dementia.

Main Variable	OR	95%CI	p
30-Day Hospital Readmission			
Race			
Asian/Pacific Islander (reference group: Caucasian)	3.59	1.60-8.08	0.002
Comorbidity			
>7 comorbidities (reference group: ≤1 comorbidity)	3.98	1.85-8.55	<0.001
Body Mass Index (BMI)			
Underweight (reference group: normal weight)	0.30	0.13-0.70	0.006
Marital Status			
Divorced (reference group: married)	1.58	1.04-2.41	0.03
Income			
$0 (reference group: $50000-$100000)	1.89	1.08-3.32	0.03
$10000-$50000 (reference group: $50000-$100000)	1.89	1.11-3.25	0.02
Hospital LOS			
> 1 month (reference group: ≤ 1 day)	2.27	1.17-4.42	<0.000
< 1 week (reference group: ≤ 1 day)	0.36	0.21-0.61	<0.000
> 2 weeks (reference group: ≤ 1 day)	3.93	2.11-7.34	<0.000
>3 weeks (reference group: ≤ 1 day)	5.2	2.55-10.62	<0.000

Note: Only significant predictors of 30-day hospital readmissions are included in the table. Non-significant predictors include gender, age, smoking status, insurance, employment, education, physical functioning, mental functioning, cognitive impairment, and all specific baseline diseases.

Main Variable	**OR**	**95%CI**	**p**
Emergency Department Visit			
Age			
>65 (reference group: <65)	1.34	1.01-1.78	0.04
Smoking Status			
Current smoker (reference group: non-smoker)	1.52	1.15-2.02	0.004
Insurance			
Medicaid (reference group: Medicaid)	1.49	1.12-1.98	0.006
Education			
Bachelor's degree (reference group: high school)	0.60	0.40-0.91	0.02
Marital Status			
Never Married (reference group: married)	1.74	1.13-2.67	0.01
Comorbidity			
Depression (reference group: no disease)	1.43	1.09-1.88	0.01
CVD (reference group: no disease)	1.31	1.01-1.68	0.05
Asthma/COPD (reference group: no disease)	1.52	1.22-1.89	<0.001
Osteoporosis (reference group: no disease)	1.69	1.09-2.63	0.02
Physical Functioning (SF-12)			
25%ile (refence group: 75^{th}%ile)	1.68	1.21-2.33	0.002
50%ile (refence group: 75^{th}%ile)	1.60	1.22-2.08	0.001
BMI			
Obese (reference group: normal weight)	0.71	0.54-0.93	0.01

Note: Only significant predictors of emergency department visits are included in the table. Non-significant predictors include gender, race, income, employment, mental functioning, cognitive impairment, number of comorbidities, hypertension, musculoskeletal disorder, arthritis, stroke, stomach/GI disorder, colon cancer, diabetes, thyroid disease, cancer(any), kidney, disease, dementia.

Exploratory Research Question 2a-c. Using the significant predictors from RQ1, how accurate is a model at predicting:

RQ2a: Hospitalizations in persons with HF?

RQ2b: 30-Day Readmissions in persons with HF?

RQ2c: Emergency Department visits in persons with HF?

SPSS Modeler v18 was used to create a predictive model for each outcome variable in research questions 2a-c. The SPSS Modeler package is an addition to the original SPSS software program. To complete the model setup, each outcome variable was identified as a "target" variable, while each predictor variable was identified as an "input" variable. The data sample (N=1714) was split into an 80% (n=1371) training and 20% (n=343) testing group for each model. The larger 80% of the sample was used to "train" or create a model, while the remaining 20% of the sample was used to test the accuracy of the model that was developed. SPSS modeler uses an auto-classifier model design, whereby the system can run the given data through up to fifteen varying types of machine learning modeling processes to arrive at a best fit. For the purposes of this study, only four of the most used models in the current HF literature were applied to the data for analysis. These models include logistic regression, Random Forest, Bayesian, and Neural Networks. Examples of studies that use applications of these machine learning models can be found in Chapter 2**.** Outcomes for each model are displayed in Tables 5-7 below. Results for each model's fit and effectiveness include the AUC statistic and model accuracy, sensitivity and specificity, positive and negative predictive value, as well as a list of the most significant predictors within the model represented by percent variance

(r^2). Definitions and formulas for the given comparison statistics are presented below in Table 5.

Table 5 – Definitions and Formulas for Machine Learning Model Comparisons

	Definition	**Formula**
Area Under the Curve (AUC)	The ability of a classifier to correctly distinguish and properly place positive and negative cases on a curve within a data set.	AUC of 0.0 = model cannot correctly classify cases AUC of 0.5 = model is not better than random chance AUC of 1.0 = model can correctly classify cases 100%
Accuracy	The proportion of true positives and true negatives in the given data set.	Correct Predictions / Total Predictions
Sensitivity	The ability of a test to correctly classify an individual with a condition/outcome based on a gold-standard referent test.	True Positive / (True Positive + False Negative)
Specificity	The ability of a test to correctly classify an individual without a condition/outcome based on a gold-standard referent test.	True Negative / (True Negative + False Positive)
Positive Predictive Value (PPV)	The probability an individual will have the condition when the test is positive as compared to other individuals being assessed in a population.	True Positive / (True Positive + False Positive)
Negative Predictive Value (NPV)	The probability an individual will not have the condition when the test is negative as compared to other individuals being assessed in a population.	True Negative / (True Negative + False Negative)

Hospitalizations. The average AUC statistic for the four machine learning models predicting hospitalizations was 0.58, indicating the model performed a slightly better than random chance at correctly predicting hospitalizations in individuals with HF. Among the four models, level of education accounted for at least 20% of the variance within each of the models. In the analysis of machine learning models for the prediction of hospitalizations in persons with HF, the top performing model was a Random Forest approach (AUC 0.592) and influential predictors by percent variance explained include level of education (0.50), asthma/COPD at baseline (0.24), score on the SF-12/PCS (0.14), and cardiovascular disease at baseline (0.11). Though the AUC was just above average, the model was only able to correctly classify those who actually experienced a hospitalization a third of the time.

30-Day Hospital Readmissions. The machine learning analysis of 30-day hospital readmissions produced the strongest predictive ability for any health service use model, with the average AUC statistic of 0.78 between the models. Hospital LOS was the number one influential predictor across all models, with some models assigning nearly 26% of the variance to the LOS variable. The top performing model for 30-day hospital readmissions was the Bayesian approach (AUC 0.747). The most influential predictors of 30-day readmissions in this model by percent variance explained include insurance (0.21), BMI (0.17), race (0.16), marital status (0.13), hospital LOS (0.13) and number of comorbid conditions (0.13).

Emergency Department Visits. The model analysis for emergency department visits produced similar outcomes to the model for hospitalization. The average AUC statistic for the four machine learning models predicting emergency department visits was 0.60, indicating a modest predictive ability. The top performing model was a Neural Network approach (AUC 0.619). Influential predictors of the emergency department model by percent variance explained include score on the SF-12/PCS (0.19), type of insurance (0.14), BMI (0.14), level of education (0.12), asthma at baseline (0.11), marital status (0.10), and smoking status (0.09). The model's predictive ability, as determined by sensitivity and PPV, on average, was better than in the predictive models for other health service use outcomes.

Table 6 - Machine Learning Model Performance Based on Significant Predictors of Hospitalizations

	AUC	Accuracy	Sensitivity	Specificity	PPV	NPV	Predictor	Percent Variance
Random Forest	0.592	59.04%	25.6%	60.3%	48.2%	36.0%	Education Asthma SF-12/PCS CVD	0.50 0.24 0.14 0.11
Logistic Regression	0.584	57.93%	29.8%	82.2%	54.8%	61.7%	CVD Education SF-12/PCS Asthma Kidney Age BMI	0.22 0.20 0.20 0.12 0.09 0.09 0.07
Neural Networks	0.549	60.89%	13.2%	90.3%	46.7%	61.8%	Education CVD BMI Kidney Age SF-12/PCS	0.33 0.25 0.13 0.08 0.08 0.07
Bayesian	0.586	54.46%	35.8%	78.8%	53.5%	64.5%	BMI Kidney Age CVD	0.13 0.09 0.09 0.03

Table 7 - Machine Learning Model Performance Based on Significant Predictors of 30-Day Hospital Readmissions

	AUC	Accuracy	Sensitivity	Specificity	PPV	NPV	Predictor	Percent Variance
Random Forest	0.710	27.0%	11.5%	85.2%	28.9%	64.7%	Marital LOS Income Race	0.05 0.05 0.04 0.04
Logistic Regression	0.730	29.3%	16.3%	89.8%	44.7%	67.9%	LOS Comorbidities Income BMI Marital	0.15 0.14 0.14 0.14 0.14
Neural Networks	0.630	24.4%	30.0%	67.0%	8.9%	85.4%	LOS Race Income Comorbidities	0.26 0.17 0.14 0.14
Bayesian	0.747	26.5%	21.2%	91.1%	57.1%	67.5%	BMI Race Marital LOS Comorbidities	0.17 0.16 0.13 0.13 0.13

Table 8 – Machine Learning Model Performance Based on Significant Predictors of Emergency Department Use

	AUC	Accuracy	Sensitivity	Specificity	PPV	NPV	Predictor	Percent Variance
Random Forest	0.578	58.30%	59.3%	70.0%	57.8%	70.3%	Education Insurance SF-12/PCS Depression Asthma	0.35 0.24 0.16 0.11 0.08
Logistic Regression	0.615	63.83%	46.9%	75.1%	51.7%	71.4%	Insurance Asthma SF-12/PCS Smoker Stroke	0.26 0.26 0.18 0.16 0.15
Neural Networks	0.619	57.93%	47.4%	63.6%	41.3%	69.1%	SF-12/PCS Insurance BMI Education Asthma	0.19 0.14 0.14 0.12 0.11
Bayesian	0.587	58.72%	45.7%	65.9%	41.7%	69.5%	Depression Education Insurance SF-12/PCS Stroke	0.10 0.10 0.10 0.10 0.10

Research Question 3: What is the likelihood and average time after enrollment in the study that a person with HF will experience a hospitalization or emergency department use?

To answer this question a Cox Survival Analysis was completed, and results were reported as hazards ratios. Survival analysis is a methodological approach for analyzing the expected duration of time until one or more outcome events occur (Wright, 2000). The Cox proportional hazards model is a set of regressions used to assess survival based on predictive covariates (Kraisangka and Druzdzel, 2018) and an outcome event. The survival analysis is typically described through a hazards ratio, often denoted as h(t), and is the estimated risk that an individual under observation experiences the outcome of interest (Clark et al., 2003). For the purposes of this study, the following assumptions of a Cox hazards model were met, (1) the event must be mutually exclusive and collectively exhausted states from one another (i.e. experiencing an event vs not experiencing an event), (2) independent-censoring of variables; meaning censoring is a random process unrelated to event occurrence and instead related to planned study end before event occurrence (i.e. participants admitted to the hospital after the end of the twelve months are censored data), (3) all participants are event free at the start of the study, and (4) survival times are measured precisely in months (Klabfiesh and Prentice, 1980). For the current study, the Cox model was used to determine the relative risk that a patient would experience a hospitalization or emergency department visit during the study duration of twelve months, starting from the time they completed the baseline survey. A separate model was run for each dependent outcome variable of hospitalization and emergency department visit. The time dependent variable in the models included number of months from enrollment in the study to an emergency department visit or hospitalization within a 12-month period, while the status dependent variable included the outcome of experiencing at least one emergency department visit or hospitalization during the study period. The independent predictor variables included the original variables as described in the study model (see Figure 3). Results are outlined by outcome variable below. Results were interpreted such that a hazard ratio of >1 indicated an increased risk of health service use as well as a shortened time to event, while a hazard ratio of <1 indicated a decreased risk of health service use and longer time to event.

Hospitalizations

For the Cox proportional hazards analysis on hospitalizations, 40.3% of the total study sample (n=691) experienced at least one hospitalization within the twelve-month study period. Within this sub-sample of persons who were hospitalized, the average number of months to a hospitalization event was 5.3 months after study enrollment (SD=3.2), with 11.5% (n=79) of individuals experiencing a hospitalization within the first month after enrollment. Just over 38% of hospitalizations in this sub-sample occurred within the first three months of study enrollment. The fewest number of hospitalizations occurred in month 11 (n=38, 5.5%). Interestingly, 4% of the sample (n=86) experienced three hospitalizations during the study period, while 22% (n=383) experienced only one hospitalization. Of further importance, of those who were hospitalized at least once, nearly 30% (n=203) experienced an unplanned 30-day readmissions. Those who experienced a hospitalization in or after month twelve (n=1032) were considered to have no hospitalization during the study-period and were counted as right-censored data. Of note, due to limits in the study data, we were unable to differentiate between datapoints that occurred in month twelve. In other words, it is unclear whether a datapoint in month twelve was related to a hospitalization in month twelve or related to no event during the study duration. For this reason, all data points in month twelve were considered as right-censored data.

The Cox analyses for hospitalizations were first run with all independent predictor variables and then run with just the nineteen comorbidity predictors. There were no cases with negative time, and no cases that were censored prior to the end of the study period. Results from the Cox proportional hazards model are presented in Table 9. Results of significant predictors are presented as (HR, CI, p). The overall model fit was significant (89.29, 32, $p<.000$) when looking at all predictors simultaneously. The delta-change in the -2Log-Likelyihood between the base model with no predictors (8716.98) and the final model with predictors (8619.68) was minimal indicating a good fit of the model that includes independent predictor variables versus the null model. Among the independent predictor variables in the model, significance was seen in the following variables; age greater than 65 (HR=1.23, 95% CI=1.01-1.55, $p<0.07$), SF-12 scores on the physical functioning component both in the 25th (HR=1.73, 95% CI=1.33-2.26, $p<0.001$) and the 50th percentile (HR=1.61, 95% CI=1.27-2.04, $p<0.000$), having a graduate degree (HR=0.60, 95% CI=0.36-1.02, $p<0.06$), overweight BMI (HR=0.71, 95% CI=0.56-0.91, $p<0.006$), as well as having asthma/COPD (HR=1.29, 95% CI=1.09-1.53, $p<0.003$), diabetes (HR=1.18, 95% CI=1.01-1.34, $p<0.04$), stroke (HR=1.65, 95% CI=1.06-2.49, $p< 0.03$) and cardiovascular disease (HR=1.32, 95% CI=1.11-1.57, $p<0.002$) at baseline. A graphic of the cumulative hazard function of all independent predictors for hospitalizations is displayed below in Figure 6. Based on the hazard function of all categorical predictors within the sample, just over 10% of the sample was hospitalized within the first three months of study enrollment, 25% of the sample was hospitalized within the first seven months of study enrollment and nearly 60% of the sample never experienced a hospitalization during the study duration. Based on the study results, individuals at increased risk for hospitalization with shorter time to event include those with poor physical functioning (SF-12), and those who reported baseline asthma/COPD, cardiovascular disease, stroke, and diabetes. Individuals at decreased risk for hospitalization with longer time to event include those with overweight BMI and a graduate degree. Individual Kaplan Meier curves were run for variables in the hazards model that were significant (Figures 8-14). The Kaplan Meier analysis is useful when comparing the relative risks of two groups simultaneously, for example, the difference in months to hospitalization or emergency department visit in those with or without a given condition. For the current study, the time to event between groups were compared using the sample mean (in months). We found that individuals over 65 experienced a hospitalization (M=6 months) on average, two months earlier those under the age of 65 (M=8 months). Interestingly, individuals who scored in the 25th or 50th percentile on the SF-12/PCS experienced a hospitalization at 7 months, however those who scored in the 75th percentile experienced an event in, or after month 12. The sample average time to event for persons with stroke (M=4 months) was significantly shorter than those persons without stoke (M=10 months) at baseline. Similarly, those who had diabetes (M=5 months) experienced a hospitalization roughly 2 months before a person without diabetes (M=7 months).

Model comparisons for identifying the significant difference between groups within factors include Log-Rank, Breslow and Tarone-Ware statistical tests, which are presented below as (chi-square, df, p). The Log-Rank comparison is the most widely accepted model comparison method as it gives equivalent weights to all data points, however, all three model comparison statistics will be presented for each individual Kaplan Meier curve. Significance in the model comparison statistics indicate a true difference in risk between groups for the outcome of interest. The Kaplan Meier curve for individuals over the age of 65 showed significance across comparisons including Log-Rank (9.52, df=1, $p<0.002$), Breslow (9.26, 1, $p<0.002$), and Tarone-Ware (9.49, df=1, $p<0.002$). The model comparison for the SF-12/PCS component included Log-Rank (38.51, df=2, $p<0.001$), Breslow (37.72, df=2, $p<0.001$), and Tarone-Ware (38.40, df=2, $p<0.001$). For individuals with asthma/COPD at baseline, the model was significant for the Log-Rank (9.07, df=1, $p<0.003$), Breslow (10.86, df=1, $p<0.001$), and Tarone-Ware (10.01, df=1, $p<0.002$). For individuals with cardiovascular disease at baseline, model comparisons were significant for Log-Rank (15.79, df=1, $p<0.001$), Breslow (20.13, df=1, $p<0.001$), and Tarone-Ware (18.02, df=1, $p<0.001$).

Table 9 - Hazards Ratios by Independent Predictor Variable for Hospitalizations – COX Survival Results

Main Variable	HR	95%CI	P
Hospitalizations			
Age			
>65	1.33	1.14-1.56	<0.001*
Gender			
Female	1.02	0.88-1.19	0.77
Race			
Asian/Pacific Islander	0.85	0.6-1.22	0.38
African American	1.01	0.85-1.20	0.91
Hispanic	0.82	0.60-1.12	0.22
Education			
Less than 8th grade	1.18	0.93-1.49	0.18
Less than high school	0.94	0.74-1.82	0.59
Graduate degree	0.60	0.36-1.03	0.06*
Some college	0.93	0.75-1.15	0.49
Bachelor's degree	0.80	0.57-1.12	0.19
Employment			
Not working d/t disability	1.23	0.91-1.68	0.19
Retired	0.96	0.70-1.13	0.39
Unemployed	1.04	0.75-1.44	0.83

Main Variable	HR	95%CI	p
Marital Status			
Never Married	1.10	0.83-1.44	0.52
Widowed/er	1.00	0.78-1.29	0.97
Divorced	1.18	0.96-1.44	0.12
Insurance			
Private	0.89	0.72-1.01	0.30
Tricare	0.89	0.55-1.46	0.65
Uninsured	0.84	0.58-1.23	0.37
Medicaid	1.04	0.82-1.31	0.77
Family Income			
$0	1.13	0.83-1.54	0.44
$10000-$50000	1.21	0.87-1.68	0.25
$50000-$100000	1.22	0.97-1.55	0.10
>100000	0.65	0.37-1.13	0.13
Current Smoker			
Current Smoker	0.97	0.67-1.42	0.89
Cognitive Impairment			
Presence of Impairment	0.94	0.77-1.14	0.51

Main Variable	HR	95%CI	p
Physical Functioning (SF-12)			
25%ile (M=23.79)	2.30	1.80-2.94	<0.001*
50%ile (M=31.72)	1.91	1.52-2.40	<0.001*
Mental Functioning (SF-12/MCS)			
25%ile (M=39.2)	1.11	0.85-1.44	0.44
50%ile (M=56.58)	0.99	0.79-1.24	0.91
Comorbidity			
More than 7 comorbidities	1.18	0.79-1.78	0.42
4-6 comorbidities	1.04	0.74-1.45	0.83
2-3 comorbidities	0.83	0.60-1.61	0.28
Depression	1.16	0.83-1.46	0.20
Asthma/COPD	1.29	1.09-1.53	<0.003*
Cardiovascular Disease	1.32	1.11-1.57	<0.002*
Stroke	1.63	1.06-2.49	0.03*
Stomach/GI	1.08	0.86-1.35	0.50
Colon Cancer	1.29	0.91-1.82	0.15
Diabetes	1.18	1.01-1.34	0.04*
Kidney Disease	1.81	0.89-3.70	0.15

Main Variable	**HR**	**95%CI**	**p**
Colon Cancer	1.29	0.91-1.82	0.15
Diabetes	1.13	0.96-1.34	0.10
Hypertension	0.99	0.83-1.21	0.99
Muscular Disease	1.09	0.89-1.32	0.42
Arthritis	0.84	0.68-1.03	0.10
Osteoporosis	1.22	0.85-1.75	0.28
Thyroid Disease	1.05	0.84-1.31	0.69
Cancer	1.06	0.85-1.34	0.60
Urinary Disease	1.01	0.78-1.32	0.92
Dementia	0.92	0.53-1.60	0.76
Hyperlipidemia	0.89	0.75-1.05	00.17
Obesity	1.39	0.66-2.95	0.39
Body Mass Index (BMI)			
Underweight <18%	1.09	0.65-1.82	0.76
Overweight >25%	0.82	0.67-0.99	0.05*
Obese >30%	0.90	0.72-1.12	0.34

