SOCIOECONOMICS STATUS AND HOSPITALIZATION CHARACTERISTICS IN THE UNITED STATES: A RETROSPECTIVE STUDY

By

Adiebonye E. Jumbo

A Dissertation Submitted to

the Rutgers University - School of Health Related Professions

in partial fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Biomedical Informatics

Department of Health Informatics

Fall 2015

SOCIOECONOMICS STATUS AND HOSPITALIZATION CHARACTERISTICS IN THE UNITED STATES: A RETROSPECTIVE STUDY

BY

Adiebonye E. Jumbo

Dissertation Committee:

Approval by the Dissertation Committee:

Shankar Srinivasan PhD

Jared Sender MD MBA

Frederick Coffman PhD

ABSTRACT

For several decades, social disparities in access to health care remain a major debate in the U.S. health care system. Despite growing attention to health inequalities, different social classes, especially, minority or ethnic groups, and those without health insurance coverage continue to face challenges to health care. To date, due to the complexities of Socioeconomic Status (SES), it is unclear how SES impacts health and income inequality. The purpose of this dissertation was to examine the association of SES and median household income groups with hospitalization outcomes in the United States from 2008 to 2010. To examine the generalizability of this phenomenon, a retrospective study was used to analyze the pattern of care for hospitalized patients between the ages of 18 and 89, using the National Inpatient Sample (NIS) data of the Healthcare Cost and Utilization Project (HCUP). The study sample consisted of 500,000 admission records and stratified and regression analysis were computed to determine the differences by age, sex, race or ethnicity, income, location, diagnoses, procedures, length of stay, payer, and costs affecting each of the defined income categories. Total hospital costs were examined within the categorical income groups by residential zip code and top 10 diagnoses and procedures showed that high medical costs is an issue across SES groups. Descriptive and inferential statistical analyses were performed. Mean, median, standard deviation, and range were used to calculate continuous variables while frequency counts and chi-square tests of association were conducted to evaluate differences in proportion for categorical variables. Linear regression modeling and multivariable modeling techniques were undertaken to test the hypotheses. Measurement and structural models were tested through structural equation modeling statistical techniques using SPSS version 22.0. When compared the SES differences among the four categorical income groups, the results show that people at the lower quintile were more likely to face higher hospitalization due to their income. Each year, many programs are designed to reduce hospital admissions, but regardless of these efforts the rates of hospitalization continue to increase in U.S population. This study recommended scientific approach in understanding of the role SES and income as they impact health disparities, which will potentially help health providers, researchers, policy makers, and public health planners to design individualized and community-wide programs and policies related to income inequality and hospitalization for high risk populations.

KEYWORDS: Health disparities; Health inequality; Socioeconomic status; Environmental factors; Behavioral factors; Access to health care; Health resources

ACKNOWLEGEMENTS

I gratefully acknowledge the support of and patient of my Ph.D. committee members, especially Dr. Shankar Srinivasan, my advisor, who untiringly provided immense encouragement and extensive feedback throughout the journey of my Ph.D. engagement. I also thank Dr. Jared Sender and Dr. Charles Perry who provided incredible support and feedback to my study.

I am most grateful to my loving husband, Morris and my daughter, Kimberly for their endless prayers that I would complete my Ph.D. and bearing with me when I say to them I cannot make dinner due to my immense research activities.

Finally, I was lucky to have many supportive friends, especially Jean Paul as well as the members of my congregation. I am also thankful to my co-worker, Yi-Ting Chiang, who endlessly edited and provided feedback on my dissertation. I am really fortunate to work with her. The invaluable support I received while pursuing my Ph.D. was so great, I am certain that I shall reciprocate by helping others in need throughout my life.

DEDICATION

To my late mother, Mrs. Violet Jumbo

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF GRAPHS

LIST OF EQUATIONS

CHAPTER I INTRODUCTION

For years, public health statistics have been routinely collected and reported for groups defined by social class, generally measured by ranking according to occupational hierarchies reflecting differences in social standing.[1-4] Despite the ample health programs implemented to reduce health disparities, hospital admission and readmission, and social economics status (SES) nevertheless continues to impact hospitalization rates in the U.S. The purpose of this dissertation was to examine the association of socioeconomic status by median household income groups with hospitalization outcomes in the United States from 2008 to 2010.

Current health research in the U.S increasingly recognizes the existence of a robust, linear relationship between socioeconomic status (SES) and health and that SES is often measured as a combination of income, education, occupation and other broader determinants such as environmental exposure, behavioral factors. SES as it is known is commonly conceptualized as the social standing or class of an individual or group. Also, research suggests that "resilient psychosocial factors also relate directly with health" and that "stress combined with low resources can generate further stress and resource deficiencies, fostering spiraling stress and losses".[5] Thus, such health gradients are problematic in health care system, therefore, indicating potential problems with unequal access to care. And in this study, one of the determinants of SES, income (categorical variables by zip code) of patients' median income is used to obtain hospital admissions information to study the pattern of patient hospitalization as it impacts the four income groups.

Researchers have investigated the relationship between SES and health and reported that health differences across groups defined by components of SES (income, education, and occupation) as well as the impacts on hospitalization characteristics across the U.S have been examined less frequently and that nonetheless, improvements are not consistent across all population. Also, when differences in income and education were reported, the number of groups being compared has often been limited to two or at most three, or to common morbidities such as cancer, HIV, asthma, low birth weight, cardiovascular diseases, hypertension, arthritis, and diabetes and mortalities, and yielding conflicting results even when measured with various SES determinants. So it is important to compare the socioeconomic factors against health gradient across these categorical income groups and to examine the associated hospital costs on each group. In health care, a small proportion of patients tend to account for a majority of the care and costs, often known as the 80/20 principle[6] and SES as a gradient or continuous variable reveals inequities in access to and distribution of resources.[7] Sociodemographic factors, such as age, race, health insurance, length of stay, types of morbidities when analyzed showed the association with hospitalization costs and the lower income group contributed strongly to the total inpatient care and costs. In retrospect, there have been policies, financial incentive, and numerous health programs that have been implemented for health interventions and disease management programs, but some have proven ineffective in the reduction of the rates of hospitalization. The high morbidity and mortality associated with SES, income disparities with hospitalization make understanding its epidemiology a research priority.

1.1 Background

The significant impact of SES on disease and its complexity make its definition and measurement of critical importance. SES as indicated is a complex phenomenon predicted by a broad spectrum of factors that is often conceptualized as a combination of financial, occupational, and educational influences.[8,9] However, recent evidence suggests that treating SES as a unified concept is inappropriate. SES has not one but many dimensions, which relate to health in diverse ways.[10] The measures of SES operate through different mechanisms and that it is useful to examine these mechanisms independently.

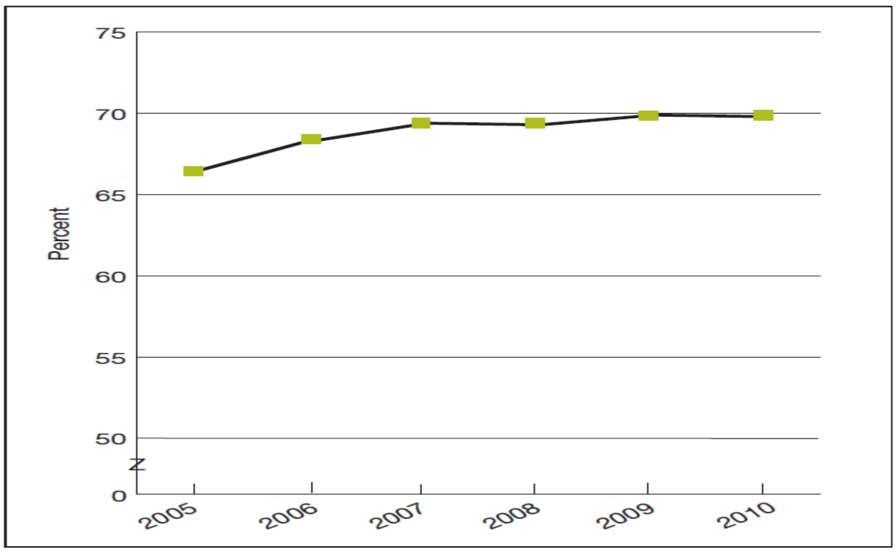


Figure 1 Average proportion of recommended care received across a panel of quality of care measures, 2005-2010
Source: National Healthcare Disparities Report, 2013

As the issues of universal coverage and health system reform become more apparent in the public health debates in the United States. Figure 1 above shows that Americans received about 66% of health care services they should have received, by 2010, the number has risen to 70% services, it is critical that policy-makers, researchers, and healthcare providers understand the effect of SES on a person's morbidity, mortality, and health care costs. It is especially important to understand how medically indigent patients use the healthcare system and how barriers to medical care can affect their health status and resource use. Myriad of studies suggest that many Americans are in poor health and do not get the best medical care they need. This lack of timely and quality care have resulted in higher hospitalization rates in the United States, especially among racial and ethnic minorities with low income. Various researches and reports also documented that minorities face "poorer health, experience more significant problems accessing care, are more likely to be uninsured, and often receive lower quality health care than other Americans".[11] Also, studies that compared variations in the rates of hospitalization among nations, regions, states, communities, and neighborhoods[12] have found that low income groups are faced with more health disadvantages than other groups. While a broad range of explanations has been offered, studies have concentrated on the impact of the differences in social determinants such as race or ethnicity, education, income, employment, immigration, housing, access to health care, neighborhood conditions, and common comorbidities, but have infrequently measured the indirect pathways (biological factors, environmental exposure, behavioral factors and lifestyles) by which SES influences hospitalization. However, with respect to disparities in health care whether it is due to clinical differences among race, cultural differences and their effect on behavior or decision making, or socioeconomic factors, studies have consistently demonstrated that disparities exist and are an integral factor in obtaining access to preventive care.

SES is an important source of health inequity, as there is a very robust correlation between SES and health. This correlation suggests that it is not only the poor who tend to be sick when everyone else is healthy, but that there is a continual SES gradient phenomenon from the top to the bottom of socioeconomic ladder, relating status to health. Specifically, lower SES has been linked to chronic stress, heart diseases, ulcers, type 2 diabetes, rheumatoid arthritis, certain types of cancers, and premature aging, among other conditions. Because of these reasons, there is a dramatic disparities between the rates of the lowest and highest social classes. Furthermore, the average life expectancies at birth in 2011 were 4.5 years shorter for black than white men and 3.1 years shorter for black than white women.[13] The Healthy People 2020 states that

> "disparities often is interpreted to mean racial or ethnic disparities, many dimensions of disparity exist in the US, particularly in health. If a health outcome is seen in a greater or lesser extent between populations, there is disparity. Race or ethnicity, sex, sexual identity, age, disability, socioeconomic status, and geographic location all contribute to an individual's ability to achieve good health. It is important to recognize the impact that social determinants have on health outcomes of specific populations."[14]

Health disparities have been investigated for decades, and although improvements of health care have led to advancement in the management of diverse diseases, the gaps between race or social classes have shown little changes, therefore, the goal of this research is to examine SES and the four categorical variables of median income from the US Census Bureau data linked to NIS that are associated with patient demographic and hospitalization characteristics.

Table 1 Categorical Variables (ZipInc) Quartile Median Income

Category 1	$1 – 38, 999
Category 2	$39,000 – 47, 999
Category 3	$48,000 – 62, 999
Category 4	$63,000 or more

Table 1 describes the income levels data associated with patients' zip codes, which are combined with clinical, and hospitalization characteristics such as discharges, mortalities, comorbidities, procedures, costs, payer types, length of stay outcomes across gender, race or ethnicity, across regions, or states, and hospitals types (urban, rural, teaching, non-teaching, and size). Nevertheless, in many studies, the results indicated that patients of lower income were generally associated with longer stays and higher hospital charges than care for those of higher income levels due to poverty and unemployment status. A large German study that engaged in a cross-sectional design found that "not only low-SES but, specifically, single parent families were strongly associated with the extent of psychosocial distress among children".[15] A study suggested that mentally ill individuals gravitate to low-income communities as a result of their disability, perhaps drawn by lower living costs. Nonetheless, the consistent findings of the researches favoring the various hypotheses mostly accept that no single interpretation of the relationship has been supported by the available data.

1.2 Goals and Objectives

The overall goal of the study was to compare the four categories of income groups (ZipInc median household income) and SES associated with demographic and hospitalization characteristics in the U.S. from 2008 to 2010. The research objectives of the project are the following:

- Identify the income categories vary by race or ethnicity, sex, and age
- Identify the total hospital costs incurred by the SES groups
- Identify clinical factors such as types of comorbidities and procedures that the groups encounter
- Identify the health insurance coverage amongst the different income groups
- Determine whether there are differences in the length of stay amongst the median income groups

1.3 Purpose of the Study

The purpose of this quantitative retrospective study was to examine the association of SES by median household income groups with hospitalization outcomes of patients' admission records across the U.S. community hospitals between 2008 and 2010. The study assesses the total hospital costs derived from hospital charges across the income levels which were used to outline the resource implications and opportunities for improving outcomes hospitalizations by patients within the lower income group. The study compared the causal directions, generalizability across the U.S. adult populations and diseases, and relations with health for different indicators of SES. The study results can help to describe the healthcare related issues that plague the people, disparities in the regionalization of care access, utilization, determine quality of care, and suggested solutions that ensure equal healthcare delivery and aid in the evaluation of medical costs, length of stay, and the effectiveness of procedural provision. The study contributed new knowledge to the research body, also provided workable information to policy makers and healthcare providers to better mitigate the health disparity issues that exist among the racial or ethnic groups with different incomes, especially within the low income groups. Finally, we hope that the study and the discussion provided information for future research expansion on the improvement of health and healthcare costs of all Americans. This study examines the patterns of socioeconomic differences in a wide array of important health indicators in the United States. This study also examines the contributors to hospital utilization. Therefore, by understanding the influence of hospitalization and health needs in a systematic manner, the study seeks to help decision makers to implement social and economic policies that will impact access to care resources.

The results either agrees or disagrees with the results of other studies and the findings can help biomedical informatics, policy makers, healthcare practitioners, and researchers to understand the health issues that the American people face and the contributors of disparities in hospital use and formulate solutions that will help to alleviate the ongoing national health challenges at both individual and societal levels.

1.4 Need and Rationale

A survey of the literature reveals that there are many studies on SES associated with patients' demographics and hospital characteristics. Nonetheless, improvements in the health care access for low income group are not consistent across the population strata. For instance, reduction in cancer deaths in persons with low SES and in African Americans remains slow and such lack of improvement in selected segments of the US population may suggest underlying disparities in the access to care.[11 12] In addition; studies report that health status improves as SES increases, however, the underlying mechanisms of the association are not properly clear. Given these findings and the critical information that is still missing, this study seeks to assist in unveiling some of these mechanisms.

1.5 Significance of the Problem

A survey of literature reveals that although there are many studies that relate to this topic, the concentration is mostly on race or ethnicity, comparing certain morbidities and mortality. Also, many of the early publications regarding the effect of SES and hospitalization may be outdated. This study may provide significant contribution to the healthcare industry.

There have been dissertations, books, and articles that touch on some portions of this present study, but none of those reviewed encompass the proposed scope of this dissertation. Finally, it is hoped that the results of this study will make a scholarly contribution to the body of knowledge and may allow health care decision makers in addressing and implementing adequate social and economic policies geared toward improving low income groups' access to needed health resources.

1.6 Research Question and Hypothesis

1.6.1 This research poses several questions and the research questions were:

1.6.1.1 How does socioeconomic status, median household income vary by patient demographic characteristics, gender and age?

1.6.1.2 What are the groups of patients in the four categorical median household income and cost? Is there significant difference between the four categorical median household income and hospital charge in the 2008-2010 discharge records?

1.6.1.3 Which clinical factors, such as the types of comorbidities and procedures that affect median income groups most prevalently? Is there significant difference among the median income groups?

1.6.1.4 What are the types of health insurance coverage the groups have?

1.6.1.5 Are there differences in the cost and length of stay amongst the different median income groups?

1.6.2 The hypotheses investigated in this study were:

- $H_0$1: There is no statistically significant difference between the four categories of the median income groups and hospitalized patients' gender and age in the 2008-2010 NIS data set.
- H_a1: There is statistically significant difference between the four categories of the median income groups and hospitalized patients' gender and age in the 2008-2010 NIS data set.
- $H_0$2: There is no statistically significant difference between the categorical median income groups and hospital charge in the 2008–2010 discharge records.

- H_a2: There is statistically significant difference between the categorical median income groups and hospital charge in the 2008–2010 discharge records.
- H_03: There is no statistically significant difference in the prevalence of clinical factors, such as the types of comorbidities and procedure, among median income the groups.
- H_a3: There is statistically significant difference in the prevalence of clinical factors, such as the types of comorbidities and procedure, among the median income groups.
- H_04: There is no statistically significant difference in the types of health insurance coverage the groups have.
- H_a4: There is statistically significant difference in the types of health insurance coverage the groups have.
- H_05: There is no statistically significant difference in the costs (TOTCHG) and length of stay (LOS) amongst the different median income groups.
- H_a5: There is statistically significant difference in the costs (TOTCHG) and length of stay (LOS) amongst the different median income groups.

1.7 Theoretical Framework Perspective

In this study, the conceptual framework developed by[65] helps to outline a conceptual guide in choosing the model and to study, a mechanism linking SES to hospitalization. This framework presents findings from a review of the literature demonstrating the characteristics related to SES and health. Researchers report that no single mechanism accounts for the SES and health, but one consistent outcome is that education, occupation and income are associated with SES on the high end of the spectrum. Therefore,

Figure 2 presents a simplified model of pathways through which SES affect health. This framework depicts how socioeconomic factors interact to affect adult health. Consequently, even though this model does not include all indirect pathways, it shows that within each layer, as SES declines, demands increase, and resources for dealing with these demands decrease. Therefore, this dissertation uses this model to show that income is one of the strong predictors of SES which greatly affects the low income group. The result of this study will either agree or disagree with this theory that the determinants of (income, education, occupation, and other indirect pathways, environmental exposure, biological, and behavioral) SES are association with hospitalization.

Particularly, the study focuses on using categorical income to examine the relationship between SES and hospitalization. Myriad of studies confirm that people with lower SES are exposed to more stress than people with higher SES. When these individuals can no longer response to these environmental challenges, they become more vulnerable to diseases or chronic conditions.[18]

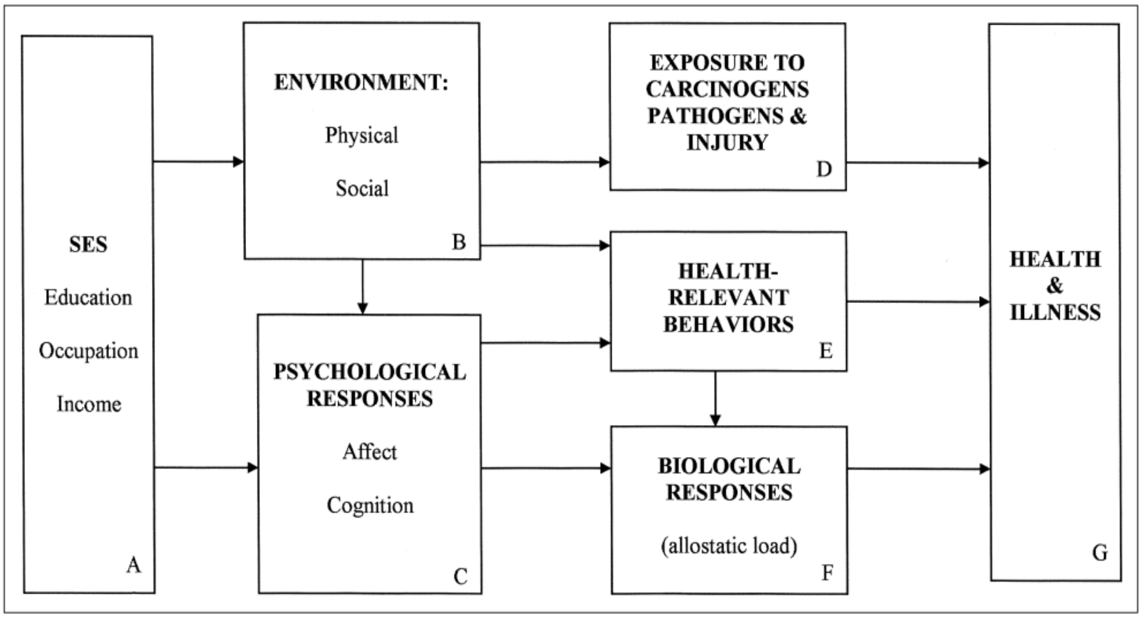


Figure 2 Simplified model of pathways from socioeconomic status to health
Source: Adler and Ostrove (1999)

1.8 Intended Results

The projected findings from the statistical analysis will show that low income groups are associated with high hospitalization between 2008 and 2010. The study intends to demonstrate that age, gender, race or ethnicity significantly affect hospitalization and costs in the US. Ultimately, lower income status has been linked to chronic diseases as well as made frequent comments about the deficiencies of the health care systems; therefore, the study will show:

- The counts and frequency of hospital use by age, gender, and race using the 2008-2010 NIS dataset
- The medical diagnosis and procedures are significantly effective to determine the hospitalization characteristics
- Relations between patients' demographic, clinical characteristics, and hospitalization characteristics can be predicted
- Lack of proper education can affect hospitalization, therefore, impact costs

CHAPTER II REVIEW OF LITERATURE

2.1 Introduction

In chapter 1, we presented the goal and objectives of the study. This chapter presents the reviews on the exiting literatures to determine the concept and the essential components of SES, examining its association with income and hospitalization characteristics in the United States. In this review, we start with the general concept of socioeconomic status, particularly as it relates to health disparities, and continue to examine the mechanisms of which SES correlates with a person's income, which plays significant role in hospitalization. We examine the generality of SES findings and the conceptual frameworks and theories that have been identified by researchers. Based on the literature review findings, we extend and refine prior work and present our findings, gaps, and conclusions.

2.2 Socioeconomic Status and Health

Health disparities have become a major concern in the U.S and globally. The important of socioeconomic conditions for health was studied extensively in the 19th century, William Farr investigated this issue for 40 years to document the socioeconomic differences in diseases in England.[16] Differences in health status among ethnic or social groups are pronounced. However, one of the causes of this health gap is the lack of timely, appropriate ambulatory care navigation which may result in bad health conditions that require hospitalization. A significant body of research has been dedicated in investigating and explaining the disparities between SES and health for decades now. Although, there is disagreement about what constitute health disparities or inequalities[17-19], findings from different studies indicate that recent progress has been made in characterizing the relationship between the various measures of SES and health over a person's lifetime and in understanding the relative importance and the causal pathways. Studies also suggest that SES typically is measured based on educational attainment; occupational characteristics; and income. Other studies, however, argue that apart from the core measures of SES other factors such as behavioral, biological, environmental, accumulated wealth, and health insurance interact to influence SES and health status among populations, resulting in health disparities.

In the U.S. populations, SES is frequently implicated as a contributor to the disparities in health among various causal pathways.[20] Researchers also cited that these indirect factors that contribute to SES; lifestyle, cultural, environmental, biological, behavioral, living and working conditions, social, and community networks have generally received less attention. SES is a commonly used concept in health research, and results from studies show that there are numerous ways of measuring the complexity of the construct. A variety of other terms, such as social class, social stratification, social, socioeconomic position or socioeconomic status are often used interchangeably[22, 22] to define SES. As it is defined, socioeconomic status refers to the social and economic equity that influence what positions individuals or groups hold within the structure of a society, and encompasses concepts with different historical and disciplinary origins.[23] There is a growing evidence that describes SES as a combination of education, income or wealth, and occupation and that is commonly conceptualized as the social standing or class of an individual or group in a society.

The question of whether SES is seen as gradient or continuous variables that reveal inequities in access to and distribution of care resources derives more studies. So, there is little doubt that more research into the association of SES and hospitalization would assist any policy decision in the matter. Regardless of how SES is defined, the challenges that we face today are the "inequities in wealth and quality of life that are increasing."[2, 7] According to a study, it is "unquestionable that government plays a major role in the gains in health care access that have occurred in the last half century, yet today, all Americans do not have the same opportunities for health, access to care, or quality of care when they receive it."[24] Some researchers are of the opinion that despite calls to distribute resources across all income levels, SES continues to be one of the contributors to "health status disparities (differences in life expectancy) and health care disparities (differences in coverage and quality of care)"[24] in the U.S. A team of researchers presented three causal pathways through which SES impacts health, which include its association with healthcare, environmental exposure, and health behaviors and lifestyles. Together, these pathways are estimated to account for up to 80% of premature mortality.[25]

Current literatures indicate that while it is true that the U.S. has seen incredible progress in health in the last century because of the intervention of the public health policies and programs; the benefit of these achievements have not been distributed evenly across different population groups. There are still disparities in the health of the people, especially among the lower SES groups. Therefore, studies that would help shed light on how to bridge this gap in health access are essential.

In addition to the core mechanisms of SES, many studies have also shown that the role of bias, miscommunication, lack of trust, and financial and access barriers are factors that encourage health disparities in the U.S. Also, ample studies have confirmed that bad health behaviors such as alcohol abuse, lack of exercise, cigarette smoking, psychosocial and environmental risk factors, and biological factors are considered the indirect pathways to SES. It is suggested that a study be undertaken to identify solutions to these barriers to the population health. Another study suggested that disparities may be a function of the overall performance of the health system where an individual lives. Furthermore, to better understand SES, researchers have introduced several concepts in classifying SES and health:

- **Social Determinants** of health are factors beyond our genetic make-up and our access to medical care, including social status, employment and income, education, housing and neighborhoods, and access to nutritious foods.
- **Health disparities** are differences in the incidence, prevalence, mortality, and burden of diseases and other adverse health conditions that exist among specific population groups.
- **Health inequities** are health differences that are directly related to social inequities, which are systematic, socially produced (and therefore modifiable), and unfair.
- **Health equity** is the highest level of health for all people, which requires the absence of disparities and is therefore an issue of fairness and justice with far-reaching implications in our society.