Note: Predictor significance is represented with (*)

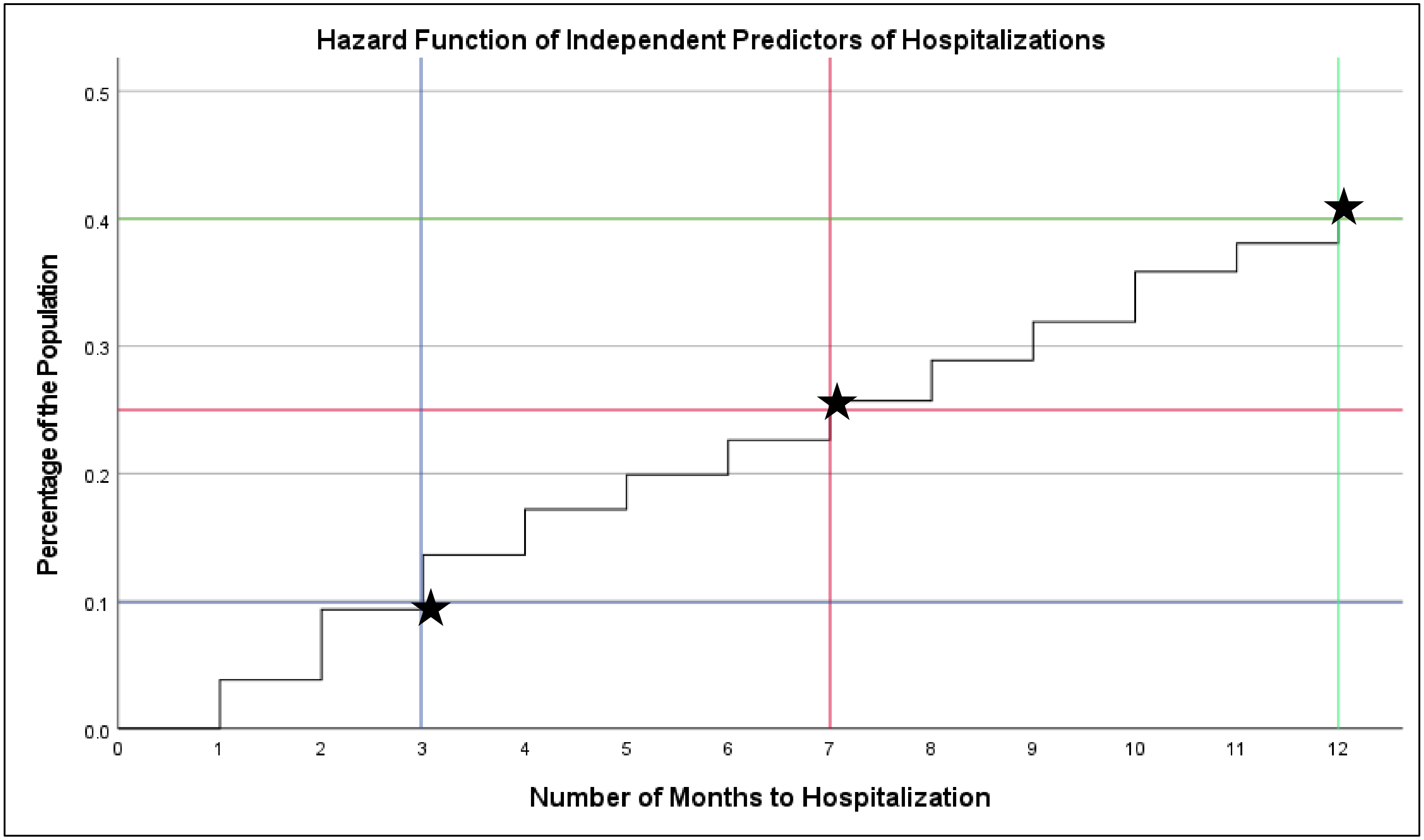


Figure 6 – Hazard Function for Hospitalizations from the Total Sample (N=1714)

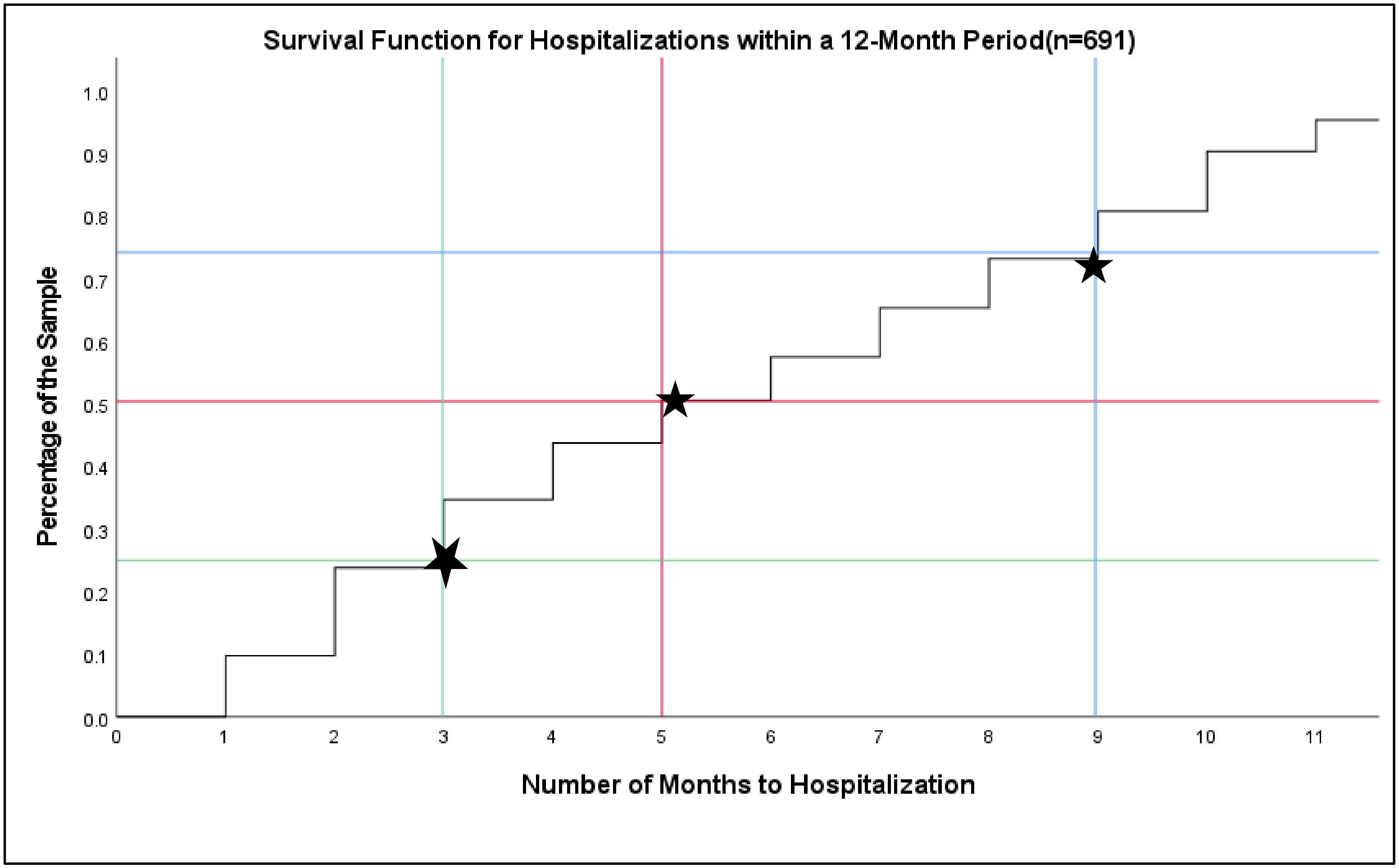


Figure 7 - Hazard Function for Hospitalizations from the Sub-sample (n=691)

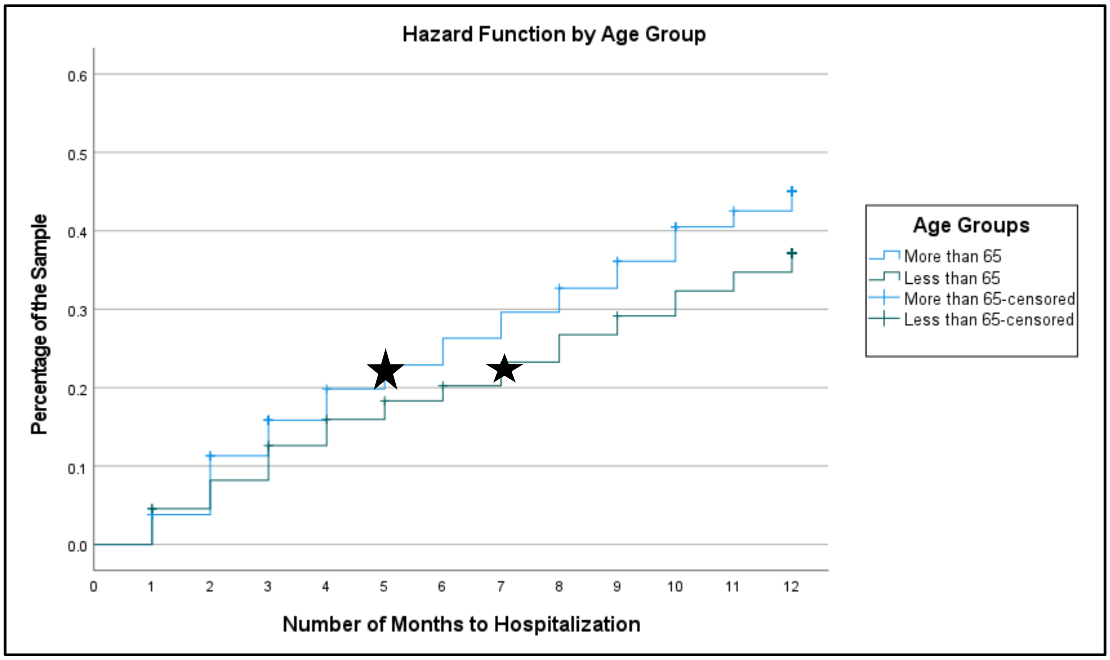


Figure 8 – Hazard Function for Hospitalizations by Age Group

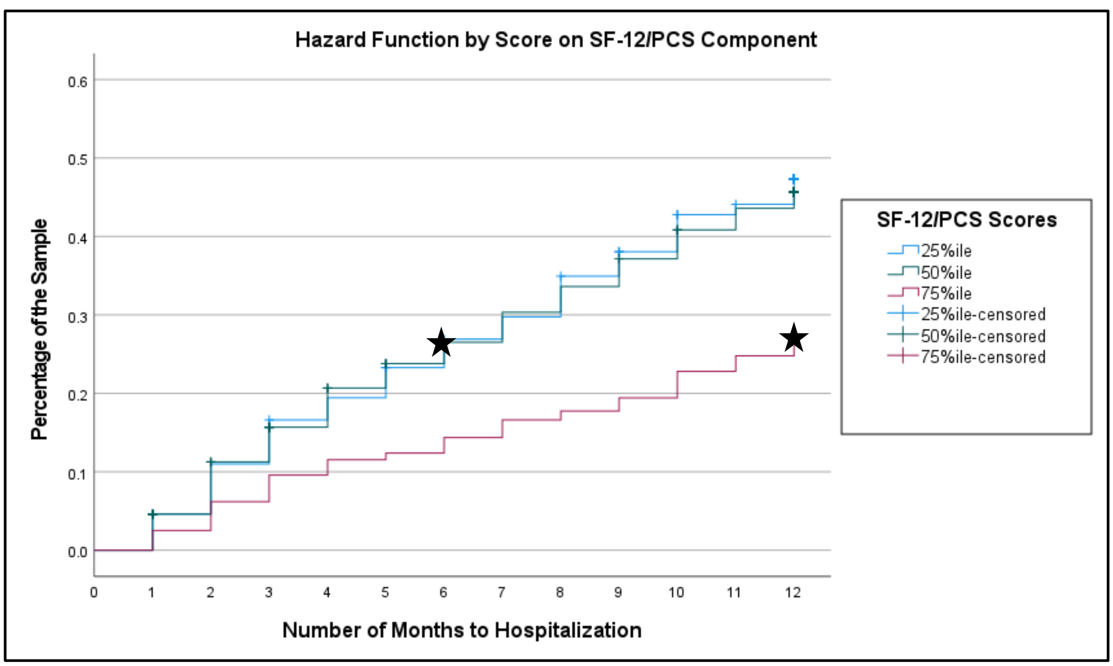


Figure 9 – Hazard Function for Hospitalizations by Score on SF-12/PCS Component

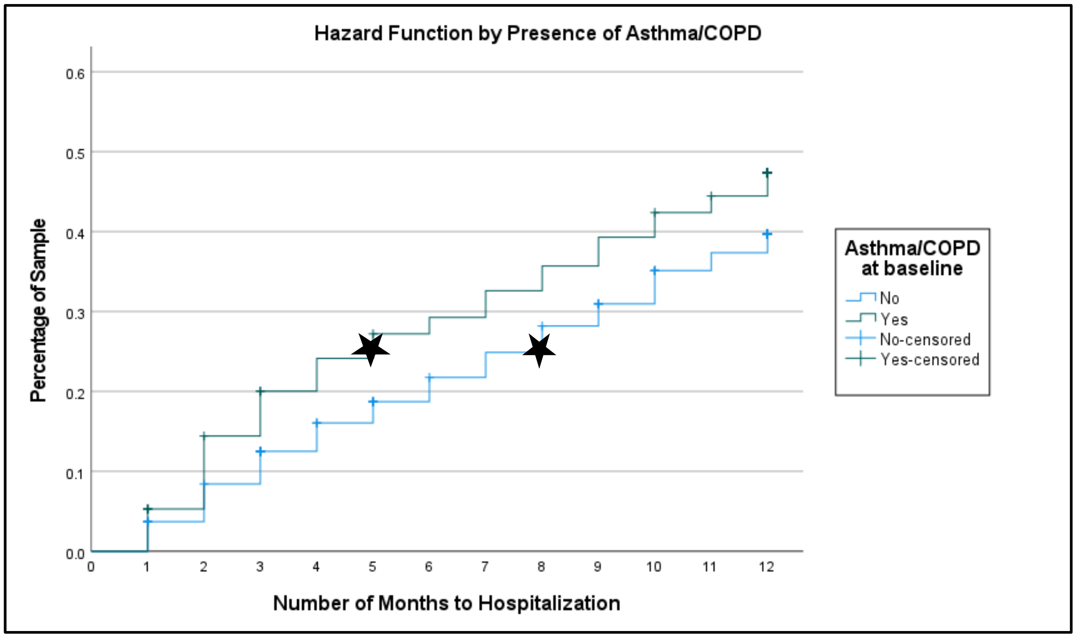


Figure 10 – Hazard Function for Hospitalizations by Presence of Asthma/COPD

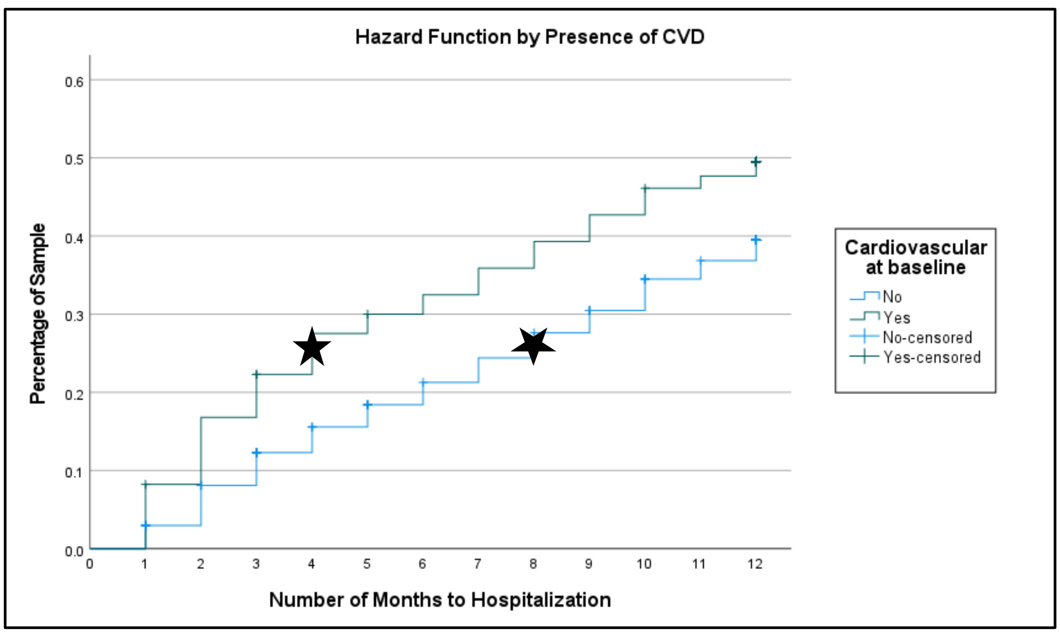


Figure 11 – Hazard Function by Presence of CVD

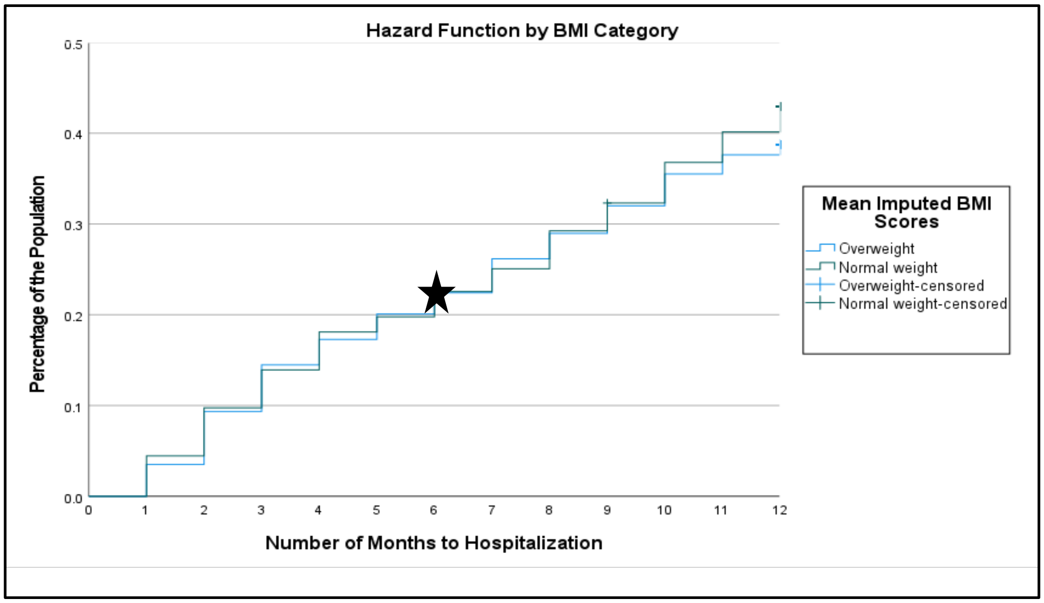


Figure 12 – Hazard Function for Hospitalizations by BMI Category

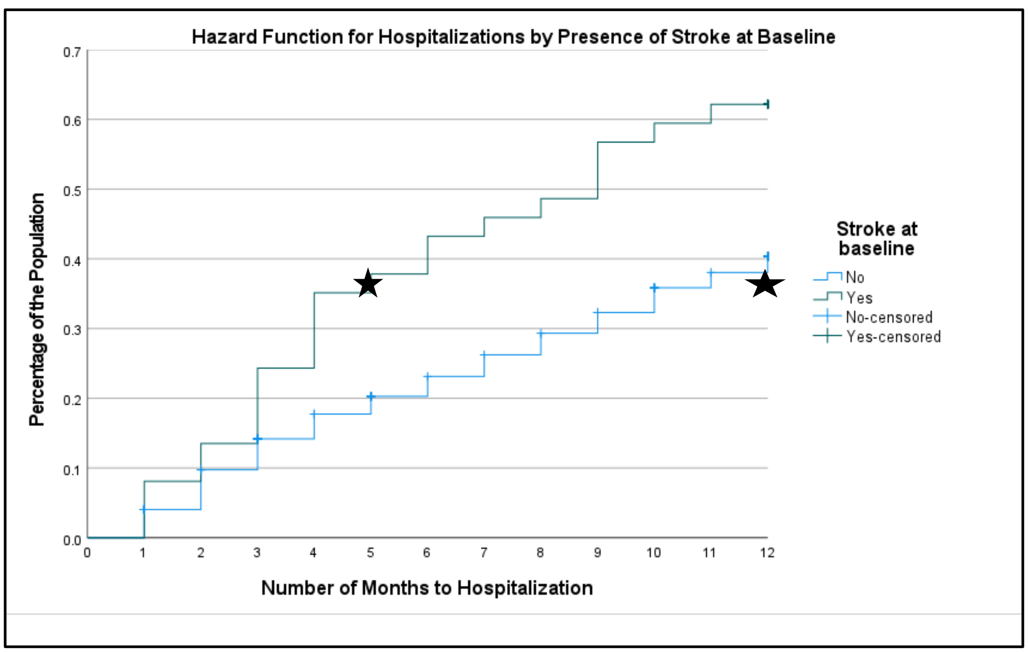


Figure 13– Hazard Function for Hospitalizations by Presence of Stroke

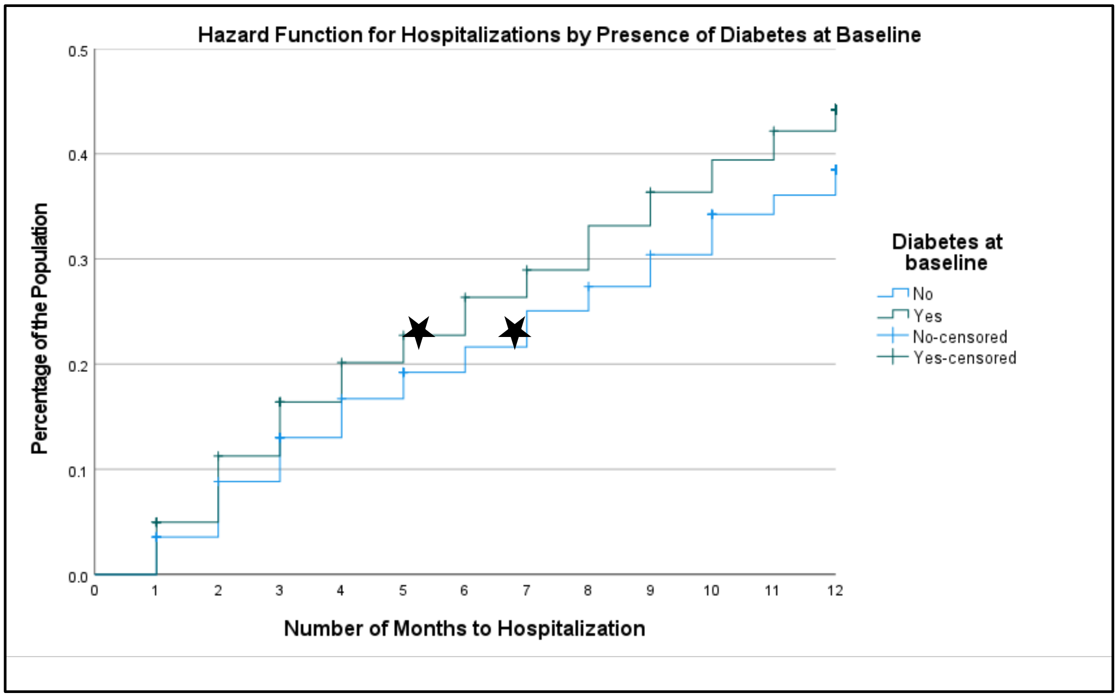


Figure 14 – Hazard Function for Hospitalizations by Presence of Diabetes

Table 10 - Comparison of Average Time to Event for Significant Predictors of Hospitalizations (in Months)

	Average Time to a Hospitalization (in Months)	
	Those with the Condition	Those without the Condition
Over 65 vs under 65	6	8
25th/50th percentile on SF-12/PCS vs 75th percentile	7	12+
Overweight vs normal BMI	6	6
Having asthma/COPD vs not having asthma/COPD	5	8
Having CVD vs not having CVD	4	8
Having diabetes vs not having diabetes	5	7
Having stroke vs not having stroke	4	10

Emergency Department Visits

The Cox proportional hazards model was also used to analyze the relative risk for emergency department visits within a twelve-month period from study enrollment. In the total study sample 41.8% (n=717) of participants experienced an emergency department visit event within the 12-month study period, while the remaining 59% (n=1001) of participants did not. Among the individuals (n=717) who experienced an emergency department visit within the study period, 12.6% experienced an event in the first month after enrollment and just over 33% experienced an event within the first three months after study enrollment. The average number of months to an emergency department visit was 5.5 (SD=3.25). The fewest number of emergency department visits (n=46, 6.6%) occurred in month 8. Like hospitalizations, about 4% of the sample (n=63) experienced three emergency department visits within the study period, while 25% (n=426) experienced only 1 visit. The overall model fit was significant (72.37, 32, 0.000). The delta change in the -2Log-Likelihood between the base model with no predictors (8818.22) and the final model with predictors (8740.39) was minimal indicating a good fit of the model that includes independent predictor variables as compared to the null model. Significant predictors from the Cox Proportional Hazards model are presented in Table 6 and are displayed as (HR, CI, p). Among independent predictor variables present in the model, significance was found for the following variables: scores on the SF-12/PCS component both in the 25th (HR=1.54, 95% CI=1.1.17-2.03, $p<0.002$) and 50th percentile (HR=1.46, 95% CI=1.14-1.86, $p<0.003$), overweight BMI (HR=0.77, 95% CI=0.62-0.97, $p<0.02$), current smokers (HR=1.33, 95% CI=1.06-1.67, $p<0.03$), those who were never married (HR=1.41, 95% CI=1.02-1.94, $p<0.04$), those with a Bachelor's degree (HR=0.65, 95% CI=0.43-0.99, $p<0.05$), those with Medicare (HR=1.35, 95% CI=1.05-1.74, $p<0.02$) and in individuals who reported baseline asthma/COPD (HR=1.43, 95% CI=1.21-1.68, $p<0.0001$), stroke (HR=1.75, 95% CI=1.15-2.66 $p<0.009$), depression (HR=1.39, 95% CI=1.13-1.71, $p<0.002$), and osteoporosis (HR=1.45, 95% CI=1.06-1.98, $p<0.02$). A graphic of the hazard function for emergency department use in the sample is displayed in Figure 7**.** According to the hazard function, 40% of the study sample experienced an emergency department visit within the twelve-months of the study duration. Based on the results, those with an increased risk for emergency room use and shorter time to event include individuals who were never married, were current smokers, were insured by Medicare, had poor physical functioning (SF-12/PCS), and reported baseline asthma/COPD, stroke, depression, and osteoporosis. Conversely, those with decreased risk and longer time to event include those with a bachelor's degree and overweight BMI. Individual Kaplan Meier curves were run for each of the significant predictors of emergency department visits and are presented in Figures 14-25. We found that on average, persons with lower education attainment (high school/GED) experienced an emergency department visit (M=5 months) at roughly two times the rate of persons with advanced educational attainment (M=10 months). Individuals who reported being never married (M=6 months) and those who reported being a current smoker (M=8 months) experienced an event nearly two months before those who reported being married (M=8 months) and non-smokers (M=10 months). Like the hospitalization outcome, persons who scored in the 25th or 50th percentile in the SF-12/PCS (M=6 months) experienced an emergency department visit on average, halfway through the study, while those in the 75th percentile experienced an event in or after the twelfth month of the study period. Persons with asthma/COPD (M=5 months) and stroke (M=4 months) at baseline experienced an emergency department visit roughly five months before individuals who did not have asthma/COPD (M=11 months) or stroke (M=10 months) at baseline. We also found that individuals who reported baseline depression (M=7 months) experienced an emergency department visits nearly three months before a person who did not report depression (M=10 months). Lastly, the current study found that individuals who were overweight at baseline (M=8 months) experienced an emergency department visits approximately one month after individuals of normal weight (M=7 months). Model comparisons for significance among the Kaplan Meier analyses including Log-Rank, Breslow and Tarone-Ware, are presented below as (chi-square, df, p). A significant in the tests indicates a notable difference between groups. The Kaplan Meier curve for individual scores on the SF-12/PCS component showed equivalent significance across comparisons including Log-Rank (11.81, df=2, $p<0.003$), Breslow (10.34, df=2, $p<0.006$), and Tarone-Ware (11.10, df=2, $p<0.004$). For individuals who reported the presence of stroke at baseline model comparisons include Log-Rank (9.89, df=1, $p<0.002$), Breslow (9.08, df=1, $p<0.003$), and Tarone-Ware (9.51, df=1, $p<0.002$). For individuals who reported asthma/COPD at baseline model comparisons include Log-Rank (7.94, df=1, $p<0.005$), Breslow (9.19, df=1, $p<0.002$), and Tarone-Ware (8.61, df=1, $p<0.003$).

Table 11 - Hazards Ratios by Independent Predictor Variable for Emergency Department Visits			
Main Variable	**HR**	**95%CI**	**p**
Emergency Department Visits			
Age			
>65	1.27	0.99-1.63	0.06*
Gender			
Female	1.05	0.88-1.26	0.57
Race			
Asian/Pacific Islander	1.06	0.77-1.45	0.73
African American	0.99	0.83-1.17	0.89
Hispanic	0.92	0.69-1.24	0.58
Education			
Less than 8th grade	0.90	0.70-1.17	0.45
Less than high school	0.96	0.76-1.20	0.70
Graduate degree	0.91	0.60-1.39	0.67
Some college	0.98	0.78-1.21	0.82
Bachelor's degree	0.65	0.43-0.99	0.05*
Employment			
Not working d/t disability	0.90	0.67-1.21	0.48
Retired	0.77	0.57-1.04	0.09
Unemployed	0.94	0.70-1.28	0.78

Main Variable	HR	95%CI	p
Marital Status			
Never Married	1.41	1.02-1.94	0.04*
Widowed/er	1.12	0.91-1.38	0.30
Divorced	0.94	0.75-1.17	0.56
Insurance			
Private	0.97	0.78-1.21	0.78
Tricare	0.84	0.50-1.41	0.52
Uninsured	1.05	0.73-1.50	0.81
Medicaid	1.35	1.05-1.74	0.02*
Family Income			
$0	1.21	0.91-1.61	0.18
$10000-$50000	1.25	0.87-1.79	0.23
$50000-$100000	1.19	0.91-1.57	0.21
>100000	0.64	0.38-1.07	0.09
Current Smoker			
Current Smoker	1.33	1.06-1.67	0.02*
Cognitive Impairment			
Presence of Impairment	1.02	0.85-1.21	0.81

Main Variable	**HR**	**95%CI**	**p**
Physical Functioning (SF-12/PCS)			
25%ile (M=23.79)	1.54	1.17-2.03	p<0.002*
50%ile (M=31.72)	1.46	1.14-1.86	p<0.003*
Mental Functioning (SF-12/MCS)			
25%ile (M=39.2)	1.21	0.95-1.54	0.13
50%ile (M=56.58)	1.08	0.88-1.33	0.47
Comorbidity			
More than 7 comorbidities	1.36	0.94-1.97	0.10
4-6 comorbidities	1.23	0.92-1.66	0.17
2-3 comorbidities	0.95	0.70-1.28	0.74
Depression	1.39	1.13-1.71	p<0.002*
Asthma/COPD	1.45	1.06-1.98	P<0.000*
Cardiovascular Disease	1.16	0.98-1.38	0.09
Stroke	1.75	1.15-2.66	P<0.01*
Stomach/GI	0.97	0.79-1.20	0.79
Colon Cancer	1.30	0.95-1.79	0.10
Diabetes	1.12	0.96-1.31	0.14
Kidney Disease	1.32	0.70-2.50	0.39

Main Variable	HR	95%CI	p
Muscular Disease	1.04	0.87-1.25	0.66
Arthritis	0.90	0.75-1.08	0.26
Hypertension	1.07	0.90-1.28	0.45
Osteoporosis	1.45	1.06-1.98	0.02*
Thyroid Disease	0.98	0.79-1.21	0.83
Cancer	1.03	0.83-1.28	0.76
Urinary Disease	1.02	0.80-1.30	0.87
Dementia	1.22	0.76-2.00	0.41
Hyperlipidemia	0.93	0.79-1.09	0.36
Body Mass Index (BMI)			
Underweight <18%	0.93	0.53-1.62	0.79
Overweight >25%	0.77	0.62-0.97	0.03*
Obese >30%	0.87	0.71-1.07	0.20

Note: Asterisk (*) Indicates Significance

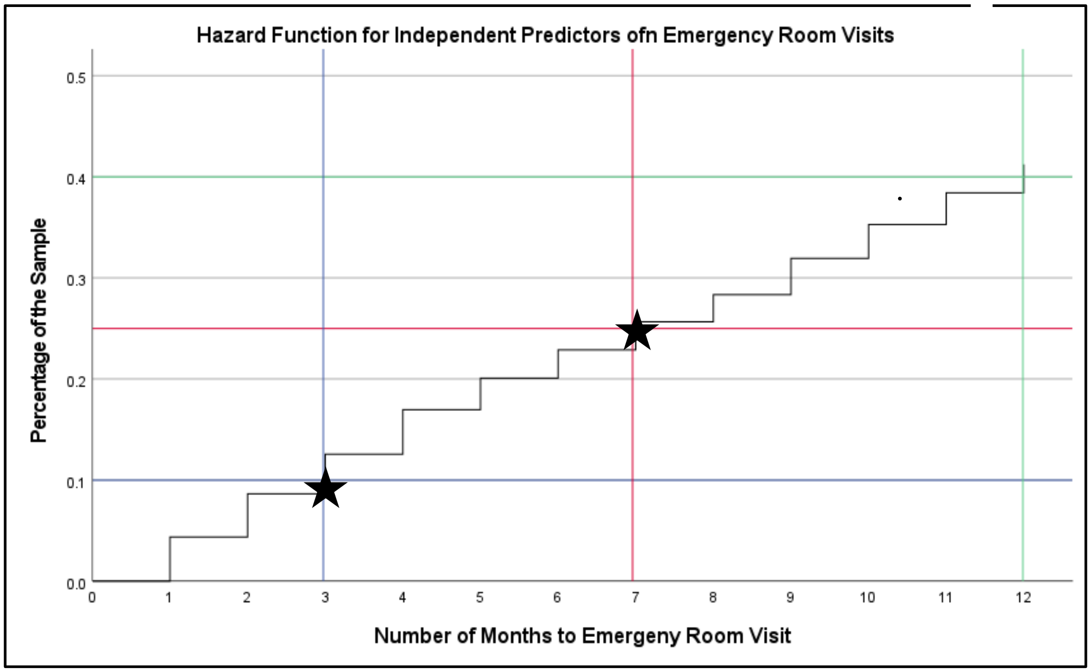


Figure 15-Hazard Function for Emergency Department Visits in the Sample (N=1714)

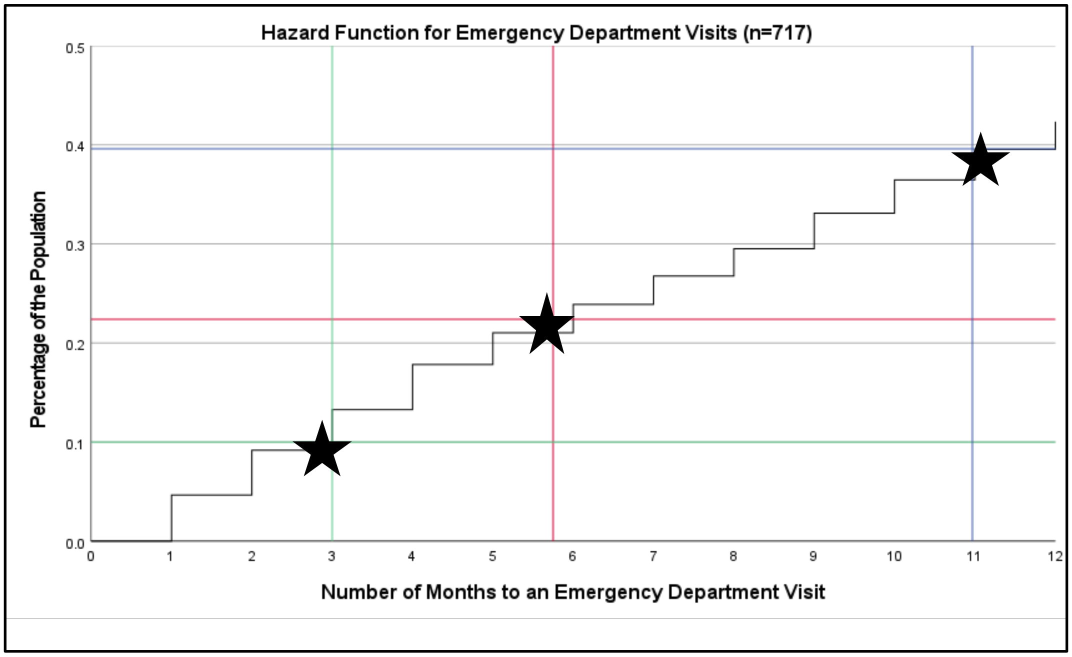


Figure 16-Survival Function for Emergency Department Visits in the Sub-sample (N=717)