However, the question of just how universal health disparity is has increased the number of cross-cultural investigations of adults' health status and determinants, utilization of health resources, health care resources, and health expenditure in hospitalization in the U.S. Despite these challenges, increasing innovation in the application of technologies, intervention programs, and policies have simplified the access of care to those that were underprivileged and unable to afford it in the past. As the healthcare system is becoming more integrated with the help of technology, the utilization of care becomes more transparent and accessible. Even though the integration and consolidation of data across the health facilities are not complete, we are able to use technology and clinical data for analysis and provide trends in biomedical informatics statistics, genetic manipulations that can help advance the provision of care and reduce health disparities. In more recent studies, SES is defined as "a broad concept that refers to the placement of persons, families, households and census tracts or other aggregates with respect to the capacity to create or consume goods that are valued in our society."[26]

2.3 SES and Characteristics of Health Care Utilization

Disparities should be examined in all aspects of health care, including the allocation of resources for care, utilization of services, care quality, and the associated cost. With that said, city and state hospital discharge data have been used to show trend in hospital utilization by the socioeconomic positions, however, middle and lower income groups are less likely to receive the appropriate care and are more likely to encounter delays in their care or have an irregular source of care. A research study suggests a method in evaluating the utility of various individual determinants of health services; a theoretical framework for health services utilization which presents the characteristics of the health services delivery system, changes in medical technology and social norm, and individual determinants of utilization.[27] The framework demonstrates the factors in achieving equitable distribution of health care services. Prior studies emphasize that these determinants of medical care utilization are the results of the societal values and perceptions that people have a right to medical care regardless of their ability to pay for this care. However, despite this belief, certain population groups such as the poor or the underserved and inner city and rural residents are not getting the same quality care compared to the rest of the population groups. For example, in the aftermath of Hurricane Katrina in New Orleans, the city was deeply divided along racial and class lines with large historical pockets of poverty and social problems experienced over multiple generations, the scenario clearly demonstrate social disparities in the city. This devastating event affected the poor much more deeply than other groups and more than five years later, many of them have yet to recover from the catastrophe. Therefore, many studies emphasized that the greatest threat to our health in the U.S. today is the large income gap. This problem can never be ignored or underestimated and it is necessary to evaluate this large income gap. Researchers lament that although racial inequality is a significant source of health inequity, socioeconomic status or class may have a greater influence, the rich are healthier than the middle class, and the middle class are healthier than the poor and can navigator care better.

Research shows that from 1979 to 2007, the income gap tripled between the richest 1% of Americans and the middle and poorest fifths of the country and because of the deterioration of health among the poor appears to offset any gains in health experienced by the rich. In addition, researchers also mentioned that the recent economic downturn in the recession of 2007 to 2008 might even increase health inequalities further. Figure 3 below describes the relationship of the major components of the framework for viewing health services; it defines the unit of health service utilization.

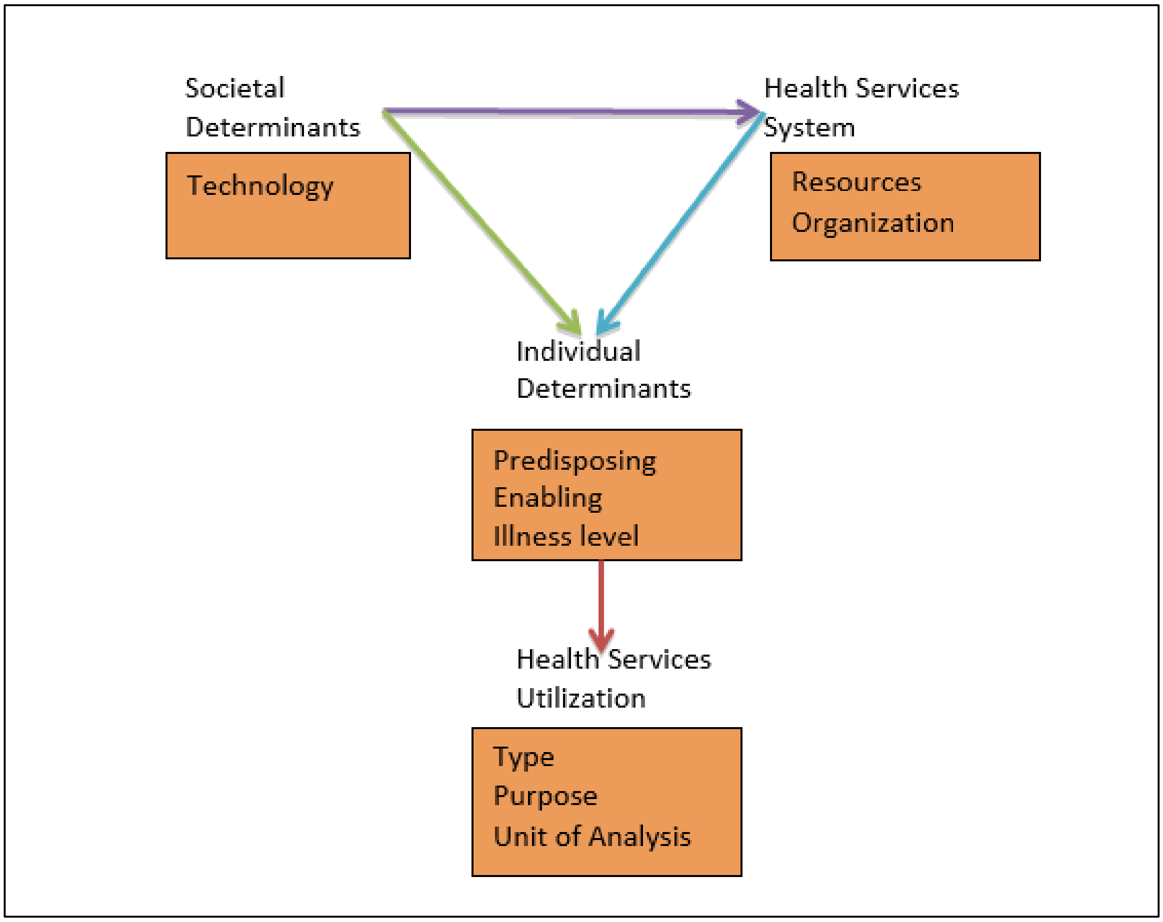


Figure 3 Framework for View Health Services Utilization
Source: Adler. N, Newman K, "Socioeconomic Disparities In Health: Pathways And Policies

Based on Figure 3, the utilization of health services are viewed as a function of the characteristics of the individual behavior, characteristics of the environment where he or she lives, and interaction of societal forces. To date, most empirical studies and theories dealing with health services utilization emphasized the individual characteristics while less attention has been given to the societal impact.[28] Ultimately, the characteristics of prime importance are the type (hospital, physician, drugs and medications, dentist, nursing home, and other), purpose (primary care, secondary care, tertiary care, and custodial care), and unit of analysis (contact, volume, episodic care). From the Figure 3 health services system is defined as the resources available to patients and the organization is a process in which patients receive care, while the societal determinant of health utilization are technology and norms. The national health care system (resources and organization) helps to shape the provision of health care services to the individuals. The resources of the system are the labor and capital devoted to health care, while organization simply describes what the system does with it resources, the manner in which medical personnel and facilities are coordinated and controlled in the process of providing medical services. Research on traditional health care system and its navigation is a vital step towards providing knowledge which will strengthen health practices and extend support and services to the people of the U.S. However, the influences on health also include the living conditions, neighborhood, workplaces, and communities which are likely to impact health.

2.4 Mechanisms through which SES Influences Health

There has been many debates as to what constitutes the mechanisms of SES and in other to understand this concept, we provide the current literatures on the various components between health and SES.

2.4.1 Education

Generally, educational attainment is found to be one of the most commonly used predictors that have a causal effect on health and mortality [29, 30, 31, 32] in the United States; therefore impacts hospitalization rates. In addition, researchers note that both education and income are strongly related to most measures of health and health-related behaviors across the span of life. Also, there is an argument that the mechanisms to which education affects health are unclear and that while consumption, behavior and curative, and preventive care can partly explain the effect of education on health, it remains largely unclear why more educated individuals behave in a heathier manner.[33] So, it is highly necessary to regularly conduct research that will highlight the importance of the methods that impact SES and health. Further, many research findings agree that education opens doors to opportunities and resources that lead to a higher socioeconomic status or position in a society. In general, more education is seen to be associated with higher-paying jobs and the related benefits such as financial stability, health insurance protection and investment, better working conditions, and social connections. Researchers also emphasize that compliance to prescribed treatment and managing diseases is greater in more educated and affluent patients than their counterparts, and may reduce the risk of comorbidity and the need for hospital care.[34 35] Finally, most studies indicated that low birth weight individuals have been associated with worse schooling outcomes.

2.4.2 Income

Empirical results from studies generally support that income is the predictors of SES that mostly measures the material resources component as with other indicators such as education has association with health,[36, 37] and can influence a wide range of material circumstances with direct implications for health.[38, 39] However, studies reveal that those with higher income levels are more likely to live longer, have healthier lives, and have less comorbidity. It has been reported that higher income leads to resources that promotes good health, nutritious food, safe housing, and nurturing neighborhoods. While, on the other hand, those living in poverty face numerous hardships that ultimately lead to poor health or even mortality. Also, researchers postulate that poverty has a significant impact on children across generations, although children are not included in this study, but it worth mentioning. Generally, poverty breeds poor health and therefore affects the ability for people to do well academically and eventually earn a decent wage.

	<$24K (low income)	$24K-<$90K (middle class)	$90K+ (high income)	Gap, low vs. high income
Well-Being Index score	57.2	67.7	74.3	17.1

Figure 4 Well-Being Index Scores, by Income Group (Annual Income)
Source: Gallup-Healthways Well-Being Index – January 1-September 28, 2010

Numerous studies mention that income is one of the major contributors to health and health care disparities in the U.S. As shown on Figure 4 above, the income gap between the low verses the high is significant. Based on the positive relationship between income and health, it is assumable that the United States is the richest country in the world today, and would have the best health status in the world. Yet, the U.S. face high income disparity. Most people within the low income groups are unable to purchase health care coverage and some not having health insurance at all. There is a consensus that higher wages are associated with greater opportunity costs, hence, patients with higher income can afford better nutrition, housing, schooling, recreation, and live in better community, which would reduce the amount of time devoted to health maintenance and ultimately increase hospitalization needs.

The underinvestment in public infrastructures and welfare and the experience of inequality are both greater in more stratified societies and that these, in turn, affect health. Numerous studies also found that regardless of race, patients with annual income of lower $45000 or lower were more frequently treated at the hospitals, and certain diseases were treated more frequently for whites than blacks and other race. Despite policies implemented in the U.S. to help reduce economic inequality, there are still huge income gaps between the population groups. Examples of policies are the progressive income tax in 1913, Social Security and welfare policies in the 1930s; all these policies were targeted toward bridge of income inequality.

Low SES has been shown to be strongly associated with excess hospitalization, and patients within this category are more likely to have exposure to toxic substances and are less likely to have access to appropriate medical care.[25]

Additionally, income depends on individual's stock of knowledge which is based on years of schooling, years of working experiences, and life exposures. For instance, a study showed that Congestive Health Failure patients with lower income are known to experience a greater rate of hospitalization.[40] It is also believed that healthy people are more productive, earn higher wages, and are able to accrue wealth. But researchers also noted that attempting to investigate the role of categorical income groups and hospitalizations in respect to health may differ in hospital costs.

2.4.3 Occupation

So far the literature review on occupation indicate occupational status to be one of the major predictors of SES and a complex variable. The measurement of this determinant varies depending on one's point of view about the significance of various aspects of work life.

Growing evidence suggests the significance of whether or not one is employed, since the employed have better health condition than the unemployed.[41] There are many indications that unemployment and the length of unemployment affect health status and this status is a reflection of a person's social standing, income and intellect. At the same time, some types of benefits for the unemployed can buffer the adverse effects on health. A study shows that entitlement benefits appear to reduce some negative health effects, while means-tested benefits do not.[42] As supported in current literature, lower status jobs expose workers to both physical and psychosocial risks, for example transport driver and laborer.[43] Further, more affluent workers may choose safer working (associated with a lower level of job-related health stress) and living environments since safety is a normal good.[44, 45] Occupation may reflect social networks, work based stress, control, and autonomy and thereby affects health outcomes through psychosocial processes.

2.5 Indirect Methods of Assessing Factors that Influence Variations in SES and Health

A variety of indirect methods have been used to assess the contribution of SES to Health differences in the disease prevalence in the U.S. population groups. The current literatures on health care disparities discovered that gaps in life expectancy by racial or ethnic groups in recent years may have narrowed, but racial disparities in other health related measures have continued over time. In health outcomes, such as maternal mortality, this gap has increased. Some researchers advise that there may be differences in how people of different racial or ethnic groups respond to treatment, particularly to that of pharmaceutical interventions, and the variations in these forms of treatment may be justified on the basis of patient race or ethnicity. In addition, minorities are less likely than whites to seek for help and more likely to avoid or delay seeking for appropriate care. It was also found that racial differences in patients' attitudes, including preferences for treatment, do not differ by race or ethnicity; however, researchers were unsure if people are refuse medical care due to their general perceptions of healthcare providers, negative encounters, or mistrust in their doctors' lack of concern for their health. More study is needed to explore and understand treatment refusal. To further describe these methods, researchers summarized the following areas as the indirect mechanisms that impact health disparity.

2.5.1 Environmental Factors

SES is associated with the physical and social environments in which people work and live. Unfortunately, some individuals with lower socioeconomic resources tend to inhabit environments with higher levels of pathogens, toxins, noise, and occupations that subject them to more physical risks. These environments may lack certain amenities, including recreational facilities, operation of healthcare, the legal and regulatory systems, healthy foods, and social support. Also such environments are more prone to social conflict, crowding, and crime. Generally, people that live in these environments have decreased access to means for restoring and maintaining health. However, studies show that racial or ethnic minorities receive lower quality healthcare than whites, even when they are insured to the same degree and when other healthcare access-related factors, such as the ability to pay for care are the same.

2.5.2 Behavioral Risk Factors

Over the past several decades, health behaviors such as smoking, alcoholism, and leisure-time physical activity have been indicated as the major contributors of premature morbidity and mortality.[46] Lately, these behavioral differences in health outcomes by SES have been acknowledged as a persisting and increasing health-damaging and public health problem.[46] For example, smoking is now known to be a major risk factor for several diseases such as cancer, chronic bronchitis, emphysema, and cardiovascular disease. According to the U.S. Department of Health and Human Services, alcoholism is an important risk factor for numerous health outcomes, including cirrhosis of the liver and pancreatitis. Many studies acknowledge that alcohol is also a contributing factor in approximately half of all homicides, suicides, and motor vehicle fatalities. Similarly, regular physical activity and correct nutrition have been shown to lower one’s overall risk of mortality as well as being linked to reducing the risk of certain diseases such as cardiovascular disease, non-insulin-dependent diabetes, osteoarthritis, and depression. In addition, a study shows that racial and ethnic groups differ on these behaviors; see Table 2, the study used the data from the 2000 Behavioral Risk Factor Surveillance System, comparing rates of smoking, obesity, leisure-time physical activity, and alcohol consumption. On the other hand, a study mentions that racial or ethnic groups differ on these behaviors and are found to be inconsistent between the groups because the relationship between health risk behaviors and SES is still unclear. But behaviors earlier in life predict later behavior and may have long-term effects not fully accounted for.[47] As shown in the Figure 5 below behavior, cigarette smoking and educational level impact different age groups. The Figure 5 shows that education is consistently associated with risk factors especially among the lower SES groups.

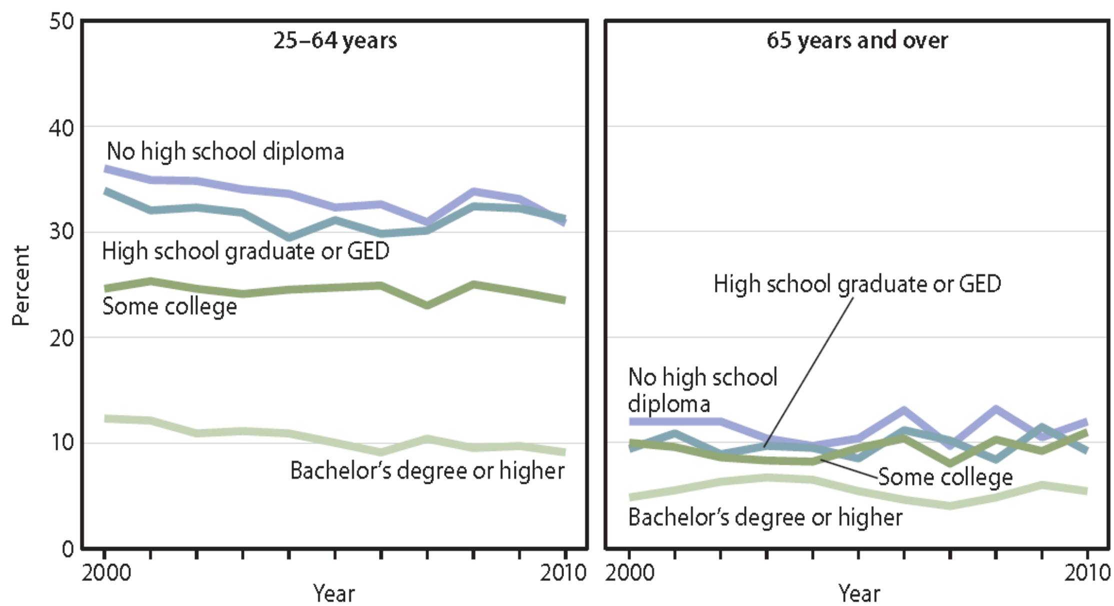


Figure 5 Current Cigarette Smoking among Adults 25 years of Age and Over, by Age and Educational Level: United States, 2000-2010
Source: CDC/NCHS, National Health Interview Survey

In a number of longitudinal studies, important SES determinants such as income and education have been shown to be inversely associated with various mortality outcomes, including premature mortality, cardiovascular mortality, and death from all causes.[46] Moreover, the low-income groups are significantly more likely to lead a sedentary lifestyles, smoke excessively, consume alcohol and be obese compared to the high income groups. These behaviors are among the major risk factors that are associated with health differences and high cost in the U.S. As shown in Table 2, socioeconomic disadvantage is one reason for differences in behavior risk factors, and less education is associated with more smoking,

Table 2 Self-Reported Health Risk Behaviors, by Race and Ethnicity

Behavior	White	Black	Hispanic	Asian	American Indian/ Alaska Native
Current Cigarette Smoking					
Median percent[a]	23.6	22.8	23.1	10.7	41.3
Percentage of states where group median exceeds white median	—	46.0	47.0	11.0	82.0
States included	—	34	35	9	11
No Leisure-Time Physical Activity in Last 30 Days					
Median percent[a]	25.1	38.2	34.2	28.9	37.2
Percentage of states where group median exceeds white median	—	92.0	81.0	100	100
States included	—	36	35	6	5
Obesity[b]					
Median percent[a]	15.6	26.4	18.2	4.8	30.1
Percentage of states where group median exceeds white median	—	97.0	73.0	20.0	91.0
States included	—	35	35	10	11
Binge Drinking[c]					
Median percent[a]	14.5	8.7	16.2	6.7	18.9
Percentage of states where group median exceeds white median	—	11.0	66.0	25.0	91.0
States included	—	35	35	8	11

*a*Median percentages across states; fewer states are represented for minorities, particularly the last two groups, because of small samples.
*b*Body mass index of 30 kg/m2 or higher.
*c*Five or more alcoholic drinks at least once in past month.
SOURCE: Data from Behavioral Risk Factor Surveillance System, 1997.

and less physical activity, and that controlling for education and income reduces the apparent behavioral disadvantages among the low income groups. However, education and income do not increase any advantage in lower alcohol consumption.

Researchers also indicate that culture is another source of behavioral differences, the degree of acculturation among the diverse groups in the U.S. The more acculturated an individual the more prone the person is to smoking and obesity, [48, 49] and again that groups that are impacted more are the lower SES groups. Various cultural beliefs have also been proposed as important factors.

An additional factor that may impact behavioral differences is the residential environment. Researchers have reported that neighborhoods provide stimuli, such as outlets for alcohol or illegal substances, limit options for healthy nutrition, facilities for leisure or exercise, or insufficient supplies of prescription drugs.[50, 51] As such the lower income groups are more likely to suffer from such institutional risks factors.

2.5.3 BioBehavioral Risk Factors

Findings from numerous studies have challenged various aspects of the relationships among stress, psychosocial risk and resilience, and racial and ethnic differences in health. And have argued that for many years, stress have been considered a potentially important component connecting an individual's experience, living and working conditions, interpersonal relations, and other behavioral variables to biological factors that more directly influence health. Stress is a known risk factor for hypertension, and the significantly higher prevalence observed in low-income groups. Some emerging studies focus on possible biobehavioral mechanisms that quantify health disparities. The important aspect of the new effort is to understand the relationship between environmental and behavioral challenges and stressors, health, and diseases.

2.5.4 Access to Healthcare

Studies examine the extent to which the spatial distribution of health care and health outcomes across the United States varies by hospital referral region. Differential health care access and quality is a major public policy challenge. The United States is the only developed country in the world that does not have national health coverage, and 40 million Americans do not have any form of health care coverage. The general consensus is that, if society adopts better use of the ambulatory care then the need for hospitalization will reduce drastically and would reduce health disparity in the U.S.

2.5.4.1 SES and Clinical Uncertainty

Physicians rely on inferences about the health condition of patients, which contributes to disparities in treatment. This helps to create uncertainty because it is "based on what they can see about the illness and on what else they observe about the patient (e.g., race)".[52] The doctors are therefore seen to be working with previous beliefs about the patients' conditions that will be different according to age, gender, socioeconomic status, and race or ethnicity. However, despite doctor's uncertainties of patients' medical conditions patients must choose whether to undergo the recommended treatment and physician must decide what test or procedure to perform. If the physician has difficulty accurately understanding the symptoms then he or she is more likely to place greater weight on prior findings to make diagnostic decisions.

2.5.4.2 SES and the Nature of Stereotypes

In the literature findings, stereotyping was defined as the process by which people use social categories e.g., race, sex in acquiring processing, and recalling information about others. The beliefs (stereotypes) and general orientations (attitudes) that people bring to their interactions help organize and simplify complex or uncertain situations and give perceivers greater confidence in their ability to understand a situation and respond in efficient and effective ways.

2.5.4.3 SES and Healthcare Provider Prejudice or Bias

Findings from different studies reveal that prejudice is defined by psychology as an unjustified negative attitude based on a person's group membership. There has been evidence that most white Americans also have the notion that particularly African Americans are less intelligent, more prone to criminal activities, and prefer to live off of social welfare. However, it is morally abhorrent that healthcare provider engages in such practice because they are required to maintain professional values, but in most cases they may not recognize manifestations of prejudice in their behavior. Further, studies have so far identify direct evidence that providers' bias affect the quality of care for minority or lower income patients, research suggest that healthcare providers' diagnostic and treatment decisions, as well as their feelings about patients are influenced by patient's race or ethnicity. For instance, physicians referred white male, black male and white female and would recommend some treatments to the patients depending on their race.

2.6 SES and Demographic Characteristics

2.6.1 The Role of Race / Ethnicity

A growing body of evidence shows dramatic differences in health outcomes among racial or ethnic groups in the United States and these differences spread across a wide range of important indication of health from the beginning of life through old age. Before diving deeper into the findings in the literature reviews, it is beneficial to mention one of the problems that was identified when examining race or ethnicity was that the definitions used by government agencies are explicitly not based on biological categories,[53] rather the system was developed to meet the political obligations of the Census. These labels "black", "white", "American Indian", and "Asian" are considered races, while "Hispanic" is a language or cultural grouping, and "Asian/Pacific Islanders" "American Indian" is often used to collapse many smaller groups. However, the good thing about these race categories is that they provided a clear pattern of common diseases in the U.S. racial or ethnic groups. Further, health disparities are mostly due to the differences in our life experiences and not to genetic differences. Unquestionably, the most consistent health disparities are generally observed among the blacks compared to the whites and this is believed to be due to a long history of racial discrimination and income inequality, which indicates how groups of individuals are seen and treated in a society associated with their socioeconomic status or class. Researchers claim that lower socioeconomic status translates into poorer health status through social isolation and unequal access to the resources that could help in bad health prevention. Therefore, the issues of race or ethnicity, socioeconomic status, and health are closely related.

Studies also examine the relationship between SES and race or ethnicity among whites, blacks, and other ethnic groups based on their income levels and found that low income groups are less likely to receive needed high quality care than the higher income groups, and the physical health problems more prevalent to them include obesity, diabetes, cancer, high blood pressure, cardiovascular diseases, mental illness, asthma, depression, headache, stroke, pneumonia/influenza, and infant mortality. Also differences in risk behaviors were found to be pronounced among the blacks due to income, educational levels and the community where they live.[54] For example, a study found that a baby born to a black mother is more than twice as likely, and an infant born to an American Indian or Alaska

Native mother almost 11/2 times as likely, to die before reaching his or her first birthday than a baby born to a white mother, an Asian, Pacific Islander or a Hispanic mother, (Figure 6).

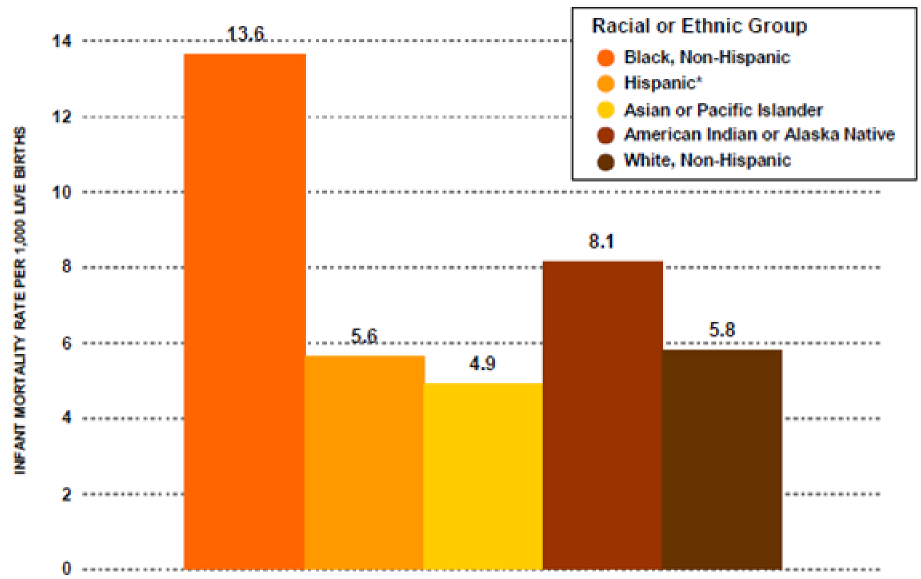


Figure 6 Black and American Indian or Alaska Native babies are much more likely than babies in other racial or ethnic groups to die in their first year of life.
Source: Mathews TJ & MacDorman MF. "Infant Mortality Statistic from the 2005 Period Linked Birth/Infant Death Data Set." National Vital Statistics, Vol 57, no 2. Hyattsville, MD: National Center for Health Statistics, 2008.