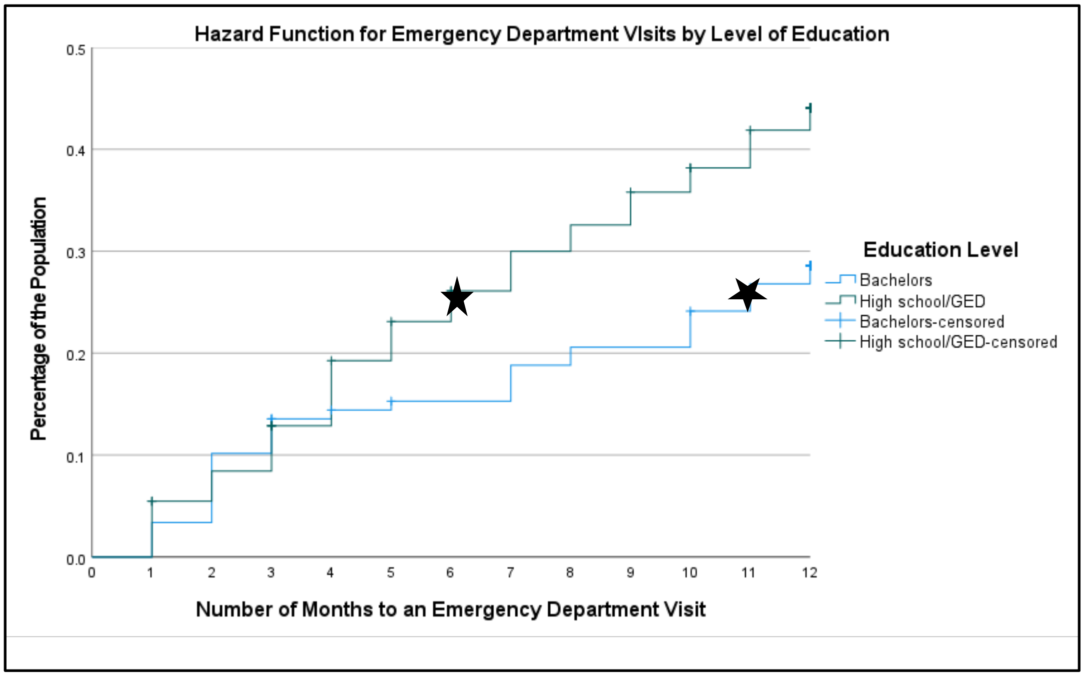


Figure 17 - Hazard Function for Emergency Department Visit by Level of Education

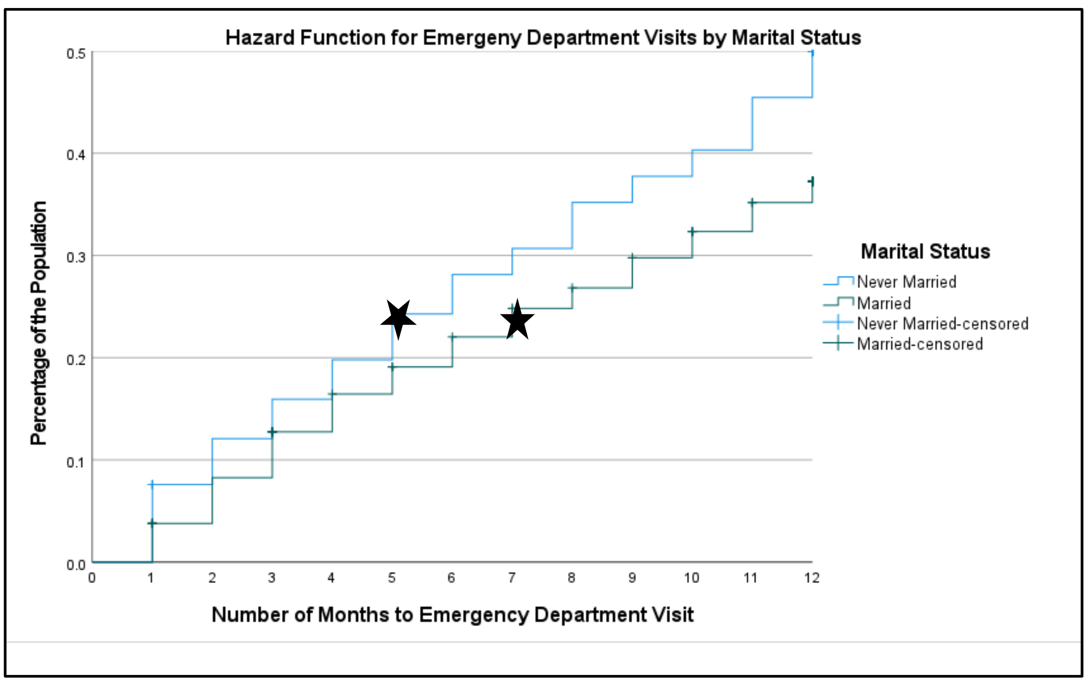


Figure 18 - Hazard Function for Emergency Department Visit by Marital Status

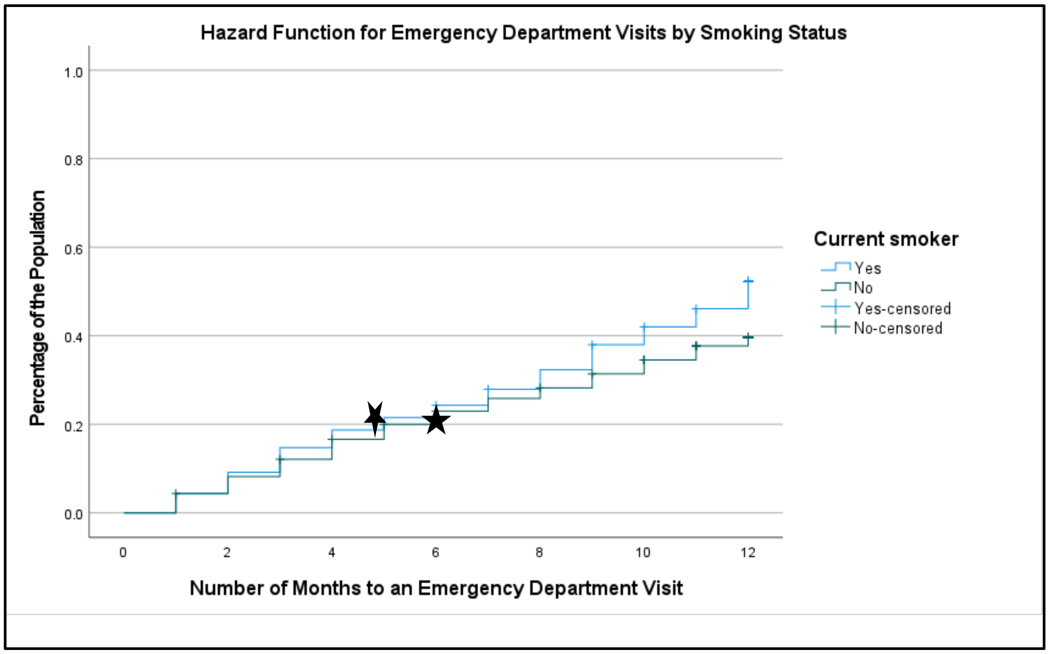


Figure 19 – Hazard Function for Emergency Department Visits by Smoking Status

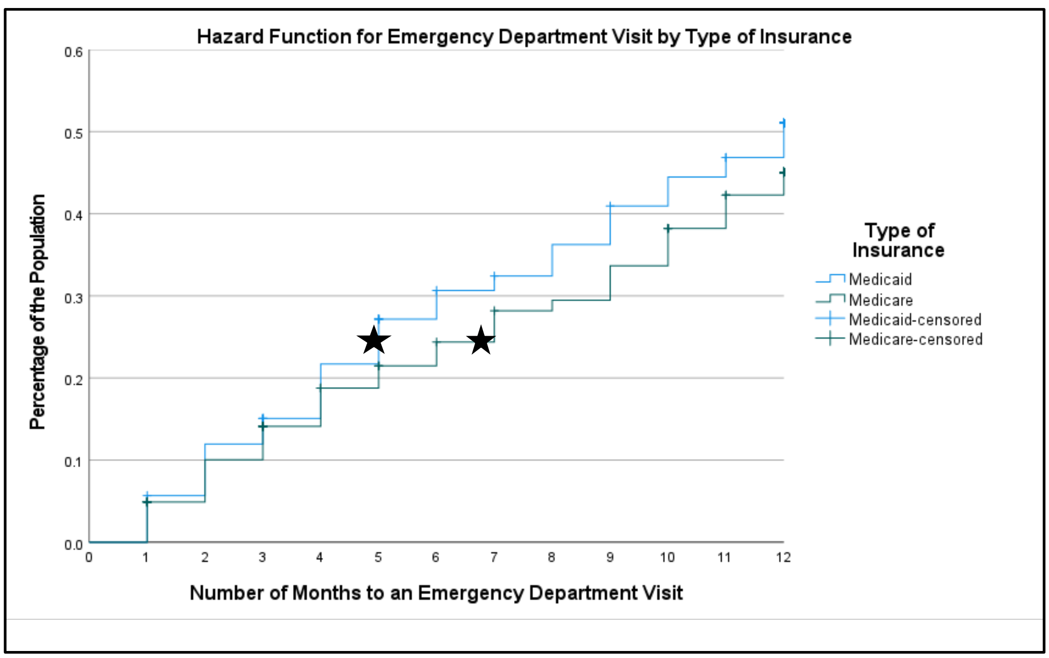


Figure 20 – Hazard Function for Emergency Department Visit by Type of Insurance

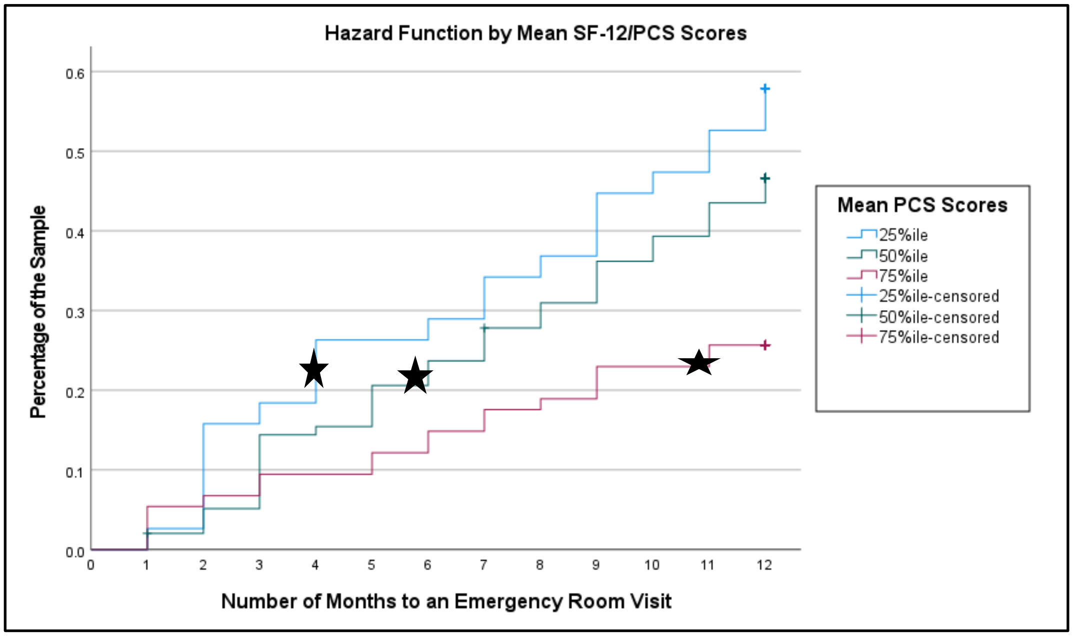


Figure 21 – Hazard Function by Mean SF-12/PCS Score for Emergency Department Visit

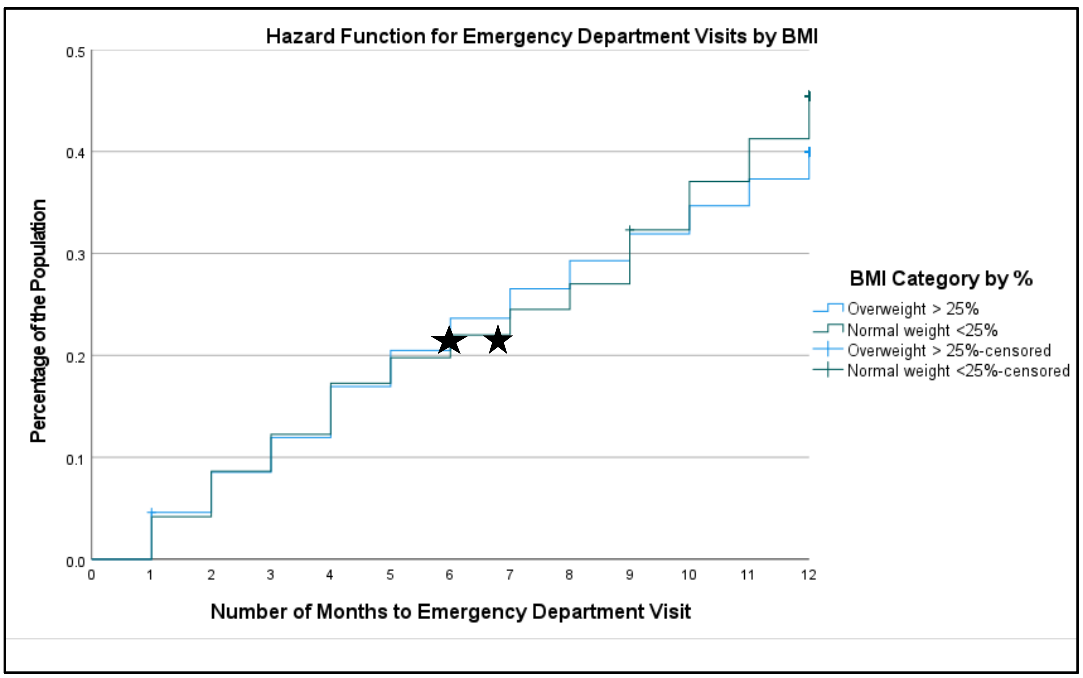


Figure 22 – Hazard Function for Emergency Department Visit by BMI

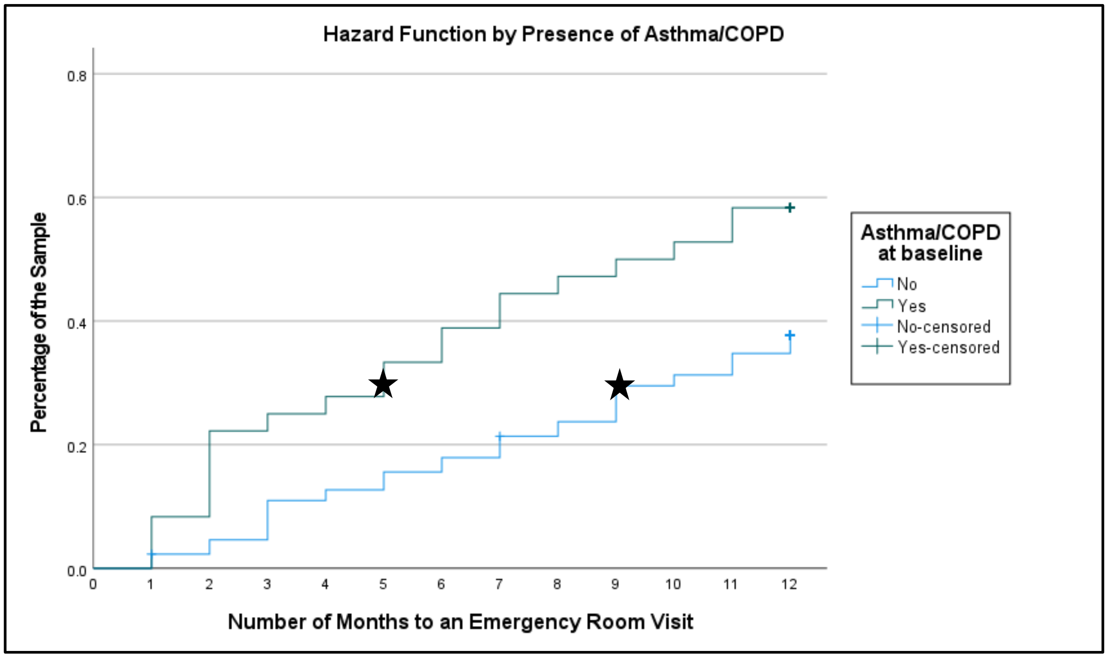


Figure 23 – Hazard Function for Emergency Department Visits by Presence of Asthma/COPD

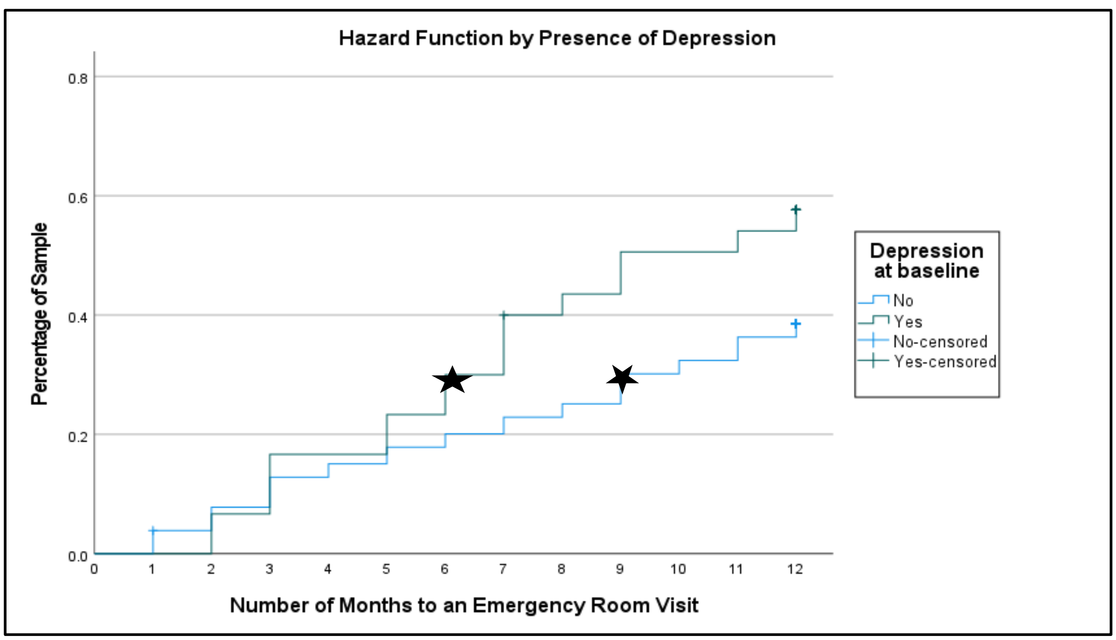


Figure 24 – Hazard Function for Emergency Department Visits by Presence of Depression

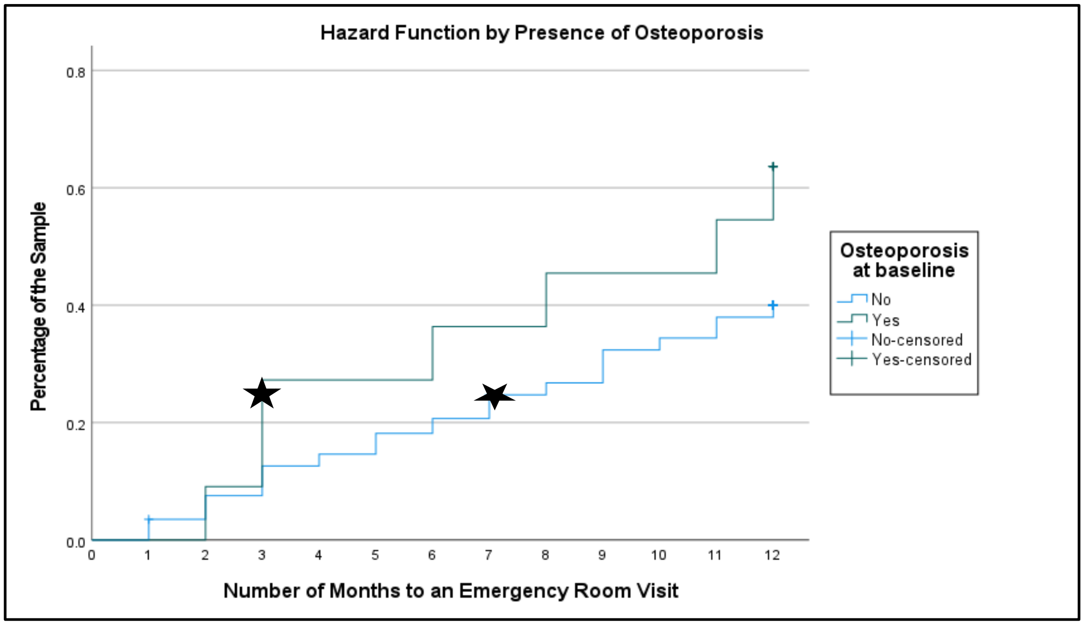


Figure 25 – Hazard Function for Emergency Department Visits by Presence of Osteoporosis

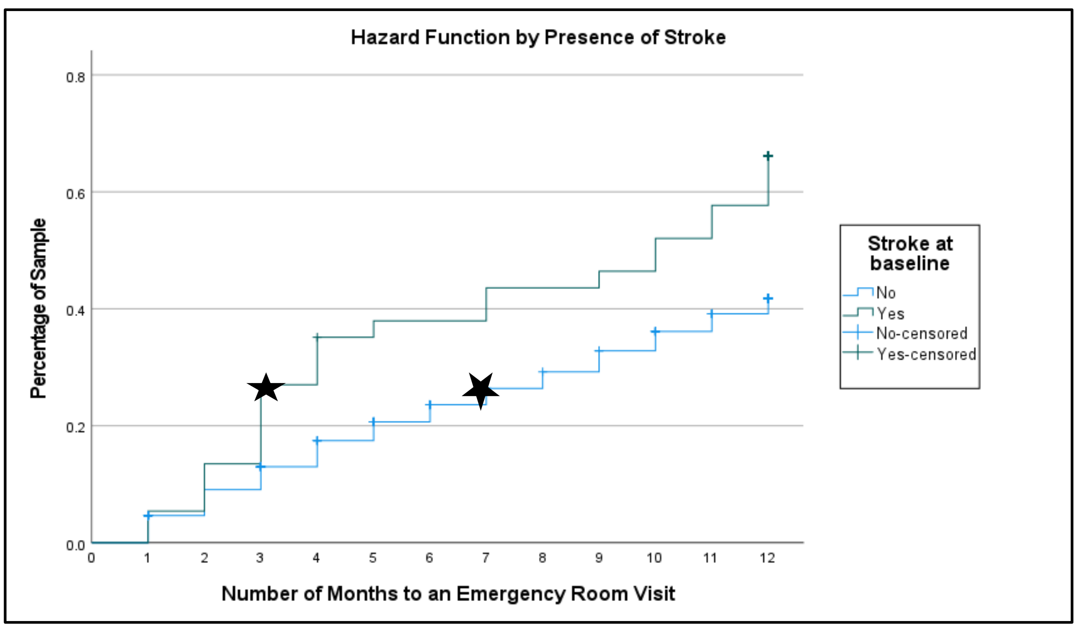


Figure 26– Hazard Function by Presence of Stroke for Emergency Department Visits

Table 12 – Comparison of Average Time to Event for Predictors of Emergency

	Average Time to Emergency Department Visit (in Months)	
	Those with the condition	Those without the condition
High school/GED vs bachelor's degree	5	10
Never married vs married	6	8
Current smoker vs non-smoker	8	10
Medicaid vs Medicare	5	7
25th/50th percentile on SF-12/PCS vs 75th percentile	6	12+
Overweight vs normal BMI	7	8
Having asthma/COPD vs not having asthma/COPD	5	11
Having depression vs not having depression	7	10
Having osteoporosis vs not having osteoporosis	6	10
Having stroke vs not having stroke	4	10

Department Visits (in Months)

Chapter V

Discussion

The purpose of the study was to (1) identify the strength of the relationship between the predisposing, enabling, perceived need, and environmental predictors and outcomes of health service use; including hospitalization, 30-day readmission and emergency room visits in persons with HF, (2) build a model to predict health service use in the HF population, and (3) determine the time to event and likelihood of health service use for persons with HF. The following chapter presents an interpretation of study findings including similarities and differences related to prior research, as well as study limitations, nursing implications, and recommendations for future research.

Study Sample

The current study used a national sub-sample of persons living in the community with HF. This type of sample differs from much of the current HF research and literature, as it does not include clinical or inpatient data such as laboratory values, NYHA class, cardiac imaging, or medications that are frequently monitored during an acute hospital stay. Instead, the sub-sample data are focused on socioeconomic characteristics of persons with HF that may influence an individual's predisposition to utilize urgent health services. Such socioeconomic data provides a unique perspective into the non-clinical characteristics of persons with HF that are often understudied and misunderstood. Of the models that do utilize only socioeconomic data, many are focused on specific groups or cultures (Ponce et al., 2018 and Tsuchihashi et al., 2001 and Mahajan et al. 2018), and may not accurately reflect the whole population of persons with HF. Among the HF literature, many of the studies attempting to create predictive models are built using samples of inpatient medical data (Awan et al., 2019 and Shameer et al., 2018) rather than community-based socioeconomic data, and tend to focus on 30-day hospital readmissions as an outcome rather than hospitalizations or emergency department use. The findings from the current study both clarify and reinforce findings from prior works, and introduces new information to the literature regarding emergency department use and hospitalizations in persons with HF.

Predictors of Health Service Use

For the purposes of the current study, a hospitalization was defined as any initial admission to a hospital within the 12-month study period, not within 30-days of a previous hospitalization and lasting at least 24 hours. A 30-day readmission was defined as any unplanned admission to the hospital, for any reason, within 30 days of a prior hospitalization and discharge. It is important to note that the reason for hospitalization, emergency department visit, or 30-day readmission was not included in the study data, simply whether an individual with a HF diagnosis experienced at least one hospital admission, visit to the emergency department, or 30-day readmission. The current study identified significant predictors of both hospitalizations and emergency department visits to include those with higher age, lower education level, poorer physical functioning, and the presence of baseline asthma/COPD, CVD, and kidney disease. Similarities in predictor significance between health service use outcomes may be related to the high number of emergency department visits nationally that are coded as HF-related, which frequently result in a direct hospital admission (Montoy et al, 2019), reaching upwards of 80% in some studies (Storrow et al., 2014).

Age. We found that individuals over 65 are at increased risk for experiencing an unplanned hospitalization. This finding is consistent with other studies that reported an increased risk of hospitalization for seniors with HF (Hamner et al., 2005). It has been well established in the medical literature (Corbretti et al., 2017 and Hawkins et al., 2015) that elderly patients, especially those with HF often experience health complications related to disease complexity, multi-morbidity, and frailty.

Gender and Race The study variables of gender and race were not found to be significant predictors of hospitalization or emergency department use in our study sample, however those who reported being of Asian descent were more likely to experience a 30-day readmission. This finding differs from some of the prior literature which states that males (Eapen et al., 2015 ad Ahmad et al., 2018 and Go et al., 2019) and African Americans (Blecker et al., 2018) are more likely to use health services. Additional studies found that women (Mortazavi et al., 2016), specifically Hispanic women (Pounce et al., 2018) are at greatest risk of health service use. Inconsistencies in the significance of gender and race in predictive models may be related to additional factors that are not considered or used in these models. It is possible that the influence of gender roles or cultural norms are often immeasurable and therefore are not included in predictive models. For example, one study (Heo et al., 2017) found that HF outcomes between men and women are related to differences in confidence, disease knowledge and functional status; factors which are not included in HF predictive models.

Education The current study also examined the influence of an individual's level of education on health service use, and found that those with a formal, upper-level degree are less likely to experience a hospitalization or emergency department visit than those with a high-school level education or less. This finding is consistent with other studies (Goals et al., 2018 and Eapen et al., 2015) and may suggest that individuals with lower educational attainment are less familiar with HF self-care or medical concepts that may prevent the unnecessary use of health services. The impact of education on health service use is especially important for persons with HF over the age of 65, as much of this population does not possess an advanced degree. Lower educational attainment is further exacerbated by the presence of cognitive impairment in persons with HF often related to disease progression and etiology. The current study did not find the presence of cognitive impairment to be a significant predictor of any health service use outcome, however cognitive impairment is present in nearly 25% of persons with HF nationally and has been found to be a significant predictor of 30-day readmissions in prior studies (Agarwal et al., 2016). The presence of cognitive impairment in addition to advanced age and low educational attainment likely hinders an individual's ability to appropriately remember complex care regimens, resulting in unplanned health service use secondary to mismanaged HF symptoms.

Comorbidity It is well known that persons with HF often suffer from multimorbidity, adding to the complexity of their disease management. The current study found multimorbidity at study enrollment, specifically having more than seven comorbid conditions, to be a significant predictor of 30-day readmissions but not hospitalizations or emergency department use. Additionally, we found the presence of specific baseline comorbidities including asthma/COPD and cardiovascular disease to be significant predictors of hospitalizations and emergency department use but not 30-day readmissions. The significance of specific conditions and multimorbidity present at baseline is consistent with other studies that identified kidney disease (Keenen et al., 2008), COPD (Au et al., 2012) and depression (Freeland et al., 2016 and Awan et a., 2019) as well as multimorbidity (Posada et al., 2019 and Hamner et al., 2004) as predictors of health service use in persons with HF. In the current study, depression was found to be a significant predictor of only emergency department use. This finding differs from prior literature which identifies depression as a predictor in 30-day readmission models (Johnson et al., 2012). Persons with depression often struggle with daily tasks including self-care management (Hawkins et al., 2016) which may lead to an increased tendency to use urgent health services for symptom management. It is important to note that the current study utilized the nineteen most common comorbid conditions as defined by the Canadian Community-Based Primary Healthcare Signature Initiative (CBPHSI) as variables in the predictive models. The use of the CBPHSI as a part of predictive models, rather than the Charlson Comorbidity Index, is relatively new in the HF literature (Fortin, Almirall, and Nicholson, 2017) and provides a unique and different perspective on the influence of comorbidity on health service use in a community-based HF population. The CBPHSI is focused on the self-reported nature of comorbid conditions often seen in administrative databases, such as the MEPS. Because the Charlson Index (Charlson et al, 1987) was originally developed for use in mortality risk prediction, it may not accurately reflect the socioeconomic aspects of health service use risk prediction from a community setting. As such, the use of the CBPHSI will likely benefit HF research as it will enable scientists to more accurately develop models which capture community-based aspects of common comorbidities that could lead to increased health service use in persons with HF.

Physical Functioning Physical functioning was also found to be a significant predictor of both hospitalizations and emergency department use in persons with HF. Notably, across all research question outcomes, individuals who scored in the 25th or 50th percentiles on the assessment were at least 1.5 times more likely to experience a hospitalization or emergency department visit. These findings were congruent with similar studies (Kitamura et al., 2019 and Yamada et al., 2012) which reported as much as a two-fold increased risk for health service use in persons with decreased physical ability. Physical mobility in persons with HF is further complicated by disease etiology which frequently results in edema, shortness of breath, and fatigue. As may be expected, lower physical mobility, specifically in the elderly, may impact a person's ability to manage the complex daily HF self-care regimen.

Marital Status Martial status was found to be a significant predictor of both emergency department use and 30-day readmissions. Specifically, persons who reported being divorced were more likely to experience a 30-day readmission, while those who reported never being married were more likely to experience an emergency room visit. Though marital status has not routinely been included in predictive models, this finding is supported by a few studies (Bradford et al., 2017 and Lu et al., 2016 and Esquivel and Spicer, 2012) which reported non-married individuals were more likely to experience a 30-day readmission. Whether an individual reported bring divorced or never married, it is likely that they lack the presence of a support person or caregiver, resulting in the need to utilize urgent health services for decisions, support and medical care. This is further supported by a meta-analysis of social factors in persons with HF which reports that those who lack social support and are unmarried are more prone to experience readmissions (Calvillo-King et al., 2013).

Insurance Status We found that persons who had Medicaid insurance were more likely to use the emergency department, but not experience an unplanned 30-day readmission. The finding of increased emergency department use is supported by other studies (Montoy et al., 2019 and Griswold et al., 2005) which report Medicaid users being more than 1.5 times more likely to experience an emergency department visit. However, the current study findings differ from prior literature regarding the significance of Medicaid in 30-day predictive models (Chen et al., 2019) with some studies reporting as many as 75% of patients (Jiang et al., 2019 and Hamner et al., 2005) being users of Medicaid during a 30-day readmission. In fact, many 30-day predictive models in the HF literature use Medicaid databases to obtain patient data. The discrepancy in results may be due to the low number of readmissions (n=203) as compared to hospitalizations (n=691) and emergency department visits (n=713) in the study sample. Though in general, persons with Medicaid are frequent users of public health services and should be monitored as a high-risk group.

Hospital LOS The current study found hospital LOS to be a significant predictor of 30-day readmissions in persons with HF. Specifically, any length of time greater than one week was associated with increased likelihood of 30-day readmission. Interestingly, persons with a hospital LOS of more than three weeks, but less than one month, had the highest likelihood (OR=5.2) of readmission. These results are supported by other studies (Au et al, 2012 and Ashfaq et al., 2019) which found that a hospital stay of longer than 11 days to be associated with unplanned 30-day readmission. Another study supported our findings, reporting a 17% increased risk for unplanned 30-day readmission for a hospital LOS between 5 and 10 days (Reynolds et al., 2015). As such, it is important for clinicians to monitor medical progress in acute HF patients and attempt to reduce hospital LOS when medically possible.

Income The current study found that persons with low income, specifically those who reported income between \$10000-\$50000 per year and those who reported \$0 per year to be at increased risk for 30-day readmissions. These findings are congruent with other studies (Lindauer et al., 2013 and Patil et al., 2019 and Chamberlin et al., 2018) which suggest that persons from low-income quartiles are more likely to experience negative outcome including 30-day readmissions. This may be because persons with low income are unable to afford important medications or low-sodium foods that help to reduce HF-related complications. Additionally, it may be possible that persons with low income, but not receiving Medicaid, cannot afford health insurance and therefore do not obtain routine primary care that may prevent unnecessary urgent health service use.

BMI The current study found that when compared to individuals of normal weight, persons who reported being obese (BMI >30%) were less likely to experience a hospitalization and those who were overweight (BMI >25%) less likely to experience an emergency department visit. These findings are not supported by much of the recent literature (Cox et al., 2017 and Przybylowicz et al., 2020 and Mandviwala et al., 2020) which report that obesity is a predictor of health service use in persons with HF. The current study findings may be supported in part by the “Obesity Paradox” (Horwich et al., 2001) which states that persons with established chronic HF who are underweight can experience worse outcomes than those who are overweight (Curtis et al., 2005 and Zadeh et al., 2005 and Sharma et al., 2015). According to the literature, this is likely due to the protective effect of classified obesity as related to body composition (Horwich et al., 2018) that may lead to improved outcomes for individuals with HF who are overweight. The improved outcomes for those who are overweight may be attributed to the inability of a frail individual to meet the increased demands of HF on the body, often caused by increased work of breathing and cardiac demand. As a result, individuals who are frail may be more likely to require medical intervention to manage the symptoms of HF. In fact, a recent national HF study found that those in a high-risk frailty group experienced longer hospital LOS (11 days) as compared to those in a low-risk frailty group (4.6 days) respectively (Kwok et al., 2020). Because increased hospital LOS is known to predict rehospitalization, it may be possible that frail individuals with low BMI are truly at an increased risk for health service use. However, the obesity paradox in HF is generally related to mortality risk rather than hospitalization risk and may not fully explain the current study findings. Additionally, further research is needed to explore the possible relationship among BMI, frailty and health service use in persons with HF.

Smoking Status We found that persons who reported being current smokers were more likely to experience an emergency department visit than non-smokers. This finding is supported by another study (Sax et al., 2017) which reported 7% of persons with HF who were current smokers experienced an emergency room visit within one week of a prior emergency department discharge. It is important to note that in the current study, emergency department visits were measured as all-cause, making it difficult to determine if this visit was related to complications from smoking versus complications from HF. However, smoking is associated with negative health outcomes, especially in those with cardiac disease.

Table 13 - Comparison of Significant Odds Ratios Among Health Service Use Outcomes

	Hospitalizations		**30-Day Readmissions**		**Emergency Department Use**	
	Less Likely	More Likely	Less Likely	More Likely	Less Likely	More Likely
Age >65		(1.37)				(1.34)
Asian				(3.59)		
Graduate Degree	(0.48)					
Bachelor's Degree					(0.60)	
Medicaid						(1.49)
Never Married						(1.74)
Divorced				(1.58)		
Current smoker						(1.52)
SF-12/PCS		(1.70)				(1.60)
Obese BMI	(0.77)				(0.71)	
Underweight BMI			(0.3)			
Asthma/COPD		(1.26)				(1.52)
CVD		(1.50)				(1.31)
Kidney Disease		(3.60)				
Depression						(1.43)
Osteoporosis						(1.69)
>7 comorbidities				(3.98)		
Hospital LOS >1 month				(2.27)		
Hospital LOS < 1 week			(0.36)			
Hospital LOS >2weeks				(3.93)		
Hospital LOS >3weeks				(5.2)		
Income $0				(1.89)		
Income $10000-$50000				(1.89)		

Note: variables that were not significant predictors of health service use include gender, employment, cognitive impairment, mental health (SF-12/MCS), and the presence of baseline hypertension, musculoskeletal disorder, arthritis, stroke, stomach/GI disease, colon cancer, diabetes, thyroid disease, any cancer, urinary/renal disease and dementia.

Health Service Use Hazards and Risk Prediction in Persons with HF

The current study used a time to event analysis via Cox Proportional Hazards and Kaplan Meier Curves to determine both the relative risk of a hospitalization or emergency department visit among significant model predictors and the relative time to a hospitalization or emergency department visit for each predictor. A hazard, by definition, refers to the probability that, at any given time, an individual will experience an event (Fields, 2008). Of note, as hazard increases the survival time to an event decreases. Much of the current HF literature focuses on time to event analyses either for the 30-day readmissions outcome or for 1-year mortality. This is because of the changes to hospital reimbursements as defined by Medicare, which emphasizes a window of 30-days after hospital discharge for a non-reimbursable readmission. However, due to the nature of our dataset regarding the lack of an available initial hospitalization date, time to event analysis for the 30-day readmission outcome was not possible. Instead, the current study examined the time to event outcome for both hospitalizations and emergency department visits within one year following study enrollment. Because a 30-day readmission must follow an initial hospitalization, it is important for HF research to examine the time to an event (i.e., hospitalization) from the community setting without limitations on duration. This approach is supported by a recent study (Chen et al., 2017) which explains that 30 days may not be an appropriate interval to determine outcomes, rather time to event analyses should include an arbitrary time interval. While 30-day time to event models are necessary for Medicare reimbursement claims research, persons with HF experience frequent health service use outside of the 30-day window. To facilitate meaningful comparisons in time between the one-year window in the current study and results from prior literature, significant variables will be compared with studies that examined time to event hospitalization and emergency department outcomes of at least 60 days after discharge.

Hospitalizations A COX Proportional Hazards Model was used to determine time to event and relative risk for hospitalizations in persons with HF. There are limited time to event studies related to socioeconomic factors that influence hospitalizations from the community setting. Instead, much of the HF literature is focused on clinical factors that impact the time to a readmission following an initial hospitalization, typically within a 30-day window. Of note, time to event models of 30-day readmissions have found on average, that persons who experience a readmission do so within he first two weeks (mean day 11) after discharge (Fudim et al., 2017). The current study found that on average, persons experienced a hospitalization 5.3 months after study enrollment. We also found that individuals who were at increased risk for hospitalization including shorter time to event includes persons who had poor physical functioning (SF-12/PCS), were greater than 65 years of age, and reported baseline asthma/COPD, CVD, stroke, and diabetes. In addition, we found that those at decreased risk for hospitalization including longer time to event included individuals who were overweight and had a graduate degree. These findings can be supported by a recent time to event study of HF which found that elderly persons with HF and diabetes are at 5% increased risk of hospitalization over a five-year period (Williams et al., 2020). It is possible that the combined symptom burden of these two conditions facilitates poor disease management, especially in the elderly, leading to the increased risk of health service over time. Another study found that individuals with functional decline, specifically related to completion of

ADLs, are nearly 1.5 times more likely to experience a hospitalization (Takabayashi et al., 2019). As previously discussed, individuals with poor physical functioning likely are not able to maintain the daily demands of HF self-care resulting in symptom exacerbations that require acute medical attention. Overall, the literature related to time to event analyses of hospitalizations in HF focus on the acute period following the initial hospitalization. However, it may be possible that hospitalizations occurring outside of the typical 30-day window are influenced more by community-based factors, like those used in the current study, rather than inpatient or clinical factors. One study related to factors influencing early verses late hospitalizations (Vader et al., 2016) found that a hospitalization occurring after 60 days of a prior hospitalization is frequently related to causes other than HF.

Emergency Department Visits No studies have been conducted which specifically examine twelve-month time to event analyses of socioeconomic factors of persons living in the community with HF, and relative risk for emergency department visits. Instead, current literature focuses on inpatient clinical factors which lead to emergency department use after an initial hospitalization or prior emergency department visit. The current study found that from a community setting, the average time to an emergency department visit after study enrollment was 5.5 months. For reference, approximately 20% of persons with HF revisit the emergency department within two weeks after a previous emergency department visit and up to 40% visit the emergency department within 30-days after a hospital discharge (Claret et al., 2018). Due to the lack of similar time to event studies, findings from the current study will be compared to related studies of risk factors for emergency department use in persons with HF. The current study found that persons at an increased risk of emergency department use and shorter time to event were those greater than 65, with poor physical functioning, who were current smokers, never married, and reported baseline asthma/COPD, depression, and stroke. Further, we found that individuals who reported an overweight BMI and graduate degree had decreased risk for emergency department use and longer time to event. Our results are comparable to one study of inpatient medical management of HF which reported that advanced age (>80) was independently associated with "uncontrolled heart failure" leading to nearly 90% of the study population experiencing at least three emergency department visits within one year after the first acute hospitalization for newly diagnosed HF (Miro et al., 2019). This may translate to the community setting, whereby medical management through frequent primary care visits and provider follow up could medicate the effects of "uncontrolled HF' and reduce the number and frequency of emergency department visits, especially those over 65. Our significant findings regarding the presence of baseline disease (asthma/COPD, depression, and stroke) can be supported by a recently conducted time to event study on mortality in HF, which found that those who had multimorbidity often experienced three or more emergency department visits within 6 months, which was independently associated with shorter time to mortality (Posada et al., 2018). As can be expected, a high number of comorbid conditions adds complexity to disease management often resulting in the need to use the emergency department for acute or mismanaged symptoms. Further, disease mismanagement, specifically self-care, can be worsened by advanced age, poor physical functioning, and depressive symptoms, all determined to be risk factors of emergency department use in the current study. These findings can be compared to a study of self-care beliefs and emergency department use in HF, which found that those of a younger age had more accurate HF self-care beliefs and reduced emergency department use (Albert et al., 2014). One additional study found that persons with HF who have combined multimorbidity and functional limitation experience nearly a 3.5 times greater risk of emergency department use (Manemann et al., 2018). Our findings of decreased risk for emergency use for those who were overweight can be substantiated by a study of frailty and health service use in HF which found that frail individuals, characterized in-part by low BMI, had a 92% increased risk of emergency department use (McNallan et al., 2013). We found that individuals with higher educational attainment experienced decreased risk for emergency department use and longer time to event. While level of education was not found to be a factor in any time to event models related to emergency department use in HF, one study (Murray et al, 2009) found that an individual's literacy and reading skills were correlated with both increased self-care abilities and decreased emergency service use. Persons with HF experience high rates of emergency department use, especially after an initial hospitalization, however further research is needed that explores socioeconomic factors and their impact on time to event from the community setting. Because the average time to event in the current study was at least 5 months, it may be necessary to extend the window for future time to event research for an emergency department visit following a prior hospitalization or emergency department visit to at least 6 months, rather than 30-days. Extension of this time window may help clinicians to differentiate between inpatient factors and community-based factors that may influence the time between emergency department visits.