According to popular opinions, racial groups are seen as physically distinguishable populations that have a common origin. Although genetics and biology account for some aspects of variations in health status among racial and ethnic groups, social science research demonstrates the powerful influence on health of risk taking and preventive behavior, social and economic inequalities, communities and environments, health policy, and racist practices. These social dynamics play a major role in explaining racial and ethnic disparities in health outcomes. Even with a growing number of researches, we are reminded that race is not an absolute category but a social category subject to change.

2.6.2 SES and Age

Age is another sociodemographic factor that is consistently related to health status. Researchers indicate that age is one of the indirect contributing factors of SES and it is a more complex variable. Evidence relating SES to mortality among age range is less consistent, however, SES and mortality relation with age has been observed in most studies which compared older and younger populations. Socioeconomic differentials in mortality persist at older ages, but in others, no associations are found. For example, studies show that the death rate for black males and females are higher than of the whites. The pitfall to this point is that age is used as a covariate instead of studying this independently. The overall life expectancy for blacks is nearly 7 years shorter than for whites.[55] Research has found consistently that lower SES is related to worse physical health[56] and self-ratings of health.

2.6.3 SES and Gender

Studies indicated that females generally have more favorable health profiles than males across the ethnic and socioeconomic groups, with the exception of health outcomes that occur in an inordinately higher frequency in one gender group that is breast and prostate cancer.[57] There are also higher hypertension prevalence rates for females who are over 55 years over compared to males. Women who live in poverty, with low levels of education, work in low level jobs, and have minimal social supports and are more likely to suffer adverse birth outcomes than are higher SES women.

2.6.4 SES and Insurance Status

According to numerous studies, Blacks, Hispanics, and some Asian populations, when compared with whites, appear to have lower levels of health insurance coverage. Sadly, Hispanics are found to face greater challenges to health insurance compared to other groups.[58] In addition, ample studies have examine the relationship between hospitalization and insurance status for each race group and cite that regardless of race, those privately insured were more frequently treated than those with Medicare, Medicaid insurances. Figure

7 below shows that more low-income Americans are uninsured compared to the high income citizens.

	<$24K (low income)	$24K-<$90K (middle class)	$90K+ (high income)	Gap, low vs. high income (pct. pts.)
% Who have personal doctor	68.5	82.3	89.1	20.5
% Uninsured	33.4	13.4	4.7	28.7
% Not enough money for healthcare	35.5	16.4	6.4	29.0
% Easy to get medicine in your area	83.8	94.0	97.2	13.4
% Visited dentist	44.3	68.6	82.5	38.2

Figure 7 Access to Medical Care, by Income Group
Source: Gallup-Healthways Well-Being Index – January 1-September 28, 2010

Low SES patients more frequently receive care in nonoptimal organizational settings such as emergency rooms and lack continuity in health care. Also Blacks and Hispanics are less likely to have insurance coverage from private employers, whether directly or through a spouse, and are more likely to have public health coverage than whites.[59]

Recently, government data show that insurance premiums rose by 3% in 2013 while wages were increased by 2.3% in the same year, which further discourage the minorities groups to purchase insurance coverage. A study report shows that adults under age 65 with low family incomes are less likely to have insurance coverage than higher income adults.

2.6.5 SES and Cost

Research shows that not only does low SES impair health, but illness can in turn impose costs and reduce earnings and wealth. According to the National Health Expenditure Accounts (NHEA) health spending in the U.S. is nearly $2.5 trillion in 2009 as indicated in Figure 8, which reached an all-time high of $8,086 per person. NHEA indicate that this spending presents an almost two-fold increase since 1997, and this is due in large part to the decline in GDP as a result of the recession from 2007 to 2008. Researchers question who were the high-cost users and show that the spending was concentrated among a small number of persons. The spending distribution is consistent with the well-documented increase in population risk factors – most notably, obesity, hypertension, diabetes, trauma, and other treated chronic diseases. Findings in the literature review indicate that the lower income group have the higher likelihood of higher hospital utilization compared to the higher income group.

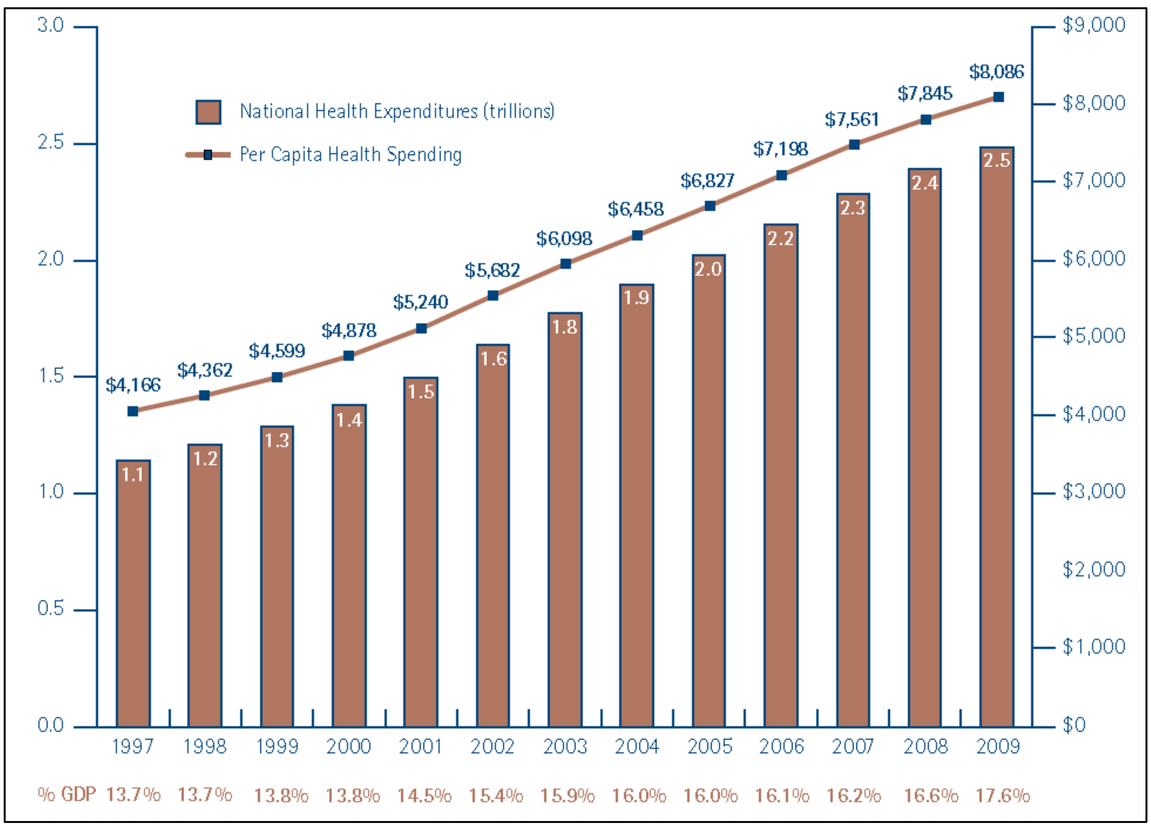


Figure 8 National Health Expenditures, Total, Per Capita, and as a Percent of GDP 1997-2009

Source: NIHCM Foundation analysis of data from the National Health Expenditure Accounts, available at https://www.cms.gov/NationalHealthExpendData/.

The Figure 9 below shows the aggregated hospital costs and stay by age in 2010, which shows that adults ages 45-64 and 65-84 have the highest hospital stay, and these two age groups summed up nearly two-thirds of the aggregate hospital costs and stay. While adults age 85 and older accounted for just 8 percent of aggregated costs and hospital stay. This report supports the notion that age is an indirect predictor of SES.

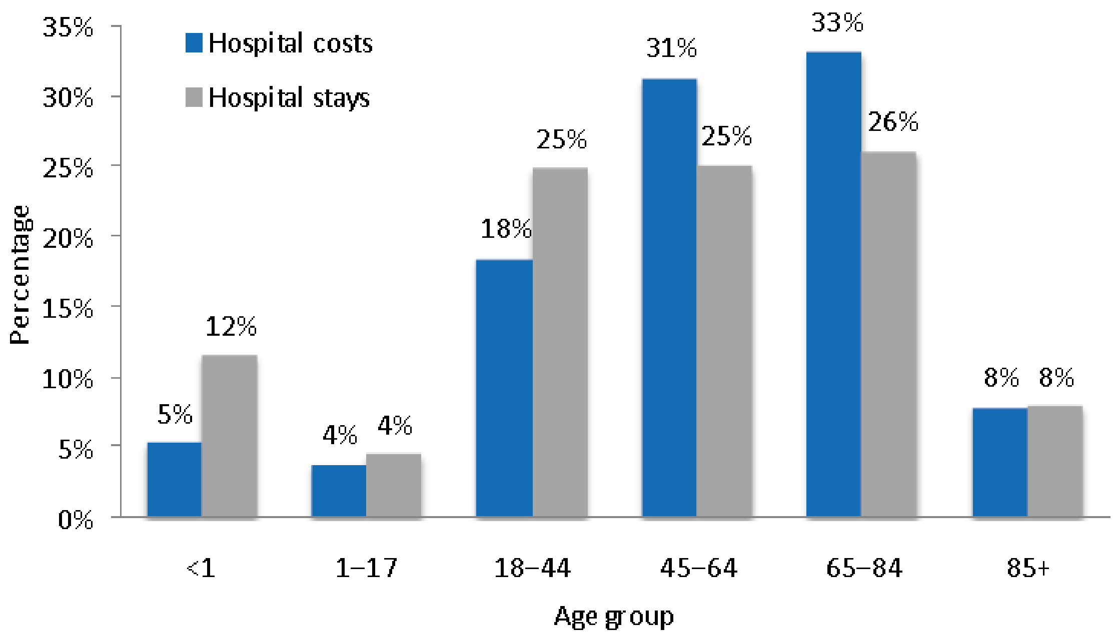


Figure 9 Distribution of Aggregate Hospital stay Costs and Stay by Age, 2010

Source: AHRQ, Center for Delivery, Organization, and Markets, Healthcare Cost and Utilization Project, Nationwide Inpatient Sample, 2010

2.6.6 Length of Stay

Studies investigate the relationship between hospital length of stay (LOS) as a determinant of health services utilization and SES, and they mention LOS to be inversely related to SES, while other studies show no effect. However, a study that examined the relationship between race and hospital LOS among the elderly African Americans were found to have a significantly shorter LOS after adjusting for age and health status. A study in Canada found a small inverse association between neighborhood-level income quintiles, an ecological indicator of SES, and LOS for patients admitted to for 14 common illnesses, while a study argued that once patients are admitted in the hospital, there was no relationship between neighborhood-level income quintiles and LOS.

2.6.7 SES and Mortality

Many studies investigated the mortality and the leading cause of health conditions and prevalence of diseases, and supported that mortality differentials by SES may arise because persons of low SES know less about healthy life styles, health problems, and when and where to seek medical care. Studies also mentioned that among the health conditions that have strong associations with SES and have been established as the overall rate of mortality are cardiovascular diseases, infant and maternal mortality, unintended injury, homicide and suicide, and prevalence of diseases include arthritis, heart disease, ulcers, diabetes, hypertension, and chronic bronchitis.[60, 61, 62] Also an individual lacking money and education may be constrained in their ability to purchase health-promoting and life-prolonging goods and services such as medical care. In this regards, this person may not go to the hospital and whose health further deteriorate, or use the emergency department (ED) room, which leads to overcrowding of the ED services.

2.7 The Challenge of Identifying Effective Interventions

Research indicates that SES is one of the strongest predictors of health status. By these many reports SES could influence health disparities directly or indirectly. Given this view, one of the primary goals of Healthy People 2010 is eliminating health disparities among all population subgroups, however, because the factors that influence SES are not all yet understood, it presents challenges in implementing effective interventions. Findings in the literature review noted that the relative failure of individual interventions was interpreted by many as evidence of the importance of environmental factors in health, individuals are products of their environment, the theory went, and thus one cannot change the individual without changing the community in which he or she lives. Therefore, interventions must be implemented in a way that it achieves its purpose. For example, people will decide not to change their risk behaviors, but changing behaviors require major change in thinking and action because health messages are very easy to ignore. In addition, a vast literature has evaluated the impact of these public policies on various risk behaviors. These public policies have been enacted to influence health behaviors, for example, the most important public policy intervention for health is in the area of cigarette smoking. Such policy helps to minimize the impact of tobacco use in the national health. But regardless of the strict policies the driver of this cigarette smoking is large demand responses to price increases.