Clinical Implications to Reduce Health Service Use in Persons with HF

The current study examined the characteristics of a national sample of persons with HF groups and included three health service use outcomes. The advantage of the current study model, is that it allows us to compare predictors of various health service use outcomes to better understand commonalities and linkages among patterns and characteristics of persons with HF. The similarities and differences among the predictors for each outcome (Table 9) can help us to both understand and identify those who are at increased risk for hospitalization, emergency department visits, and 30-day readmissions, and implement strategies to address specific risk factors.

Emergency Department Visits and Hospitalizations There were several commonalities in predictors between emergency department use and hospitalizations. This is likely the result of the high number of emergency department visits that result in a subsequent hospitalization. Overall, those at increased risk of experiencing these health service use outcomes are individuals with HF who are older, have low educational attainment, poor physical functioning and report baseline asthma/COPD or CVD. Persons over the age of 65 have notably diminished self-care abilities that result in adverse health outcomes. The decreased self-care abilities of this aged population result in rapid health decline and the subsequent need for treatment and advanced therapies. As the percentage of persons over the age of 65 is expected to double by 2050 (Census Bureau, 2020) it can also be expected that the number of hospitalizations and emergency department visits would also increase. The proposed increase in health service use in an elderly HF population will place an overwhelming and unnecessary burden on the health care system. Initiatives such as the Age-Friendly Health System (AFHS) lead by the Institute for Healthcare Improvement (IHI) seeks to establish and maintain a healthcare system that caters to the needs of an elderly population. To achieve this goal, the IHI proposed that hospitals provide effective care to elderly patients by setting appropriate care outcomes, prescribing age-specific medications, monitoring mentation, and ensuring safe mobility (IHI, 2020). The increased complexity of HF management with age coupled with lower educational attainment places elderly persons with HF in an increased risk category. Initiatives aimed at reducing health service use by bridging the HF knowledge gap include the MyROAD card (Albert et al., 2020) and the HF Discharge Initiative (Albert et al., 2021). Both programs provide individuals with audio resources that detail guidelines for improving HF self-care including medication adherence, daily weighing, sodium restrictions and physical activity. While there was no impact on readmission rates, participants in the MyROAD program saw a 27% reduction in emergency department visits at 30 days post-discharge (Albert et al., 2020). For persons with HF over the age of 65, initiatives such as these could be imperative in aiding the reduction of unnecessary hospitalizations and emergency department visit. In addition, older persons with HF often suffer from multimorbidity, further complicating their treatment regimen. Specifically, persons with HF who also have asthma/COPD, CVD, kidney disease and depression are at greatest risk for hospitalizations and emergency department use. Management of these conditions during outpatient visits is essential in reducing complications that could worsen the HF trajectory and result in the need to use urgent health services. The significant socioeconomic characteristics identified in this study can be easily monitored and managed during interactions in community setting such as clinics, home-health visits, and outpatient visits. Activities such as performing a medication reconciliation, reinforcing healthy dietary habits, and providing education on coping skills should become a routine part of outpatient care. Screening tools for IQ, physical ability, and depression (Aloisi et al., 2019) should become a routine part of primary care assessment and care management for persons with HF.

If primary care nurses recognize characteristics of those at risk, early intervention would be possible, and may help to reduce likelihood of an unnecessary health service use. Examples of primary intervention for an at-risk individual with low educational attainment and poor physical functioning may include targeted teaching related to specific aspects of HF self-care, possibly through the use of a more able-bodied caregiver, including sodium restrictions and medication adherence as well as participation in physical therapy programs. It is essential that these interventions occur prior to the unnecessary use of health services, rather than during an emergency department visit or hospitalization. Of note, though depression was only found to be a significant predictor of emergency department use, it should be considered as a risk factor for all persons with HF, as there is a large body of literature (Patel et al., 2020 and Kewcharoen et al., 2020 and Xu et al., 2018) which has established a connection between depression and negative outcomes in persons with HF. There were three factors found to be significant predictors of emergency department use but not hospitalizations. These include individuals who report the use of Medicaid, being a current smoker, and never being married. Interestingly, in one study 16% of persons who received a new HF diagnosis continued smoking (Son and Lee, 2020). Smoking is known to result in negative health outcomes, especially in those with cardiac disease. It is possible for primary care nurses to influence unnecessary health service use in persons with HF who smoke by providing continued education and information on smoking cessation programs. Additionally, the primary care nurse can provide information on community support programs for persons who were never married and may not have a support person to help with self-care decisions.

30-Day Readmissions Interestingly, there was no overlap between the factors that predicted emergency department use and hospitalizations and the factors that predicted 30-day readmissions. In fact, most of the significant predictors of 30-day readmissions are non-modifiable including race, hospital LOS, marital status, and income. This suggests that there are significant clinical components, rather than socioeconomic factors, that arise during a hospitalization that could influence a readmission. It is likely that a community-based sample which lacks clinical data is not effective at predicting readmissions in persons with HF. In fact, one study reported the use of both clinical and socioeconomic data improved predictive models by up to 5% (Walsh and Hripsak, 2017). Additional studies focused on 30-day readmissions in persons with HF cite the importance of including prior hospitalization data to improve the efficiency and accuracy of predictive models (Chen et al., 2019 and Bradford et al., 2017). The current study found that those admitted for under one week were less likely to experience a readmission. This may be related to the nature of a less serious illness requiring a shorter hospital LOS, therefore resulting in fewer post-discharge complications. One recent study (Massari et al., 2019) found that hospital LOS (on average 9 days) is significantly influenced by degree of HF exacerbation, specifically related to the level of hydration, brain natriuretic peptide levels and blood urea nitrogen levels. Because a 30-day readmission must follow an initial hospitalization, it is possible that inpatient clinicians, as well as primary care providers, can directly influence clinical factors affecting readmissions. To this end, hospital nurses must participate in multidisciplinary rounds and remain attuned to care plans and advocate for shorter hospital LOS times when medically possible. In addition, thorough discharge teaching, post-discharge follow-up visits, and assessment of disease management during routine primary care visits may help to prevent unnecessary 30-day readmissions. Further we found that those with low income were more likely to be readmitted, likely due to the financial ability to afford medically necessary resources such as medications and low-sodium foods. Both inpatient and primary care nurses can identify these individuals and initiate communications between patients and social workers, targeted at social resource programs such as WIC and unemployment. Further, the nature of the predictors of 30-day readmissions found in the current study would suggest that the 30-day readmission outcome does not align with the community-based model used in this study design. Rather, a study model focused on inpatient characteristics may be more effective and accurate at predicting unplanned readmissions.

Machine Learning Model Outcomes

Research Question 2 was answered using an exploratory modeling approach to create and compare various machine learning-based models for each health service use outcome. The goal of this question was to determine model efficacy and accuracy and compare their performance to previously developed models in the HF literature. Four models were created for each health service use outcome, and their performance was evaluated using the AUC, sensitivity, specificity, positive and negative predictive values. It is important to consider that in the HF literature, model performance was mainly reported using the AUC statistic. In addition, due to Medicaid's Hospitalization Reduction Program (HHRP) guidelines for hospital reimbursements, the current HF literature has focused mainly on predictive models for the outcome of 30-day readmissions after an inpatient hospitalization. This study is unique in that it not only explores predictive models for 30-day readmissions, but also initial hospitalizations and emergency department use for persons living in the community with HF. As such, the findings from this study both add to the current literature on predictive modeling for 30-day readmissions and brings new information to the field regarding the performance of predictive models for the other health service use outcomes. The use of simple data analytic techniques such as logistic regression, rather than complex data science techniques, has been heavily debated in the HF literature. Findings from other studies (Keenan et al., 2008 and Evans et al., 2016) support the use of traditional logistic regression to develop an effective predictive model rather than the use of complex data analytic techniques. For this reason, one of the four predictive models tested for each outcome was a simple logistic regression. The predictive models using logistic regression performed as well, or better than at least one other type of model (Random Forest, Bayesian, Neural Networks) for all the health service use outcomes that were tested (Table 8). This suggests that when building predictive models, a simplistic approach may be the most effective. Because the factors that predicted each health service use outcome differed, it is likely that one model may not be effective at predicting risk for all three health service use outcomes simultaneously, rather, individual models and variables are necessary for each health service use outcome. Further, due to the differences in factors from inpatient versus community samples of persons with HF, separate predictive models should be developed which assess the risk of health service use in each population. The model performance for each health service use outcome are discussed below.

Emergency Department Use. Though on average, the four machine learning models for emergency department use from the current study performed only modestly (AUC 0.60), the percent variance explained by each of the variables were consistent across models. Considering all criteria (Table 5) for evaluating model performance, the Random Forest model for emergency department use performed the best, however, was only able to correctly classify individuals 70% of the time, or less. The variables of education, insurance, depression, asthma/COPD, and SF-12/PCS scores individually accounted for no less than 10% of the variance in each of the various emergency department predictive models. However, the percent variance explained by the significant predictors in each model totaled no less than 70 percent. In other words, approximately 30% of the model's composition cannot be explained by the variables included in the study. Though the percent variance explained by the models was quite high, their overall performance based on sensitivity, specificity, and positive/negative predictive values was quite poor. In fact, on average the four models developed for emergency room use in the current study were only able to correctly classify persons roughly 50% of the time. The poor model performance is likely attributed to the large amount of unexplained variance. As previously mentioned, the complex nature of a HF diagnosis along with the ambiguity of socioeconomic characteristics leave a great deal of uncertainty in predictive models. The same ambiguity in predictive models is seen in other studies (Chen et al., 2017 and Allam et al., 2019 and Hirsh and Hripcsak, 2014) which report suboptimal model performance. While no models were identified in the HF literature that predict emergency department use in persons with HF from the community setting, similar studies such as the STRATIFY decision tool (Sax et al., 2021 and Colins et al., 2015), which examines adverse events after an emergency department discharge experienced similar model performance (AUC 0.70).

Hospitalizations The machine learning models for hospitalizations in the current study performed comparably to the models for emergency department use (AUC 0.578). One prior study (Lorenzoni et al., 2012) used an initial hospitalization as an outcome for the predictive model (AUC 0.80), however this study used a very small sample (n=380) which the authors report may have falsely inflated the model's predictive ability. As previously mentioned, similarities in model performance between hospitalizations and emergency department use likely stems from the frequency with which persons with HF are admitted to the hospital from the emergency department. Like the models for emergency department visits, the variables included in the models predicting hospitalizations explained roughly 70% of the variance, indicating that there are factors missing from the model. Significant factors in the models for hospitalization include education, asthma/COPD, SF12/PCS scores and baseline CVD. It is possible that the same variables that are not accounted for in the models predicting hospitalizations would also improve the model performance for emergency department visits. Based on the criteria used for model evaluation, the models performed poorly on average, and were only able to correctly classify individuals 60% of the time.

30-Day Hospital Readmission Much of the literature regarding machine learning approaches to predictive modeling uses the 30-day hospital readmission outcome. The average model performance in the current study (AUC 0.705) was slightly higher than the average model performance in the current literature (AUC 0.651). The variables of hospital LOS, comorbidity, income, and marital status were significant predictors in all four models developed in the current study. As previously discussed, the variables included in prior predictive models vary and are inconsistent, likely leading to the fluctuations in model performance. However, many prior models (Awan et al., 2019 and McLaren et al., 2015) did include the variable of hospital LOS as a significant predictor of 30-day readmission. This finding supports the findings of other authors, suggesting that hospital LOS should be included in all models attempting to predict 30-day readmissions in persons with HF. In fact, one author found that the inclusion of prior hospital admission data improved the predictive ability of the model by more than 5 percent (McLaren et al., 2015). Hospital LOS may not intuitively be a modifiable factor due to disease progression and treatment needs, however it can be implied that a shorter hospital LOS, if possible, would impact 30-day readmission rates in persons with HF, by avoiding hospital acquired complications.

Implications for the Use of Machine Learning in Nursing

The application of machine learning and predictive modeling in nursing research is valuable and should be considered for use in future research in nursing science. There are a limited number of studies conducted by nurses that use advanced data analytic techniques to better understand the HF population. It is important for the advancement of nursing science that primary care nurses be present on research teams that develop predictive models, as nurses are best able to influence the factors that affect health service use in persons living in the community with HF. To achieve this goal, it is necessary for data analytic methodologies to become a part of the nursing curriculum at both the undergraduate and graduate levels. More recently, an increasing number of courses and programs are becoming available within schools of nursing that examine the use of data science. However, to effectively influence health outcomes, nurses must not only become familiar with data science and its implications to practice, but also participate on interdisciplinary teams that are focused on health service use outcomes. The use of data science and machine learning within nursing science is crucial for progress and improved patient outcomes.

Limitations

There are several limitations to the current study which are related to original data collection, variable availability, and/or statistical constraints. Because the MEPS data are publicly available for use based on previously conducted interviews and surveys, the researcher has limited control over the methods and procedures used in original data collection. Further, the self-reported nature of the survey data as well as the availability of complete data for all variables might have influenced the reliability of data. For example, individuals were asked about sensitive personal information including the presence of decreased cognitive function, current BMI, level of education, and the presence of comorbidities. Due to original data collection procedures, there was approximately eight percent of data missing for the variables of SF-12/PCS and SF-12/MCS. Values for the missing variables were replaced with mean imputation, though this is not likely to have affected the outcome with such a large sample size. Additionally, due to changes that were made in study protocols during original data collection, there was a period where participants did not receive the SF-12 measures at baseline, which significantly affected the available sample. These missing surveys would have led to approximately fifteen percent missing data within the sample. To counteract this deficit, mean imputations of SF-12 scores for both the PCS and MCS components were used to adjust for missing values in the data. Of note, prior to data adjustment, the overall scores for the SF-12 measures for the study sample were lower than the general population mean. This may indicate that persons with HF are at a lower physical functioning status than previously measured in the population. Further, the MEPS data are based on self-report and the individual providing survey responses for the household members may not have been the same individual who had the current HF diagnosis. The self-reported approach combined with the potential second-hand nature of survey responses may have had an impact on accuracy and completeness of data, specifically related to presence of medically diagnosed comorbidities. For example, only 1% (n=17) of individuals reported baseline kidney disease, whereas in the general HF population the prevalence of kidney disease is much higher. Despite the invariance in data related to the presence of kidney disease, the variable was included in the analysis because of its repeated significance in prior studies.

Because the content of the surveys was determined ahead of time, researchers have no influence on the availability of specific variables of interest. For example, the original MEPS surveys have no clinical data such as laboratory values and included limited health data such as BMI, smoking status, and baseline comorbidities. Unlike many other inpatient databases that focus on clinical data for model development

(Solomon et al., 2007), the MEPS data do not include frequently analyzed HF-related characteristics such as laboratory values, ejection fraction, etiology, or NYHA classification. These clinical data are often only available through hospital admissions records. Because the MEPS database is a community-based sample, clinical data are unavailable and therefore the study sample may be unable to accurately represent the holistic nature of repeated readmissions in persons with HF. Additionally, there are several aspects of socioeconomic data that have been found to influence HF related self-care (Hawkins et al., 2016) were not measured in the sample that could have influenced model performance. Such variables may have included health literacy, caregiver involvement, number of medications, availability and frequency of primary care, and self-efficacy. It is unclear if or how the addition of other socioeconomic data might impact the performance of a predictive model related to health service use in persons with HF. Additionally, the current study did not examine the reason for emergency department use, or the admitting diagnosis if a person was hospitalized or readmitted. Therefore, this study examined all-cause emergency department visits, hospitalizations and 30-day readmissions rather than HF specific admissions.

Nursing Recommendations for Future Research

The current study looks deeper into the relationship among socioeconomic factors and health service use in persons living in the community with HF. Though the variables included in the model are not a comprehensive representation of all possible socioeconomic characteristics, they give a focused interpretation of significant factors leading to health service use in the HF population. However, many gaps continue to exist in the HF literature regarding predictive modeling and health service use outcomes. Many inconsistencies remain related to model success including sample composition, variable inclusion, and health service use outcome of focus. To develop a comprehensive predictive model for persons with HF, researchers should establish the nature of the relationship between concrete variables such as race and level of education and abstract socioeconomic characteristics that are not often considered in traditional models, such as the integration of cultural practices, caregiver considerations, cognition and health literacy, the obesity paradox including frailty, and self-efficacy. A recent meta-analysis of health literacy in persons with HF found that approximately 24% of individuals have inadequate health literacy, which is a leading risk factor for emergency department use and readmissions related to gaps in self-care (Fabbri, et al. 2021). Future emphasis in clinical care to identify gaps in the relationship between education and health literacy and HF knowledge may help practitioners to identify those at risk for health service use due to educational deficits. In addition, further research is needed to clarify the use of BMI as a diagnostic predictor in HF health service use models, specifically related to the obesity paradox and its relationship to frailty. While the mechanisms behind the obesity paradox in HF are unclear (Zadeh et al., 2004), It is likely that BMI alone does not effectively capture the effect of body composition on the HF disease trajectory and its subsequent effect on health service use in this population. Additional research opportunities exist related to the integration of cultural practices into predictive models, including unique healing practices, non-traditional medicinal remedies, and the influence of western medicine on the non-native health beliefs. Further, it is important to examine the influence of family caregivers in predictive models, specifically their role in reducing unnecessary health service use through support in medical management.

In addition, it is important to determine whether separate models are necessary for community and inpatient samples, as well as for each health service use outcome. It is possible that a separate model is necessary for each health service use outcome for both inpatient and community samples. This distinction is likely related to the differences in etiologies, treatment regimens, and situational components that affect the HF disease projection. Models based on community samples should integrate established theories related to health service use and behaviors into their predictive model. For example, in the current study, the study sample and multiple health service use outcomes align with the overarching Andersen Health Service Use Model (Figure 1). Andersen's original model (1968) and revised model (1995) posit that a combination of community-based, environmental factors and personal characteristics influence health service use and behaviors. Unlike prior studies which focus on clinical and inpatient data, the Andersen Model can be applied specifically to community-based populations. The predisposing, enabling, and perceived need factors, as described by the Andersen Model, include modifiable components of an individual's life that can influence the degree to which individuals utilize public health services. Several of Andersen's original model components were found to be significant predictors of health service use in our community-based HF sample. It is interesting to note that many of the significant predictors identified in the current study align with the perceived need and environmental categories from Andersen's Model. Perceived need factors included the presence of baseline comorbidity, physical functioning, and BMI, while environmental factors included hospital LOS. This is of importance, as the perceived need factors are believed by Andersen (1968) to be modifiable aspects of a person's life. If true, it can be assumed that some degree of intervention or adjustment to these factors could influence the use of health services. However, there are no studies specifically examining the relationship between Andersen's perceived need and environmental factors and health service use, or ways in which to modify these factors to impact health outcomes in persons with HF. While Andersen's Model is neither a complete representation of all persons with HF, nor a comprehensive summary of all settings in which persons with HF are found, it reflects many of the characteristics of individuals living in the community with HF who are at an increased risk of using public health services. Andersen's Model provides a foundation to guide future studies focused on community-based samples of persons with HF.

There are gaps in the literature related to persons with HF using the emergency department from the community setting. This study provides insight into several socioeconomic factors that influence emergency department use, however the models and factors identified are not an exhaustive list. There are some studies examining adverse events of persons with HF once in the emergency department, however several opportunities exist to further examine factors related to predictors of initial emergency room use. Researchers should consider the implications of emergency department use in persons with HF, recognizing that a significant number of emergency department visits result in hospitalizations. Further, it is important that future research distinguishes between persons using the emergency room for a HF-related concern rather than a secondary problem. It may be beneficial to build and compare predictive models for all-cause health service use and HF-specific health service use.

Future research on time to event models related to hospitalizations and emergency department use should expand the time window to include events that occur outside of the traditional 30-day window. This approach would allow for risk predictions to be developed for the non-acute individual living in the community with HF. The current study found that individuals experienced hospitalizations and emergency department visits throughout the twelve-month study period, suggesting that clinicians can provide interventions for the modifiable socioeconomic risk factors outside of the acute window where they are traditionally addressed.