Although long documented and well-known, income disparities in health status continue in the U.S. Separate provisions will expand healthcare coverage and make it more affordable to more Americans. However, to better understand the social, psychological, and biological processes that link the existence, quantity, structure, or content of social relationships, it is recommended that research is needed to further explore this phenomenon.

CHAPTER III RESEARCH METHODOLOGY

3.1 Introduction

This chapter describes the research design and method used in this study to collect data and analyze the statistical processes. To address the research questions, national inpatient sample hospital discharges data sets were used. The data were explored to examine the association of socioeconomic status (SES) and median household income groups with hospitalization outcomes in the United States from 2008 to 2010. The data showed the pattern in SES, income and hospitalization outcomes, explaining the impact of health in the different income (SES) groups. The chapter also presents general overview of quantitative research, an explanation of why this research method was selected, and the study design approach delineates the diagnoses, procedures, hospital costs, and mortality that affect the different social groups. Further, this chapter explains the research design and approaches; highlight the strategies, and techniques. Finally, this chapter provides descriptions of the data set and covers the data limitations.

3.2 Research Design and Approach

In order to explore the potential answers to the research questions, the research approach was based on quantitative research design, [63] a procedure for collecting and analyzing quantitative data at a given stage of research process, and to better understand a research-defined problem. The reason for using only this approach is that other existing research studies found that this method fits within the study inclusion criteria and the hypothesis or theories which examined the relationship among variables.

Quantitative research method is very helpful in understanding numerical data for developing knowledge, such as cause and effect thinking, reduction to specific variables, and answers questions being examined utilizing various measurements. This approach, as well as the variables and the chosen statistical analysis are the appropriate means to understand and to find answers to the research questions.

In addition, this research provides an opportunity to evaluate the effectiveness of current health program specifically, in the U.S. with the goal of improving and expanding health awareness and policies.

3.3 Data Source

The data used in this study were obtained from the National (Nationwide) Inpatient Sample (NIS),[64] which contain hospital discharge information in the U.S. The NIS is a national database of hospital inpatient stays from community hospitals (all nonfederal short-term general and other specialty hospitals, excluding hospital units of institutions). NIS is part of the Healthcare Cost and Utilization Project (HCUP) family, sponsored by the Agency for Healthcare Research and Quality (AHRQ) through a federal-state partnership. The NIS is known to be the largest all-payer inpatient care database in the U.S. and the data repository stores discharge abstracts from approximately eight million hospitalizations annually from about 1000 hospitals across the nation. According to NIS, the database is designed to represent approximately a 20% stratified sample of all hospitals in the country, and it is publicly available in the U.S since 1988 data year. The database contains over 96 percent of the U.S. population. The NIS data set is available yearly, allowing analysis of trends over time.

The inclusion criteria used for stratified sampling of hospitals into the NIS include location (urban or rural), teaching status, geographical region, patient volume, length of stay, patient charges, and hospital ownership; hospital information, and other related information. Also, the database includes patient demographic characteristics and sociodemographic characteristics. The abstract variables used in this research are admission date, admission elective, admission type, source, gender, race, age, expected source of payment, severity, procedures, diagnosis, comorbidities, and median household income by zip code. The NIS combined files (core, hospital and severity) contain over 150 data elements, however, not all elements were used for this study, the description of the elements are included in the Appendix page. All three files comprise of unique key identifiers (HOSPID and KEY) that were used to perform record linkage and join. NIS data included up to 25 inpatient diagnoses and procedures per hospitalization, and all procedures and diagnoses are coded using the International Classification of Disease, 9th revision, Clinical Modification (ICD-9). The NIS data set did not contain educational and occupational information which is consider to be limitations to the study.

We analyzed hospital discharge data from 2008 to 2010 for the general US population. Due to some state-specific restrictions and confidentiality laws some data elements were not included in the NIS data. These data include specific medical conditions and procedures, HIV/AIDS, behavioral health, and abortion. Another limitation is that the dataset does not have patient level information or patient identified zip code. This means that trends reflecting comparison of living standards between states cannot be determined.

The NIS database house 20% of all inpatient visits from the U.S community hospitals across 45 states in 2010. HCUP partnered with the community organizations within these states that collect and maintain the nationwide inpatient visits data (see Table 3 in Appendix that shows the participating states).

3.4 Study Sample Population

The 3 years dataset consist of approximately (n=24 million) data of the aggregated data. The NIS database provides patients age at admission from 0 to 124 for each hospitalization record for all data years. All admissions with the age range between 18 and 89 were included in our study. In addition to the age limit, the dissertation proposed a subset sampled population of (n=500,000) hospitalization discharge records which were pulled from the 3 years datasets from 2008 to 2010. The project was approved by the Institutional Review Board of Rutgers, The State University of New Jersey. The method that was used to identify the study population was an automated process in which sample selection criteria, a program in SPSS statistical application abstracted the subset sample data from the hospital discharge records.

The study population included 289,950 (58%) White, 60,874 (12.2%) Black, 44,837 (9%) Hispanic, 10,411 (2.1%) Asian or Pacific Islander, 2,916 (.6%) Native American, 13,424 (2.7%) other, and 77,855 (15.5%) missing data. Before the subsample population were taken from the 3 years discharge records from 2008 to 2010. There were 55.9% White, 12% Black, 10.4% Hispanic, 2.3% Asian or Pacific Islander, .6% Native American, 3% other, and 15.7% missing data. So, looking at the population and the subset of the sample population there were not much differences in the number of survey participants. However, the primary study outcome is concentrated on the income groups, which encompasses different race. Section 3.5.2 of this chapter illustrates the socioeconomic income groups. The patients were covered by Medicare, Medicaid, Private including HMO, Self-pay and no charges and others, about .2% were missing data. The overall population when broken down by gender, 59.5% females and 40.5% were males. The NIS data were used to measure numerous variables concerning the care of each subject leading up to their hospitalization.

3.5 Measures

Independent variables included demographic characteristics, hospital characteristics, and socioeconomic status. The dependent or outcome variables that were examined included hospital admission, hospital costs, length of stay, comorbidities, and mortality. The variables were measured as follows.

3.5.1 Demographic Characteristics

The demographic characteristics available for analysis included age, gender, race, and primary source for payment. Age was computed in birth year at admission. Race was coded as white, black, Hispanic, Asian/Pacific Islanders, Native Americans, and other race in the NIS. Race was missing for 15% of discharge abstracts. Gender and primary source for payment or insurance was computed at admission. To evaluate whether there is significant difference in demographic characteristics and hospital discharge rates, we used the defined variables.

3.5.2 Socioeconomic Status

For this study, income was the indicator chosen to represent SES. In our analyses, income is the primary predictor of interest. We focused on this SES predictor to investigate the study hypotheses. Patient's residential zip code was selected as the appropriate locational measure, and NIS classified zip codes into quartiles based on median household income (mean income within a census tract). These income classifications are as follow: quartile 1 = $1 to $38, 999; quartile 2 = $39, 000 to $47, 999; quartile 3 = $48, 000 to $62, 999; quartile 4 = $63, 000 or more. The chosen age population was from 18-89 by race, gender and median household income. The SES quartiles represent 28.2%, 26%, 22.9%, 20.2% respectively and

2.7% were missing data. The residential zip code based on classification of SES is known to reflect aggregate characteristics of its residents and an insight into environmental attributes such as healthcare resources that may have a direct or indirect impact on its residents' health. Unfortunately, besides this aggregated income, several components that may be useful in defining composite measures of SES such as education and occupation data were not available in NIS database. However, variables related to SES included age, race, income, insurance status, and urban-rural designation for the patient's county residence—all of which were included in the analysis.

3.5.3 Hospital Characteristics

Hospital characteristics such as hospital types (teaching, non-teaching), geographical location (rural, urban), bed size (small, medium, large), and region (Northeast, Midwest, South, West) were included. (See Table 2 in Appendix that illustrates the data elements with their description.) However, the data elements included in this study are shown in table 3. These variables were used to compare the hospitalization variance by geographical location. The hospital factors were categorized in respect to teaching or nonteaching status, urban or rural location.

Table 3 Data Variables Used for Analysis

Study Variables	Original Variable Name in the NIS Data Set	Variable Description
Age	AGE	Age in years, Numerical Variable
Admission Type	Elective	Indicates elective admission: (1) elective, (0) non-elective admission, Categorical (binary) variable
Mortality	DIED	Patient did not die during hospitalization (DIED=0); Patient died during hospitalization (DIED=1), Categorical (binary) Variable
GENDER	FEMALE	Gender of patient FEMALE = 1 is Male; FEMALE=0 is female, Categorical (binary) Variable
TOTAL CHARGE	TOTCHG	Total charges , Numerical Variable
RACE	RACE	1 = White, 2 = Black, 3 = Hispanic, 4 = Asian/Pacific, 5 = Native Am., 6 = Other, Categorical Variable
INSURANCE TYPE	PAY1	1=Medicare, 2=Medicaid, 3=Private insurance,4=Self-pay,5=No charge,6=Other, Categorical Variable
NUMBER OF PROCEDURES	NPR	The number of procedures performed while patient was hospitalized, Numerical Variable
SOCIO_ECONOMIC STATUS (SES)	ZIPINC	Median household income for patient's ZIP Code, 1=$1-38,999, 2=$39,000-47,999, 3=$48,000-62,999, 4=63,000 or more, Categorical Variable
COMORBIDITIES	CM_DRUG, CM_ALCOHOL, CM_DRUG ABUSE, CM_DPRESSION, CM_DM, CM_HTN	Comorbidities (drug abuse, alcohol abuse, Depression, diabetes, hypertension), Categorical (binary) Variables
LENGTH OF STAY	LOS	The number of days patient was hospitalized, Numerical Variable
NUMBER OF DIAGNOSES	NDX	The number of diagnoses on the patient record, Numerical Variable
REGION	REGION	Four regions are included Northeast = 1, Midwest =2, South = 3, west =4 , Categorical Variable
HOSPITAL BEDSIZE	HOSP_BEDSIZE	Bed size of hospital (STRATA): (1) small, (2) medium, (3) large
TYPE	TYPE	Teaching status of hospital: (0) non-teaching, (1) teaching Categorical Variable

3.5.4 Study Outcomes Measures

The study outcomes were based on the differences in cost of hospitalization stratified by residential zip code quartile which was used to determine the current trend in time among the different SES groups. The NIS database provides the total charges associated with each hospital stay that was claimed by the hospitals. Further, the total charges of each hospital stay were converted to cost estimates. Use of appropriate analytic methods suitable to the outcome measure and sensitivity analysis to address varying primary study outcomes are needed to draw robust and reliable inferences.

3.5.5 Clinical Factors

The study included the clinical factors or characteristics as one of the outcomes. The NIS sample size enables analyses of existing conditions, and the discharge information includes about 15 inpatient diagnoses and procedures per hospitalization since the incept in 1988 which was increased to 25 diagnoses in 2010. Also, all medical procedures and diagnoses are coded using the International Classification of Disease 9th revision, Clinical Modification (ICD-9-CM). The ICD-codes will be used to compare the trend between patients' conditions and their use of resources by looking at diagnosis-related groups (DRGs), hospital utilization and costs (hospital admissions and total charges). Besides these outcomes, we also analyzed the hospitalization rates, top 10% of medical diagnosis, severities, and procedures, by patient demographics and socioeconomic characteristics. Top 10% analysis were used to compare the prevalence and mortalities among the groups.

3.5.6 Recategorization

Recategorizing some of the elements in the data sets allowed the use of regression analysis in the study. Also allowed the database to run faster. The age element has been grouped into categorical elements to enable adequate performance, so the AGECAT (new element) was coded into 5 categories: ages from 18-28 years= 1, 29-43 years = 2, 44-58 years = 3, 59-73 years= 4, to 74-89 years = 5. Due to the size of the data set it was necessary to recategorize these data elements to enhance optimal processing of the data. This process allowed us to conduct regression analysis. Therefore, simple regression analysis and multiple regression analysis were performed to predict how a specific or multiple data elements affect hospitalization outcomes with 1 for elective admission and 0 for non-elective admission. Variables such as AGE, RACE, GENDER, INSURANCE, and MEDIAN INCOME were used as the independent variables and ELECTIVE as the dependent variable.

3.6 Statistical Modeling Analysis

In this section, statistical modeling analysis of the 2008 to 2010 NIS data sets were performed. In order to make national estimates of the discharge records more accurate, the crude distribution was weighted according to the provided NIS population weights. All subsequent analyses were performed on the weighted population. Descriptive statistics appropriate to the level of measurement were generated for all study variables, comparing the discharge data per payer, region, median income groups, gender, diagnosis, and procedures. This method of analysis revealed general trends and variations in hospitalization among different income groups.

In addition, univariate analysis was computed using x^2 and chi-square tests for categorical and continuous variables, respectively. Also chi-square tests at alpha = 0.05 as the level of significance was computed.

In comparing the means of continuous variables between 3 or more categories, one-way analysis of variance (ANOVA) was computed. ANOVA compares the differences in the means of variable of more than 2 groups. The results show that significance differences were detected and pairwise comparisons were performed utilizing Bonferroni correction for multiple comparisons.