Conclusion

The aims of the current study were threefold, to (1) determine the relationship among sociodemographic factors from a community-based HF sample and outcomes of health service use including hospitalizations, 30-day readmissions, and emergency department use, (2) build a predictive model for health service use based on significant socioeconomic factors, and (3) examine the relative risk and time to event trajectory of health service use for persons living in the community with HF. Significant predictors of increased health service use among the hospitalization and emergency department outcomes were similar and included individuals who were over 65 years of age, with poor physical functioning, and baseline asthma/COPD, and cardiovascular disease. Significant predictors of decreased health service use among the hospitalization and emergency department outcomes were also similar and included individuals who have high levels of educational attainment (at least a bachelor's degree), and high BMI. Similarities in the predictors for hospitalizations and emergency department use are likely due to the high number of hospitalizations that result from a visit to the emergency department for both HF-related and all-cause visits. Significant predictors of 30-day hospital readmission differed from the other health service use outcomes and included individuals of Asian descent who were divorced, made less than $50,000 per year, had seven or more comorbid conditions, and had a prior hospitalization of greater than one week. Predictive models developed for each health service use outcome performed modestly at best. The strongest performance was seen in the 30-day hospital readmission model, with hospital LOS being the most influential predictor across models. The average time to an event for hospitalizations (M=5.5) and emergency department use (M=5.3) after study enrollment were similar. Factors that increased the risk for a hospitalization and shortened the time to an event included advanced age (>65), poor physical functioning, and the presence of asthma/COPD, cardiovascular disease, stroke, and diabetes at baseline. Factors that increased the risk for an emergency department visit and shortened the time to an event included divorced marital status, having Medicare, being a current smoker, poor physical functioning, and the presence of asthma/COPD, depression, osteoporosis, and stroke at baseline. Our findings build upon previously developed predictive models for hospitalizations and 30-day hospital readmissions persons with HF, however, model performance remains ineffective. While the model did not perform well, the current study findings related to predictors of emergency department use, especially for the time to event models, add new information to the literature. In general, persons with HF who are of advanced age, physically frail, and managing multiple comorbidities are at greatest risk for health service use. Nurses in primary care settings within the community have the ability to identify these at-risk individuals and provide routine interventions to reduce community-based health service use. It may be beneficial to implement aa national HF registry that can track and monitor individual risk factors electronically across health systems and providers. This type of screening data could be used to develop predictive models in real-time for individuals based on risk factors present at each primary care visit. The development and use of predictive modes through machine learning approaches is essential in the reduction of health service use. However, further research is needed to examine the importance and influence of socioeconomic factors in risk prediction models, specifically related to the inclusion of abstract variables that are not routinely measured.

References

2020, H. P. (2018). Heart Disease and Stroke. Retrieved from https://www.healthypeople.gov/2020/topics-objectives/topic/heart-disease-and-stroke

(MEPS), M. E. P. S. (2012, August,2018). Medical Expenditure Panel Survey. Retrieved from https://www.ahrq.gov/data/meps.html

Aday, L. A., & Andersen, R. (1974). A framework for the study of access to medical care. Health Serv Res, 9(3), 208-220.

Albert, N. M., Levy, P., Langlois, E., Nutter, B., Yang, D., Kumar, V. A., Nykun, E. (2014). Heart failure beliefs and self-care adherence while being treated in an emergency department. Journal of Emergency Medicine, 46(1), 122-129. doi:10.1016/j.jemermed.2013.04.060

Allam, A., Nagy, M., Thoma, G., & Krauthammer, M. (2019). Neural networks versus Logistic regression for 30 days all-cause readmission prediction. Sci Rep, 9(1), 9277. doi:10.1038/s41598-019-45685-z

Aloisi, G., Zucchelli, A., Aloisi, B., Romanelli, G., & Marengoni, A. (2019). Depression and heart failure: an intricate relationship. Monaldi Arch Chest Dis, 89(3). doi:10.4081/monaldi.2019.1029

Amarasingham, R., Moore, B. J., Tabak, Y. P., Drazner, M. H., Clark, C. A., Zhang, S., Halm, E. A. (2010). An automated model to identify heart failure patients at risk for 30- day readmission or death using electronic medical record data. Med Care, 48(11), 981-988. doi:10.1097/MLR.0b013e3181ef60d9

Ambrosy, A. P., Fonarow, G. C., Butler, J., Chioncel, O., Greene, S. J., Vaduganathan, M., Gheorghiade, M. (2014). The global health and economic burden of hospitalizations for heart failure: lessons learned from hospitalized heart failure registries. Journal of the American College of Cardiology, v63(12), 1123- 1133. doi:10.1016/j.jacc.2013.11.053

Andersen, R. M. (1995). Revisiting the behavioral model and access to medical care: does it matter? J Health Social Behavior, 36(1), 1-10.

Angraal, S., Mortazavi, B. J., Gupta, A., Khera, R., Ahmad, T., Desai, N. R., Krumholz, H. M. (2019). Machine Learning Prediction of Mortality and Hospitalization in Heart Failure with Preserved Ejection Fraction. JACC Heart Fail. doi:10.1016/j.jchf.2019.06.013

Aranda, J. M., Jr., Johnson, J. W., & Conti, J. B. (2009). Current trends in heart failure readmission rates: analysis of Medicare data. Clin Cardiology, 32(1), 47-52. doi:10.1002/clc.20453

Ashfaq, A., Sant'Anna, A., Lingman, M., & Nowaczyk, S. (2019). Readmission prediction using deep learning on electronic health records. Journal Biomed Inform, 97, 103256. doi:10.1016/j.jbi.2019.103256

American Heart Association. (2019). Heart Failure. Retrieved from https://www.heart.org/en/health-topics/heart-failure

American Nurses Association. (2015). Code of ethics for nurses with interpretive statements. Retrieved from https://www.nursingworld.org/practice-policy/nursing-excellence/ethics/code-of-ethics-for-nurses/coe-view-only/

Au, A. G., McAlister, F. A., Bakal, J. A., Ezekowitz, J., Kaul, P., & van Walraven, C. (2012). Predicting the risk of unplanned readmission or death within 30 days of discharge after a heart failure hospitalization. Am Heart J, 164(3), 365-372. doi:10.1016/j.ahj.2012.06.010

Awan, S. E., Bennamoun, M., Sohel, F., Sanfilippo, F. M., Chow, B. J., & Dwivedi, G. (2019). Feature selection and transformation by machine learning reduce variable numbers and improve prediction for heart failure readmission or death. PLoS One, 14(6), e0218760. doi:10.1371/journal.pone.0218760

Awan, S. E., Bennamoun, M., Sohel, F., Sanfilippo, F. M., & Dwivedi, G. (2019). Machine learning-based prediction of heart failure readmission or death: implications of choosing the right model and the right metrics. ESC Heart Failure, 6(2), 428-435. doi:10.1002/ehf2.12419

Bekhet, A. K., & Zauszniewski, J. A. (2008). Theoretical substruction illustrated by the Theory of Learned Resourcefulness. Research Theory and Nursing Practice, 22(3), 205-214.

Bieler, G., Paroz, S., Faouzi, M., Trueb, L., Vaucher, P., Althaus, F., Bodenmann, P. (2012). Social and medical vulnerability factors of emergency department frequent users in a universal health insurance system. Academy of Emergency Medicine, 19(1), 63-68. doi:10.1111/j.1553-2712.2011.01246.x

Birmingham, L. E., Cochran, T., Frey, J. A., Stiffler, K. A., & Wilber, S. T. (2017). Emergency department use and barriers to wellness: a survey of emergency department frequent users. BMC Emergency Medicine, 17(1), 16. doi:10.1186/s12873-017-0126-5

Blecker, S., Sontag, D., Horwitz, L. I., Kuperman, G., Park, H., Reyentovich, A., & Katz, S. D. (2018). Early Identification of Patients with Acute Decompensated Heart Failure. J Cardiac Failure, 24(6), 357-362. doi:10.1016/j.cardfail.2017.08.458

Borenstein, M., Hodges,L.V., Higgins, J., Rothstei, H.R. (2009). Introduction to meta-analysis. Chichester, U.K.: John Wiley & Sons.

Bradford, C., Shah, B. M., Shane, P., Wachi, N., & Sahota, K. (2017). Patient and clinical characteristics that heighten risk for heart failure readmission. Research in Social & Administrative Pharmacy, 13(6), 1070-1081. doi:10.1016/j.sapharm.2016.11.002

Brookfield, S. (1987). Developing critical thinkers: challenging adults to explore alternative ways of thinking and acting (1st ed.). San Francisco, Calif.: Jossey-Bass.

Bytyçi, I., & Bajraktari, G. (2015). Mortality in heart failure patients. Anatolian journal of cardiology, 15(1), 63-68. doi:10.5152/akd.2014.5731

Calvillo-King, L., Arnold, D., Eubank, K. J., Lo, M., Yunyongying, P., Stieglitz, H., & Halm, E. A. (2013). Impact of social factors on risk of readmission or mortality in pneumonia and heart failure: systematic review. J Gen Intern Med, 28(2), 269-282. doi:10.1007/s11606-012-2235-x

Capp, R., Kelley, L., Ellis, P., Carmona, J., Lofton, A., Cobbs-Lomax, D., & D'Onofrio, G. (2016). Reasons for Frequent Emergency Department Use by Medicaid Enrollees: A Qualitative Study. Academy of Emergency Medicine, 23(4), 476-481.doi:10.1111/acem.12952

Chamberlain, R. S., Sond, J., Mahendraraj, K., Lau, C. S., & Siracuse, B. L. (2018). Determining 30-day readmission risk for heart failure patients: The Readmission After Heart Failure scale. Int J Gen Med, 11, 127-141. doi:10.2147/ijgm.S150676

Chamberlain, R. S., Sond, J., Mahendraraj, K., Lau, C. S., & Siracuse, B. L. (2018). Determining 30-day readmission risk for heart failure patients: the Readmission After Heart Failure scale. Int J Gen Med, 11, 127-141. doi:10.2147/IJGM.S150676

Chen, J., Sadasivam, R., Blok, A. C., Ritchie, C. S., Nagawa, C., Orvek, E. Houston, T. K. (2020). The Association Between Patient-reported Clinical Factors and 30-day Acute Care Utilization in Chronic Heart Failure. Med Care, 58(4), 336-343. doi:10.1097/mlr.0000000000001258

Chen, S., Kong, N., Sun, X., Meng, H., & Li, M. (2019). Claims data-driven modeling of hospital time-to-readmission risk with latent heterogeneity. Health Care Management Science, 22(1), 156-179. doi:10.1007/s10729-018-9431-0

Claret, P. G., Calder, L. A., Stiell, I. G., Yan, J. W., Clement, C. M., Borgundvaag, B., Rowe, B. H. (2018). Rates and predictive factors of return to the emergency department following an initial release by the emergency department for acute heart failure. CJEM, 20(2), 222-229. doi:10.1017/cem.2017.14

Connelly, R., Gayle, V., & Lambert, P. S. (2016). A review of educational attainment measures for social survey research. Methodological Innovations, 9, 2059799116638001. doi:10.1177/2059799116638001

Cook, C., Cole, G., Asaria, P., Jabbour, R., & Francis, D. P. (2014). The annual global economic burden of heart failure. International Journal of Cardiology, 171(3), 368-376. doi:10.1016/j.ijcard.2013.12.028

Cornell, S. E., & Hartmann, D. (1998). Ethnicity and race: making identities in a changing world. Thousand Oaks, Calif.: Pine Forge Press.

Cox, Z. L., Lai, P., Lewis, C. M., & Lindenfeld, J. (2020). Body mass index and all-cause readmissions following acute heart failure hospitalization. International Journal of Obesity. (London), 44(6), 1227-1235. doi:10.1038/s41366-019-0518-6

Curtis, J. P., Selter, J. G., Wang, Y., Rathore, S. S., Jovin, I. S., Jadbabaie, F., Krumholz, H. M. (2005). The obesity paradox: body mass index and outcomes in patients with heart failure. Arch Internal Medicine, 165(1), 55-61. doi:10.1001/archinte.165.1.55

Duero Posada, J. G., Moayedi, Y., Zhou, L., McDonald, M., Ross, H. J., Lee, D. S., & Bhatia, R. S. (2018). Clustered Emergency Room Visits Following an Acute Heart Failure Admission: A Population-Based Study. Journal of the American Heart Association, 7(7). doi:10.1161/JAHA.117.007569

Eapen, Z. J., McCoy, L. A., Fonarow, G. C., Yancy, C. W., Miranda, M. L., Peterson, E. D., Hernandez, A. F. (2015). Utility of socioeconomic status in predicting 30-day outcomes after heart failure hospitalization. Circulation Heart Failure, 8(3), 473-480.doi:10.1161/circheartfailure.114.001879

Evans, R. S., Benuzillo, J., Horne, B. D., Lloyd, J. F., Bradshaw, A., Budge, D., Lappe, D. L. (2016). Automated identification and predictive tools to help identify high-risk heart failure patients: pilot evaluation. Journal of American Medical Information Association, 23(5), 872-878.doi:10.1093/jamia/ocv197

Fabbri, M., Murad, M. H., Wennberg, A. M., Turcano, P., Erwin, P. J., Alahdab, F., Roger, V. L. (2020). Health Literacy and Outcomes Among Patients with Heart Failure: A Systematic Review and Meta-Analysis. JACC Heart Failure, 8(6), 451-460. doi:10.1016/j.jchf.2019.11.007

Failde, I., Medina, P., Ramirez, C., & Arana, R. (2010). Construct and criterion validity of the SF-12 health questionnaire in patients with acute myocardial infarction and unstable angina. Journal Evaluating Clinical Practice, 16(3), 569-573. doi:10.1111/j.1365-2753.2009.01161.x

Farré, N., Vela, E., Clèries, M., Bustins, M., Cainzos-Achirica, M., Enjuanes, C., Comín-Colet, J. (2017). Real world heart failure epidemiology and outcome: A population-based analysis of 88,195 patients. PLoS One, 12(2), e0172745-e0172745. doi:10.1371/journal.pone.0172745

Ferro, M. A. (2019). The Psychometric Properties of the Kessler Psychological Distress Scale (K6) in an Epidemiological Sample of Canadian Youth. Canadian Journal of Psychiatry, 64(9), 647-657. doi:10.1177/0706743718818414

Field, A. (2013). Discovering Statistics using IBM SPSS Statistics: Sage Publications Ltd.

Fortin, M., Almirall, J., & Nicholson, K. (2017). Development of a research tool to document self-reported chronic conditions in primary care. Journal of Comorbidity, 7(1), 117-123. doi:10.15256/joc.2017.7.122

Freedland, K. E., Carney, R. M., Rich, M. W., Steinmeyer, B. C., Skala, J. A., & Davila-Roman, V. G. (2016). Depression and Multiple Rehospitalizations in Patients with Heart Failure. Clinical Cardiology, 39(5), 257-262. doi:10.1002/clc.22520

Friedman, B. W., Serrano, D., Reed, M., Diamond, M., & Lipton, R. B. (2009). Use of the emergency department for severe headache. A population-based study. Headache, 49(1), 21-30. doi:10.1111/j.1526-4610.2008.01282.x

Frizzell, J. D., Liang, L., Schulte, P. J., Yancy, C. W., Heidenreich, P. A., Hernandez, A. F., Laskey, W. K. (2017). Prediction of 30-Day All-Cause Readmissions in Patients Hospitalized for Heart Failure: Comparison of Machine Learning and Other Statistical Approaches. JAMA Cardiology, 2(2), 204-209. doi:10.1001/jamacardio.2016.3956

Fudim, M., O'Connor, C. M., Dunning, A., Ambrosy, A. P., Armstrong, P. W., Coles, A., Mentz, R. J. (2018). Aetiology, timing and clinical predictors of early vs. late readmission following index hospitalization for acute heart failure: insights from ASCEND-HF. European Journal of Heart Failure, 20(2), 304-314. doi:10.1002/ejhf.1020

Gajulapalli, R. D., Kadri, A., Gad, M., Chahine, J., Nusairat, L., & Rader, F. (2020). Impact of Obesity in Hospitalized Patients with Heart Failure: A Nationwide Cohort Study. South Medical Journal, 113(11), 568-577. doi:10.14423/SMJ.0000000000001174

Galderisi, S., Heinz, A., Kastrup, M., Beezhold, J., & Sartorius, N. (2015). Toward a new definition of mental health. World psychiatry: official journal of the World Psychiatric Association (WPA), 14(2), 231-233. doi:10.1002/wps.20231

George, B., Seals, S., & Aban, I. (2014). Survival analysis and regression models. Journal of Nuclear Cardiology, 21(4), 686-694. doi:10.1007/s12350-014-9908-2

Go, Y. Y., Sellmair, R., Allen, J. C., Jr., Sahlen, A., Bulluck, H., Sim, D., Lam, C. S. P. (2019). Defining a 'frequent admitter' phenotype among patients with repeat heart failure admissions. European Journal of Heart Failure, 21(3), 311-318. doi:10.1002/ejhf.1348

Golas, S. B., Shibahara, T., Agboola, S., Otaki, H., Sato, J., Nakae, T., Jethwani, K. (2018). A machine learning model to predict the risk of 30-day readmissions in patients with heart failure: a retrospective analysis of electronic medical records data. BMC Medical Information Decision Making, 18(1), 44. doi:10.1186/s12911-018-0620-z

Griswold, S. K., Nordstrom, C. R., Clark, S., Gaeta, T. J., Price, M. L., & Camargo, C. A., Jr. (2005). Asthma exacerbations in North American adults: who are the "frequent fliers" in the emergency department? Chest, 127(5), 1579-1586. doi:10.1378/chest.127.5.1579

Hamner, J. B., & Ellison, K. J. (2005). Predictors of hospital readmission after discharge in patients with congestive heart failure. Heart Lung, 34(4), 231-239. doi:10.1016/j.hrtlng.2005.01.001

Health Quality Outcomes. (2017). Effect of Early Follow-Up After Hospital Discharge on Outcomes in Patients with Heart Failure or Chronic Obstructive Pulmonary Disease: A Systematic Review. Ontario Health Technology Assess Service, 17(8), 1-37. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/28638496

Heo, S., Moser, D. K., Lennie, T. A., Riegel, B., & Chung, M. L. (2008). Gender differences in and factors related to self-care behaviors: a cross-sectional, correlational study of patients with heart failure. International Journal of Nursing Studies, 45(12), 1807-1815. doi:10.1016/j.ijnurstu.2008.05.008

Holley, C. T., Harvey, L., & John, R. (2014). Left ventricular assist devices as a bridge to cardiac transplantation. Journal of Thoracic Dis, 6(8), 1110-1119. doi:10.3978/j.issn.2072-1439.2014.06.46

Horwich, T. B., Fonarow, G. C., & Clark, A. L. (2018). Obesity and the Obesity Paradox in Heart Failure. Prog Cardiovasc Disease, 61(2), 151-156. doi:10.1016/j.pcad.2018.05.005

Horwich, T. B., Fonarow, G. C., Hamilton, M. A., MacLellan, W. R., Woo, M. A., & Tillisch, J. H. (2001). The relationship between obesity and mortality in patients with heart failure. Journal of the American College of Cardiology, 38(3), 789-795. doi:10.1016/s0735-1097(01)01448-6

Hosenpud, J. D., & Greenberg, B. H. (2000). Congestive heart failure: pathophysiology, diagnosis, and comprehensive approach to management (2nd ed ed.). Philadelphia: Lippincott Williams & Wilkins.

Howie-Esquivel, J., & Spicer, J. G. (2012). Association of partner status and disposition with rehospitalization in heart failure patients. American Journal of Critical Care, 21(3), e65-73. doi:10.4037/ajcc2012382

Inouye, S., Bouras, V., Shouldis, E., Johnstone, A., Silverzweig, Z., & Kosuri, P. (2015). Predicting readmission of heart failure patients using automated follow-up calls.