Bivariate and multivariate logistic regression were carried out to examine relationships between SES, comorbidities, and hospital-levels factors. The presence of the outcome variables served as the dependent variables, while patient demographic and hospital characteristics between data year were included in the models. Model predictions come from a single dataset that pooled standardized variables and observations from each NIS dataset. The study also tested for changes between NIS dataset 2008-2010 by calculating age and gender of each outcome in each SES stratum.

All statistical analyses were performed using the statistical software, Statistical Package for the Social Sciences (SPSS, version 22.0, IBM Corp) running on Window. Descriptive analysis was presented using the data element in Table 3 below.

Table 4 Descriptive Analysis per Sub-Categorical data Elements Groups

AGE	PAY 1	HOSP_REGION	FEMALE	ZIPINC_QRTL	RACE
AGE	**PAYER**	**REGION**	**GENDER**	**INCOME**	**RACE**
18-28	Medicare	Northeast	Male	1= $1 to $38, 999; quartile	white
29-43	Private	Midwest	Female	2= $39, 000 to $47, 999; quartile	black
44-58	Medicaid	South		3= $48, 000 to $62, 999; quartile	Hispanic
59-73	Self	West		4 = $63, 000 or more	Asian/Pacific Islanders
74-89					Native Americans, and other race

Multiple linear regression analyses was conducted to test the hypotheses. Multiple regression analysis provides us with an equation explaining the linear relationship between a dependent variable and independent variables. Also, regression analysis provides summary statistics (R^2 and Adjusted R^2) which allow us to evaluate the level of variability with which the regression equation can predict values of the dependent variables. In addition, the following statistical analysis methods were used where necessary for this study:

These parametric methods were used to analyze continuous data that are normally distributed:

- Correlation analysis: Pearson correlation
- Linear regression models

- One-way ANOVA and
- Mean, SD for descriptive analyses

Non-parametric methods were used to analyze variables that all are not normally distributed and are ranks and scores:

- Correlation analysis: Spearman correlation
- Kruskal-Wallis test
- Wilcoxon Rank sum test and Mann-Whitney test and
- Median, interquartile range for descriptive analyses

Categorical variables and binary analysis will be use where appropriate.

- Logistic regression models.
- Chi-square test or Fisher exact test.

3.7 2008-2010 NIS Unit of Analysis

In the 2008 to 2010 NIS datasets, each discharge or admission record is represented by a single record or a unit of analysis. The NIS contains discharge-level records instead of patient-level records, which means that individual patients who are hospitalized multiple times in one year may be present in the NIS dataset multiple times. However, these multiple records have an encounter total charge that shows inpatient admission services were rendered and some instances copay amounts, such as $10 or $20, may be in the total charges (TOTCHG). Generally, total charge does not include professional fees and non-covered charges. For instance, if a patient is admitted 10 times in the hospital, those 10 admissions were counted as 10 different admission records or units of analysis. The NIS 2008-2010 data do not include any patient-level identifier such as medical record or encounter number, therefore, do not allow for analysis at the patient-level. Each inpatient admission is linked to a hospital identifier number that is linked to the AHA survey database.

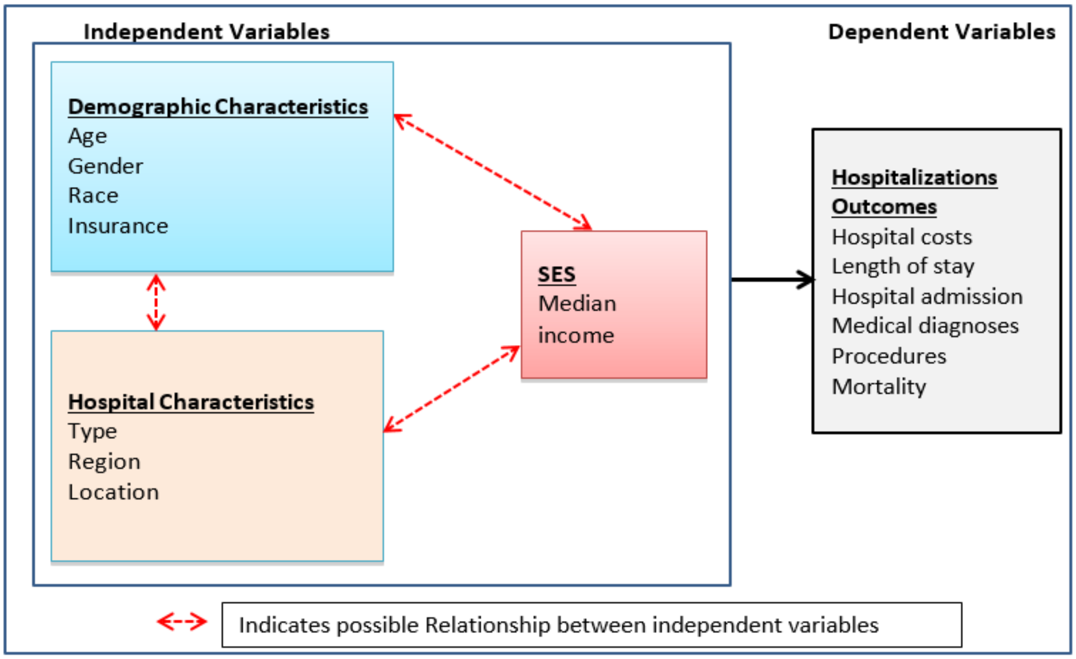


Figure 10 Conceptual Model for this study

The model shown on the **Figure 10** above was used to analyze the relationship in hospitalization and income and charges, using statistical processes such as linear regression with age, gender, median income, hospital characteristics, payer, and race as the independent variables, while the number of hospital admissions or discharges, hospital costs, LOS, and mortality are the dependent variables. A Pearson correlation coefficient was calculated to identify variables that are high correlated ($p<.05$). The broken arrows denote that when two or more independent variables are correlated the variables are considered multicollinear. That is a problem in the application of multiple regression analysis arises when two or more independent variables are high correlated with each other. Multicollinearity often occurs in particular sample when two or more of the explanatory variables are sufficiently highly correlated to make it difficult to separate the effects of one explanatory variable on the dependent variable from the effects of the other explanatory variable. Therefore, it is recommended to check for the strength of the linear relationships among the independent variables to avoid problem.

3.9 Logistic Regression Analysis

Logistic regression describes the relationship between an outcome variable (categorical) dependent variables and a single (or a group of) predictor variable (categorical or quantitative) independent variable. That is, Logistic regression seeks to model the probability of an event occurring depending on the values of the independent variables, which can be categorical or numerical. Also, regression seeks to estimate the probability that an event occurs for a randomly selected observation verses the probability that the event does not occur. It seeks to predict the effect of a series of variables on a binary response variable. Generally, logistic regression is appropriate for describing and testing hypotheses about relationships between a categorical outcome variable and one or more categorical or continuous predictor variables. Frequently, logistic regression is used to refer specifically to the problem in which the dependent variable is a binary, that is, the number of available categories is two. More so, logistic regression calculates the probability of success over the probability of failure. Therefore, using logistic regression analysis, the study test the hypothesis for statistically significant relations between patient's demographic characteristic and the variable ELECTIVE. Logistic regression analysis was useful in predicting the probability that outcome of a dichotomous dependent variable with value of 0 for non-elective admission and value of 1 for elective admission are influenced by a single or multiple independent predictors. Logistic regression was used to compute the estimated association of income groups with changes in the outcome variables. A logistic regression model estimated parameters for the outcome that age was recorded in the model and median income, hospital region, age, comorbidities, gender, and expected payer were used as predictors in the model. The model provided a positive response variable that predicted the odds ratio that a hospital elective admission occurred.

Simple regression was performed examining the probability of hospitalization as a function of one or multiple predictors. Also, simple regression was used to examine relationship between one binomial outcome and one predictor or independent variable. In addition, multiple regressions was computed to investigate the relationship between one binomial outcomes and multiple predictors or independent variables, as well as investigate the relationship between outcome variable and independent variable when adding one or multiple independent variables. By this addition, we assessed how the relationship between the outcome and independent variables is influenced. We investigated interaction terms between SES groups and admission from logistic regression models to determine if changes between the first and last differ by SES group. For instance, age can be added to a simple regression analysis in which median income was initially used as a predictor to predict the likelihood of an elective admission. This determined if age has an effect or association between income and elective admission positively or negatively. The model test the statistical significance differences between admission and median income were determined using t test and $p<0.05$ and three statistical tests, likelihood ratio test, score test, and Wald test. The Wald test confidence limits or Wald confidence interval of 95% are used to test when the population means is equal to zero, null hypothesis. The overall aim of the null hypothesis test is to determine whether the logistic regression model shown in equation 1, 2, or 3 improve over the null model. A null model only outlines the constant or intercept α without the predictor variable, that is, the regression coefficient β is not equal to 0, $Y = \alpha$. Usually, when the p-value is less than 0.0001, it suggests that the null hypothesis is satisfactory and the model evaluation is validated. The receiver operating characteristic (ROC) is valued from 0.5 to 1, while 0.5 shows that the model only predicts the outcome by change and 1 implies that the model perfectly predicts the outcome.

Therefore, using the conceptual model in section 3.7 we show examples of regression analysis models:

Equation 1 is a simple logistic regression model, where dependent variable is Y (represents elective admission) and the predictor is X (median income). We want to predict the probability (p) of occurrence of Y=1 based on some (X) independent variable. The theoretical logistic equation for this prediction is:

Equation 1 Simple Logistic Regression

Simple Logistic Regression: $Y = \alpha + \beta_1 X_1$ (Equation 1)

So, to test the relation between a positive elective admission and median income, we can formulate it as equation 2:

Equation 2 Hospital admission

Elective admission = $\alpha + \beta_1$(Median income) (Equation 2)

To test more than one predictor, $X_1, X_2 \ldots X_n$ will be:

Equation 3 Multiple Logistic Regression

Multiple Logistic Regression: $Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \varepsilon$ (Equation

Testing this model for Elective admission, Age, and Gender will derive:

Equation 4 Hospital admission

Elective admission = $\alpha + \beta_1$(Age) + β_2(Gender) (Equation 4)

Using these equations, Y represents the dichotomous dependent or outcome variable, X represents the predictors or independent variable(s), α presents the Y intercept or constant of the equation, β represents the coefficient of the predictor or independent variable(s), and n represents the number of the last predictor or independent variable(s). Alternatively, the logistic regression equation can be more complex as shown below:

Equation 5 Logit

$$\text{Logit (Y)} = \text{natural log (odds)} = \ln (\pi/1 - \pi) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta n X n \quad \text{(Equation 5)}$$

Equation 5 indicates that, $\boldsymbol{\pi}$ represents the probability of the event Y = 1 (i.e., elective admission for a hospitalization condition), 1 – $\boldsymbol{\pi}$ represents the probability of Y = 0 (i.e., a non-elective admission for a hospitalization condition).

So far, we have demonstrated how the logistic regression model fits in our study to perform data analysis using the 2008-2010 NIS datasets. In conclusion, the research approach and methods presented in this chapter enabled us to investigate, on a quantitative level, the differences in accessing health care system by the various social classes, minority or ethnic groups, and those without insurance coverage in the U.S. The next chapter illustrates the results from the statistical analysis.

CHAPTER IV RESULTS

4.1 Introduction

In this chapter, we present the results of the different statistical analyses that were outlined in the previous chapter. The results of descriptive statistics analysis, analysis of variance, and logistic regression were presented and discussed. We used both descriptive and inferential statistics to answer the research questions and tested the study hypothesis.

The NIS dataset yielded approximately 24 million total hospital admissions across the U.S. for the study period 2008 to 2010. From the total admissions, a subset sample of 500,000 hospital discharge data were randomly selected for the study. A program in SPSS helped to perform the random sample. The subset data extracted from the NIS dataset were used for all further analysis.

4.2 Descriptive Analysis – Hospitalization Rates

This section provides the descriptive statistics that outline the numerical representation of the 2008-2010 NIS data set. Table 1 through Table 20, Figure 11 through Figure 12, and Graphs 1 through Graph 6 provided the baseline sociodemographic characteristics in the study samples. This analysis helped to undercover general trends and variations among different median income groups and inpatient hospitalizations.