BMC Medically Informed Decision Making, 15, 22. doi:10.1186/s12911-015-0144-8

Jiang, W., Siddiqui, S., Barnes, S., Barouch, L. A., Korley, F., Martinez, D. A.,Levin, S. (2019). Readmission Risk Trajectories for Patients with Heart Failure Using a Dynamic Prediction Approach: Retrospective Study. JMIR Medical Information, 7(4), e14756. doi:10.2196/14756

Jiang, W., Siddiqui, S., Barnes, S., Barouch, L. A., Korley, F., Martinez, D. A., Levin, S. (2019). Readmission Risk Trajectories for Patients with Heart Failure Using a Dynamic Prediction Approach: Retrospective Study. JMIR Medical Information, 7(4), e14756. doi:10.2196/14756

Johnson, T. J., Basu, S., Pisani, B. A., Avery, E. F., Mendez, J. C., Calvin, J. E., Jr., & Powell, L. H. (2012). Depression predicts repeated heart failure hospitalizations. Journal of Cardiac Failure, 18(3), 246-252. doi:10.1016/j.cardfail.2011.12.005

Kakarmath, S., Golas, S., Felsted, J., Kvedar, J., Jethwani, K., & Agboola, S. (2018). Validating a Machine Learning Algorithm to Predict 30-Day Re-Admissions in Patients with Heart Failure: Protocol for a Prospective Cohort Study. JMIR Res Protocol, 7(9), e176. doi:10.2196/resprot.9466

Kalantar-Zadeh, K., Anker, S. D., Coats, A. J., Horwich, T. B., & Fonarow, G. C. (2005). Obesity paradox as a component of reverse epidemiology in heart failure. Arch Internal Medicine, 165(15), 1797; author reply 1797-1798. doi:10.1001/archinte.165.15.1797-a

Katz, S., Ford, A. B., Moskowitz, R. W., Jackson, B. A., & Jaffe, M. W. (1963). Studies of illness sin the aged. The index of ADL: as standardized measure of biological and psychological function. JAMA, 185, 914-919. doi:10.1001/jama.1963.03060120024016

Keenan, P. S., Normand, S. L., Lin, Z., Drye, E. E., Bhat, K. R., Ross, J. S., Krumholz, H. M. (2008). An administrative claims measure suitable for profiling hospital performance on the basis of 30-day all-cause readmission rates among patients with heart failure. Circulation Cardiovasc Quality Outcomes, 1(1), 29-37. doi:10.1161/CIRCOUTCOMES.108.802686

Kessler, R. C., Green, J. G., Gruber, M. J., Sampson, N. A., Bromet, E., Cuitan, M., Zaslavsky, A. M. (2010). Screening for serious mental illness in the general population with the K6 screening scale: results from the WHO World Mental Health (WMH) survey initiative. International Journal of Methods Psychiatric Research, 19 Suppl 1, 4-22. doi:10.1002/mpr.310

Kewcharoen, J., Tachorueangwiwat, C., Kanitsoraphan, C., Saowapa, S., Nitinai, N., Vutthikraivit, W., Banerjee, D. (2020). Depression is associated with an increased risk of readmission in patients with heart failure: a systematic review and meta-analysis. Minerva Cardioangiology. doi:10.23736/S0026-4725.20.05346-3

Kitamura, M., Izawa, K. P., Yaekura, M., Mimura, Y., Ikeda, Y., Nagashima, H., & Brubaker, P. H. (2019). Relationship among Activities of Daily Living, Nutritional Status, and 90 Day Readmission in Elderly Patients with Heart Failure. International Journal of Environmental Research in Public Health, 16(24). doi:10.3390/ijerph16245068

Krieg, C., Hudon, C., Chouinard, M.-C., & Dufour, I. (2016). Individual predictors of frequent emergency department use: a scoping review. BMC health services research, 16(1), 594- 594. doi:10.1186/s12913-016-1852-1

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2003). The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care, 41(11), 1284-1292. Doi:10.1097/01.Mlr.0000093487.78664.3c

Kwok, C. S., Zieroth, S., Van Spall, H. G. C., Helliwell, T., Clarson, L., Mohamed, M., Mamas, M. A. (2020). The Hospital Frailty Risk Score and its association with in-hospital mortality, cost, length of stay and discharge location in patients with heart failure short running title: Frailty and outcomes in heart failure. International Journal of Cardiology, 300, 184-190. doi:10.1016/j.ijcard.2019.09.064

Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist, 9(3), 179-186.

Lindenauer, P. K., Lagu, T., Rothberg, M. B., Avrunin, J., Pekow, P. S., Wang, Y., & Krumholz, H. M. (2013). Income inequality and 30-day outcomes after acute myocardial infarction, heart failure, and pneumonia: retrospective cohort study. BMJ, 346, f521. doi:10.1136/bmj.f521

Lorenzoni, G., Sabato, S. S., Lanera, C., Bottigliengo, D., Minto, C., Ocagli, H., Pisano, F. (2019). Comparison of Machine Learning Techniques for Prediction of Hospitalization in Heart Failure Patients. Journal of Clinical Medicine, 8(9). doi:10.3390/jcm8091298

Lu, M. L. R., Davila, C. D., Shah, M., Wheeler, D. S., Ziccardi, M. R., Banerji, S., & Figueredo, V. M. (2016). Marital status and living condition as predictors of mortality and readmissions among African Americans with heart failure. International Journal of Cardiology, 222, 313-318. doi:10.1016/j.ijcard.2016.07.185

Mahajan, S., Burman, P., & Hogarth, M. (2016). Analyzing 30-Day Readmission Rate for Heart Failure Using Different Predictive Models. Studies of Health Technology Information, 225, 143-147.

Mahajan, S. M., Burman, P., Newton, A., & Heidenreich, P. A. (2017). A Validated Risk Model for 30-Day Readmission for Heart Failure. Studies of Health Technology Information, 245, 506-510.

Mahajan, S. M., & Ghani, R. (2019). Combining Structured and Unstructured Data for Predicting Risk of Readmission for Heart Failure Patients. Studies in Health Technology Information, 264, 238-242. doi:10.3233/shti190219

Mahajan, S. M., & Ghani, R. (2019). Using Ensemble Machine Learning Methods for Predicting Risk of Readmission for Heart Failure. Studies in Health Technology Information, 264, 243-247. doi:10.3233/shti190220

Mahajan, S. M., & Ghani, R. (2019). Using Ensemble Machine Learning Methods for Predicting Risk of Readmission for Heart Failure. Studies in Health Technology Information, 264, 243-247. doi:10.3233/shti190220

Mahajan, S. M., Mahajan, A. S., King, R., & Negahban, S. (2018). Predicting Risk of 30-Day Readmissions Using Two Emerging Machine Learning Methods. Studies in Health Technology Information, 250, 250-255.

Mahajan, S. M., Mahajan, A. S., & Negahban, S. (2018). Regional Differences in Predicting Risk of 30-Day Readmissions for Heart Failure. Studies in Health Technology Informatics, 250, 245-249. doi:10.3233/978-1-61499-872-3-245

Mandviwala, T. M., Basra, S. S., Khalid, U., Pickett, J. K., Przybylowicz, R., Shah, T., Deswal, A. (2020). Obesity and the paradox of mortality and heart failure hospitalization in heart failure with preserved ejection fraction. International Journal of Obesity (London), 44(7), 1561-1567. doi:10.1038/s41366-020-0563-1

Manemann, S. M., Chamberlain, A. M., Roger, V. L., Boyd, C., Cheville, A., Dunlay, S. M., Rutten, L. J. F. (2018). Multimorbidity and Functional Limitation in Individuals with Heart Failure: A Prospective Community Study. Journal of the American Geriatric Society, 66(6), 1101-1107. doi:10.1111/jgs.15336

Masetic, Z., & Subasi, A. (2016). Congestive heart failure detection using random forest classifier. Computer Methods Programs Biomed, 130, 54-64. doi:10.1016/j.cmpb.2016.03.020

Massari, F., Scicchitano, P., Ciccone, M. M., Caldarola, P., Aspromonte, N., Iacoviello, M., Valle, R. (2019). Bioimpedance vector analysis predicts hospital length of stay in acute heart failure. Nutrition, 61, 56-60. doi:10.1016/j.nut.2018.10.028

McLaren, D. P., Jones, R., Plotnik, R., Zareba, W., McIntosh, S., Alexis, J., Kutyifa, V. (2016). Prior hospital admission predicts thirty-day hospital readmission for heart failure patients. Cardiology Journal, 23(2), 155-162. doi:10.5603/CJ.a2016.0005

McNallan, S. M., Singh, M., Chamberlain, A. M., Kane, R. L., Dunlay, S. M., Redfield, M. M., Roger, V. L. (2013). Frailty and healthcare utilization among patients with heart failure in the community. JACC Heart Failure, 1(2), 135-141. doi:10.1016/j.jchf.2013.01.002

Miro, O., Garcia Sarasola, A., Fuenzalida, C., Calderon, S., Jacob, J., Aguirre, A., Group, I.-S. R. (2019). Departments involved during the first episode of acute heart failure and subsequent emergency department revisits and rehospitalizations: an outlook through the NOVICA cohort. European Journal of Heart Failure, 21(10), 1231-1244. doi:10.1002/ejhf.1567

Montoy, J. C. C., Tamayo-Sarver, J., Miller, G. A., Baer, A. E., & Peabody, C. R. (2019). Predicting Emergency Department "Bouncebacks": A Retrospective Cohort Analysis. Western Journal of Emergency Medicine, 20(6), 865-874. doi:10.5811/westjem.2019.8.43221

Mortazavi, B. J., Downing, N. S., Bucholz, E. M., Dharmarajan, K., Manhapra, A., Li, S. X., Krumholz, H. M. (2016). Analysis of Machine Learning Techniques for Heart Failure Readmissions. Circulation Cardiovascular Quality Outcomes, 9(6), 629-640. doi:10.1161/circoutcomes.116.003039

Murray, M. D., Tu, W., Wu, J., Morrow, D., Smith, F., & Brater, D. C. (2009). Factors associated with exacerbation of heart failure include treatment adherence and health literacy skills. Clinical Pharmacological Therapy, 85(6), 651-658. doi:10.1038/clpt.2009.7

Ouwerkerk, W., Voors, A. A., & Zwinderman, A. H. (2014). Factors influencing the predictive power of models for predicting mortality and/or heart failure hospitalization in patients with heart failure. JACC Heart Failure, 2(5), 429-436. doi:10.1016/j.jchf.2014.04.006

Pamela N. Peterson, L. A. A., Paul A. Heidenreich, Nancy M. Albert, Ileana L. Piña, American Heart Association. (2017). The American Heart Association Heart Failure Summit. Circulation Heart Failure, 11(10).

Pandey, A., Golwala, H., Xu, H., DeVore, A. D., Matsouaka, R., Pencina, M., Fonarow, G. C. (2016). Association of 30-Day Readmission Metric for Heart Failure Under the Hospital Readmissions Reduction Program with Quality of Care and Outcomes. JACC: Heart Failure, 4(12), 935-946. doi:https://doi.org/10.1016/j.jchf.2016.07.003

Patel, N., Chakraborty, S., Bandyopadhyay, D., Amgai, B., Hajra, A., Atti, V., Fonarow, G. C. (2020). Association between depression and readmission of heart failure: A national representative database study. Progressive Cardiovasc Disease, 63(5), 585-590. doi:10.1016/j.pcad.2020.03.014

Patil, S., Shah, M., Patel, B., Agarwal, M., Ram, P., & Alla, V. M. (2019). Readmissions Among Patients Admitted with Acute Decompensated Heart Failure Based on Income Quartiles. Mayo Clinic Proc, 94(10), 1939-1950. doi:10.1016/j.mayocp.2019.05.027

Pines, J. M., Asplin, B. R., Kaji, A. H., Lowe, R. A., Magid, D. J., Raven, M., Yealy, D. M. (2011). Frequent users of emergency department services: gaps in knowledge and a proposed research agenda. Academy of Emergency Medicine, 18(6), e64-69. doi:10.1111/j.1553-2712.2011.01086.x

Ponce, S. G., Norris, J., Dodendorf, D., Martinez, M., Cox, B., & Laskey, W. (2018). Impact of Ethnicity, Sex, and Socio-Economic Status on the Risk for Heart

Failure Readmission: The Importance of Context. Ethnicity & Disease, 28(2), 99-104. doi:10.18865/ed.28.2.99

Raphael, C., Briscoe, C., Davies, J., Ian Whinnett, Z., Manisty, C., Sutton, R., Francis, D. P. (2007). Limitations of the New York Heart Association functional classification system and self-reported walking distances in chronic heart failure. Heart (British Cardiac Society), 93(4), 476-482. doi:10.1136/hrt.2006.089656

Reynolds, K., Butler, M. G., Kimes, T. M., Rosales, A. G., Chan, W., & Nichols, G. A. (2015). Relation of Acute Heart Failure Hospital Length of Stay to Subsequent Readmission and All-Cause Mortality. American Journal of Cardiology, 116(3), 400-405. doi:10.1016/j.amjcard.2015.04.052

Ross, J. S., Mulvey, G. K., Stauffer, B., Patlolla, V., Bernheim, S. M., Keenan, P. S., & Krumholz, H. M. (2008). Statistical models and patient predictors of readmission for heart failure: a systematic review. Arch Internal Medicine, 168(13), 1371-1386. doi:10.1001/archinte.168.13.1371

Sax, D. R., Mark, D. G., Hsia, R. Y., Tan, T. C., Tabada, G. H., & Go, A. S. (2017). Short-Term Outcomes and Factors Associated with Adverse Events Among Adults Discharged From the Emergency Department After Treatment for Acute Heart Failure. Circulation Heart Failure, 10(12). doi:10.1161/CIRCHEARTFAILURE.117.004144

Sax, D. R., Mark, D. G., Huang, J., Sofrygin, O., Rana, J. S., Collins, S. P., Reed, M. E. (2021). Use of Machine Learning to Develop a Risk-Stratification Tool for Emergency Department Patients with Acute Heart Failure. Ann Emergency Medicine, 77(2), 237-248. doi:10.1016/j.annemergmed.2020.09.436

Shameer, K., Johnson, K. W., Yahi, A., Miotto, R., Li, L. I., Ricks, D., Dudley, J. T. (2017). Predictive modeling of hospital readmission rates using electronic medical-wide records: A case study using mount Saini cohort. Pac Symptom Biocomputer, 22, 276-287. doi:10.1142/9789813207813_0027

Sharma, A., Lavie, C. J., Borer, J. S., Vallakati, A., Goel, S., Lopez-Jimenez, F., Lazar, J. M. (2015). Meta-analysis of the relation of body mass index to all-cause and cardiovascular mortality and hospitalization in patients with chronic heart failure. American Journal of Cardiology, 115(10), 1428-1434. doi:10.1016/j.amjcard.2015.02.024

Shimizu, Y., Yamada, S., Suzuki, M., Miyoshi, H., Kono, Y., Izawa, H., Murohara, T. (2010). Development of the performance measure for activities of daily living-8 for patients with congestive heart failure: a preliminary study. Gerontology, 56(5), 459-466. doi:10.1159/000248628

Solomon, S. D., Dobson, J., Pocock, S., Skali, H., McMurray, J. J. V., Granger, C. B., Pfeffer, M. A. (2007). Influence of Nonfatal Hospitalization for Heart Failure on Subsequent Mortality in Patients with Chronic Heart Failure. Circulation, 116(13), 1482-1487. doi:10.1161/circulationaha.107.696906

Son, Y. J., Kim, H. G., Kim, E. H., Choi, S., & Lee, S. K. (2010). Application of support vector machine for prediction of medication adherence in heart failure patients. Healthcare Informatics Research, 16(4), 253-259. doi:10.4258/hir.2010.16.4.253

Son, Y. J., & Lee, H. J. (2020). Association between persistent smoking after a diagnosis of heart failure and adverse health outcomes: A systematic review and meta-analysis. Tob Induc Dis, 18, 05. doi:10.18332/tid/116411

Son, Y. J., & Won, M. H. (2018). Symptom Clusters and Their Impacts on Hospital Readmission in Patients with Heart Failure: A Cross-Sectional Study. Research Theory and Nursing Practice, 32(3), 311-327. doi:10.1891/1541-6577.32.3.311

Storrow, A. B., Jenkins, C. A., Self, W. H., Alexander, P. T., Barrett, T. W., Han, J. H., Collins, S. P. (2014). The burden of acute heart failure on U.S. emergency departments. JACC Heart Failure, 2(3), 269-277. doi:10.1016/j.jchf.2014.01.006

Sudhakar, S., Zhang, W., Kuo, Y. F., Alghrouz, M., Barbajelata, A., & Sharma, G. (2015). Validation of the Readmission Risk Score in Heart Failure Patients at a Tertiary Hospital. Journal of Cardiac Failure, 21(11), 885-891. doi:10.1016/j.cardfail.2015.07.010

Sun, B. C., Burstin, H. R., & Brennan, T. A. (2003). Predictors and outcomes of frequent emergency department users. Academy of Emergency Medicine, 10(4), 320-328. doi:10.1111/j.1553-2712.2003.tb01344.x

Szumilas, M. (2010). Explaining odds ratios. Journal of the Canadian Academy of Child and Adolescent Psychiatry, 19(3), 227-229. Retrieved from https://pubmed.ncbi.nlm.nih.gov/20842279

Takabayashi, K., Kitaguchi, S., Iwatsu, K., Morikami, Y., Ichinohe, T., Yamamoto, T., Nohara, R. (2019). A decline in activities of daily living due to acute heart failure is an independent risk factor of hospitalization for heart failure and mortality. Journal of Cardiology, 73(6), 522-529. doi:10.1016/j.jjcc.2018.12.014

Trevethan, R. (2017). Sensitivity, Specificity, and Predictive Values: Foundations, Liabilities, and Pitfalls in Research and Practice. Frontiers in public health, 5, 307-307. doi:10.3389/fpubh.2017.00307

Tsuchihashi, M., Tsutsui, H., Kodama, K., Kasagi, F., Setoguchi, S., Mohr, M., Takeshita, A. (2001). Medical and socioenvironmental predictors of hospital readmission in patients with congestive heart failure. American Heart Journal, 142(4), E7. doi:10.1067/mhj.2001.117964

Turgeman, L., & May, J. H. (2016). A mixed-ensemble model for hospital readmission. Artificial Intelligence Medicine, 72, 72-82. doi:10.1016/j.artmed.2016.08.005

Vader, J. M., LaRue, S. J., Stevens, S. R., Mentz, R. J., DeVore, A. D., Lala, A., de Las Fuentes, L. (2016). Timing and Causes of Readmission After Acute Heart Failure Hospitalization-Insights from the Heart Failure Network Trials. Journal of Cardiac Failure, 22(11), 875- 883. doi:10.1016/j.cardfail.2016.04.014

Valderas, J. M., Starfield, B., Sibbald, B., Salisbury, C., & Roland, M. (2009). Defining comorbidity: implications for understanding health and health services. Annals of family medicine, 7(4), 357-363. doi:10.1370/afm.983

Walsh, C., & Hripcsak, G. (2014). The effects of data sources, cohort selection, and outcome definition on a predictive model of risk of thirty-day hospital readmissions. Journal of Biomedical Informatics, 52, 418-426. doi:10.1016/j.jbi.2014.08.006

Walsh, C. G., Sharman, K., & Hripcsak, G. (2017). Beyond discrimination: A comparison of calibration methods and clinical usefulness of predictive models of readmission risk. Journal of Biomedical Informatics, 76, 9-18. doi:10.1016/j.jbi.2017.10.008

Wang, H., Robinson, R. D., Johnson, C., Zenarosa, N. R., Jayswal, R. D., Keithley, J., & Delaney, K. A. (2014). Using the LACE index to predict hospital readmissions in congestive heart failure patients. BMC Cardiovascular Disorders, 14, 97. doi:10.1186/1471-2261-14-97

Ware, J., Jr., Kosinski, M., & Keller, S. D. (1996). A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Medical Care, 34(3), 220-233. doi:10.1097/00005650-199603000-00003

Wee, C. C., Davis, R. B., & Hamel, M. B. (2008). Comparing the SF-12 and SF-36 health status questionnaires in patients with and without obesity. Health and Quality of Life Outcomes, 6(1), 11. doi:10.1186/1477-7525-6-11

Williams, B. A., Geba, D., Cordova, J. M., & Shetty, S. S. (2020). A risk prediction model for heart failure hospitalization in type 2 diabetes mellitus. Clinical Cardiology, 43(3), 275-283. doi:10.1002/clc.23298

Xu, J., Gallo, J. J., Wenzel, J., Nolan, M. T., Budhathoki, C., Abshire, M., Han, H. R. (2018). Heart Failure Rehospitalization and Delayed Decision Making: The Impact of Self-care and Depression. Journal of Cardiovascular Nursing, 33(1), 30-39. doi:10.1097/JCN.0000000000000423

Yamada, S., Shimizu, Y., Suzuki, M., Izumi, T., & collaborators, P. T. (2012). Functional limitations predict the risk of rehospitalization among patients with chronic heart failure. Circ J, 76(7), 1654-1661. doi:10.1253/circj.cj-11-1178

Yu, S., Farooq, F., van Esbroeck, A., Fung, G., Anand, V., & Krishnapuram, B. (2015). Predicting readmission risk with institution-specific prediction models. Artificial Intelligence Medicine, 65(2), 89-96. doi:10.1016/j.artmed.2015.08.005

Yu, X., Stewart, S. M., Wong, P. T. K., & Lam, T. H. (2011). Screening for depression with the Patient Health Questionnaire-2 (PHQ-2) among the general population in

Hong Kong. Journal of Affective Disorders, 134(1), 444-447. doi:https://doi.org/10.1016/j.jad.2011.05.007

Yuan, H., Fan, X. S., Jin, Y., He, J. X., Gui, Y., Song, L. Y., Chen, W. (2019). Development of heart failure risk prediction models based on a multi-marker approach using random forest algorithms. Chinese Medical Journal (England), 132(7), 819-826. doi:10.1097/cm9.0000000000000149

Kalantar-Zadeh, K., Block, G., Horwich, T., & Fonarow, G. C. (2004). Reverse epidemiology of conventional cardiovascular risk factors in patients with chronic heart failure. Journal of the American College of Cardiology, 43(8), 1439-1444. doi:10.1016/j.jacc.2003.11.039