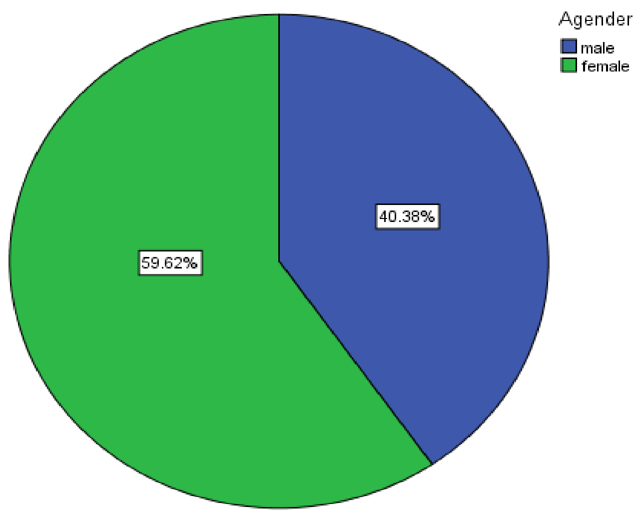
Table 5 shows the characteristics of the study population in the 2008-2010 data sets. The results show patient admission type and patient characteristics, and emergency room was one of the sources of admission that has the highest admission. Admission via the emergency room was 46.1%, elective was 24.8%, urgent was 18.2%, the remainder .4% was trauma center, and missing value was 10.5%. Children or newborn are excluded from the study. It shows the distribution of the study population by race across the 3 years data. It shows that hospitalization rate was highest for ages between 59-73 years old, 24.4%. Adults ages 74 – 89, 24% had a slightly reduce hospitalization rate compared with adults between the ages of 59-73 and 44-58, 21.6%. Others were ages 29-43, 17.1% and 18-28, 12.8%. It displays the gender population, which indicates that female had higher hospital admission. The vast majority population by race was white 58% (n=289,950) which was significantly larger than other races. In contrast black accounted for 12.2% (n=60,874) hospital admissions while Hispanic accounted for 9% (44,837) admissions. This refutes the general knowledge that the older people get the more likely they are to have poor health conditions and be more frequently hospitalized. The table shows that the South region had higher hospital admission compared to others, and the first income levels encountered higher hospitalization rates. Hospital characteristics by owner, bed size and teaching status were also displayed. As show in **Table 6**, the sample mean of length of stay was 4.76 days. The mean of the total hospital costs were $33,877.43.

Table 5 Distribution of Percentages of Admission per Type, Age, Gender, and Race, Region, Income, Owner, Bed size, and Teaching

	Admission	
Admission Type	**Frequency (n)**	**Percent (%)**
emergency	230,310	46.1%
elective	123,830	24.8%
urgent	91,193	18.2%
trauma center	2,072	.4%
other	177	.0
newborn	41	.0
Total	447,623	89.5%
AGECAT		
18 - 28	63,988	12.8
29 - 43	85,509	17.1
44 - 58	108,087	21.6
59 - 73	122,221	24.4
74 - 89	120,195	24.0
Gender		
male	201,657	40.3
female	297,718	59.5
Race		
white	289,950	58.0
black	60,874	12.2
Hispanic	44,837	9.0
Asian or Pacific Islander	10,411	2.1
Native American	2,916	.6
other	13,424	2.7
Region		
Northeast	96861	19.4
Midwest	115077	23.0
South	193036	38.6
West	95026	19.0
Income		
$1 - $38,999	140987	28.2
$39,000 - $47,999	130160	26.0
$48,000 - $62,999	114555	22.9
$63,000 or more	100847	20.2
Owners		
government, nonfederal	69782	14.0
private, non-profit	358755	71.8
private, investor-owned	66353	13.3
Bed size		
small	62668	12.5
medium	117350	23.5
large	314872	63.0
Teaching Status		
non-teaching	265457	53.1
teaching	229433	45.9

Table 6 Descriptive Statistics by Length of Stay and Hospital Charges

	N	Minimum	Maximum	Mean	Std. Deviation
Length of stay	499972	0	358	4.76	6.553
Total charges	492060	100	1,497,869	33,877.43	53,526.539
Valid N (listwise)	492044				



Graph 1 Study Population by Gender

Graph 1 presents the hospital admissions distribution of the study population by gender. The gender breakdown of the study population was females (59.62%) and 40.38% males.

Table 7 Insurance Coverage by Admissions, 2008-2010

Expected Payer	Number of Admissions	Percent
Medicare	212,901	42.6%
Medicaid	77,315	15.5%
Private including HMO	160,197	32.0%
Self-pay	28,055	5.6%
No charge	3069	0.6%
Other	17,385	3.5%
Total	498,922	99.8%
Missing	1078	0.2%
Total	500,000	100.0%

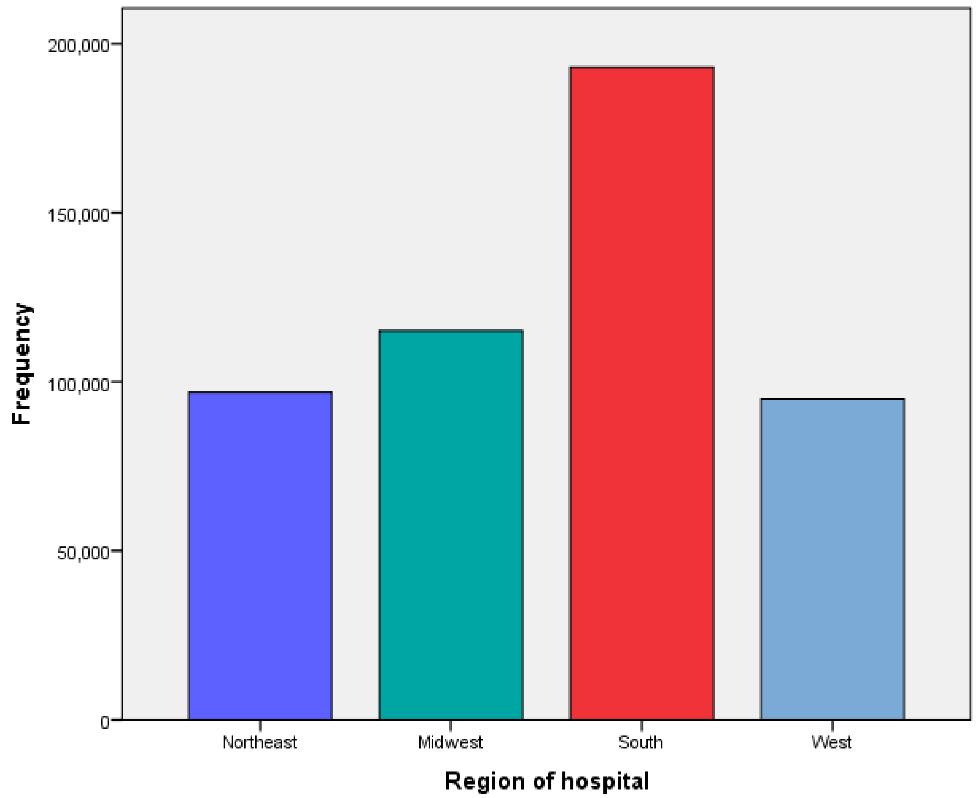
With respect to insurance status, **Table 7** shows the expected payment information, which indicated that Medicare coverage was the most frequently used plan for hospitalized patients 42.6% (n=212,901). Private insurance including HMO was approximately 32.0% (n=160,197), and 5.6% (n=28,055) of the populations paid for their own medical bills, while .6% did not incur any charges. When combined about 10% of the sample population were those without insurance coverage.

Table 8 below presented the distribution of median income in the four categorical income groups. Approximately 28% of the study population fell within the first median income groups $1 - $38,999, 26%, the second category income, $39,000 - $47,999, 22.9%, the third income category, $48,000 - $62,999 and approximately 20% were identified as the fourth income category, $63,000 or more. The percent difference between the 1st and the 4th category was a significant 8%.

Table 8 Admission by Median household income

SES by Median Income	Number of Admissions	Percent
$1 - $38,999	140,987	28.2%
$39,000 - $47,999	130,160	26.0%
$48,000 - $62,999	114,555	22.9%
$63,000 or more	100,847	20.2%
Total	486,549	97.3%
Missing	13,451	2.7%
Total	500,000	100.0%

4.2.1 Socioeconomic Status and Regional Differences in Health

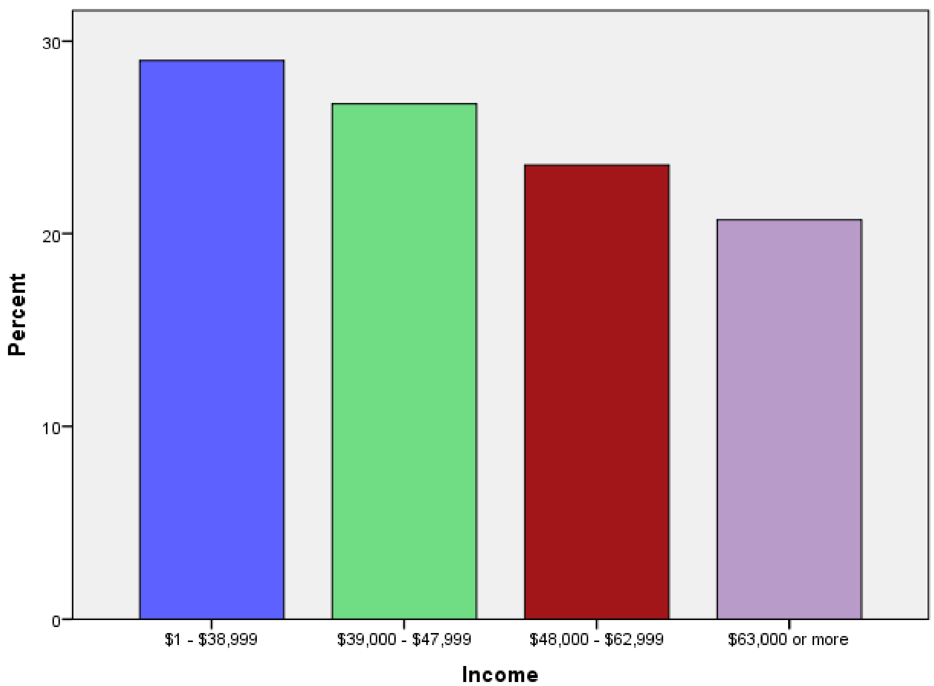


Graph 2 Distribution of Hospitalization by Region of Hospital, 2008 - 2010

Graph 2 indicates the bar chart distribution of hospitalizations across the different regions of the U.S. The figure reveals that the proportion of admission rate in the South region was almost as 2 times of the other regions between 2008 and 2010 as shown in **Table 5**. The assumption could be that health behaviors such as lack of physical activities, lack of accessible transportation, type of diet consumption, and disparity in the income in the household could have led to this increase in hospitalization.

Table 9 Median Household Income by ZIP Code

Median Income	Frequency	Percent
$1 - $38,999	140,987	28.2%
$39,000 - $47,999	130,160	26.0%
$48,000 - $62,999	114,555	22.9%
$63,000 or more	100,847	20.2%
Total	486,549	97.3%
Missing	13,451	2.7%
Total	500,000	100.0%



Graph 3 Median Household Income by Patients' ZIP Code

Neighborhood socioeconomic characteristics at a given level of income, **Table 9** and **Graph 3** show that 28.2% were in the first income level, 26% were in the second income level, 22.9% were in the third income level, while 20.2% were in the fourth income level. Further analysis of the median income groups by region across the 3 years data sets showed that at a given level of income, the first median income groups generally live in more disadvantaged neighborhoods, and the general notion from previous studies was that neighborhood characteristics such as the percentage of residents who are poor have been shown to correlate with people's health. **Table 10** shows the regional differences by the 3 years study data sets. In 2008 there were higher admissions, 171,733 compared to 2009, 164,588 in 2009; and 163,679 in 2010. As might be expected, the South region had the highest frequencies across the 3 years, 193,036 compared to other regions due to other factors such as lack of adequate transportation which may have affected their health. **Table 11** shows the results of median income groups and the regions. From the results, first median income level encountered higher hospitalization when compared the overall hospitalization across the regions. However, income group $1-$38,999 (40%) had higher hospital admission in the South, while income level $63,000 or more (33.3%) had high admission in the Northeast region, income group $48,000-$62,999 (30%) in the West region, and Income group $39,000-$47,999 (33%) had higher admission in Midwest. **Table 12** shows the frequency distribution of the rural and urban location hospital admission and the rural area had 13% and the urban 86% respectively. **Table 13** displays the results of the rural and urban by median income groups. The rural population had 52.5% discharges for the first income level ($1 – 38, 999) compared to the other income groups, while the urban population had 25.5% discharges for both the first and the third income groups compared to the other groups. Also, **Table 14** shows the distribution of the income groups by study years. From the results the lower income group was faced with highest hospitalization rate in 3 consecutive years. One explanation could be that many people lost their jobs and wealth during this time period due to the economy recession, which contributed to poor health, and resulted in high hospitalization rates. Other social factors such as a long history of racial inequality may have contributed to health disparity in the South and Midwest regions.

Table 10 Regional Differences in Discharges by Year, 2008 - 2010

		Calendar year			Total
		2008	2009	2010	
Region of hospital	Northeast	33,101	31,361	32,399	96,861
	Midwest	39,118	38,321	37,638	115,077
	South	67,493	63,389	62,154	193,036
	West	32,021	31,517	31,488	95,026
Total		171,733	164,588	163,679	500,000

Table 11 Region of hospital by Median household income, 2008-2010

Region of Hospital		Median household income national quartile for patient ZIP Code				Total
		$1 - $38,999	$39,000 - $47,999	$48,000 - $62,999	$63,000 or more	
Northeast	Count	20728	19188	21250	30472	91638
	% within Region of hospital (STRATA)	22.6%	20.9%	23.2%	33.3%	100.0%
Midwest	Count	28701	37787	29382	18178	114048
	% within Region of hospital (STRATA)	25.2%	33.1%	25.8%	15.9%	100.0%
South	Count	74862	51606	36694	25385	188547
	% within Region of hospital (STRATA)	39.7%	27.4%	19.5%	13.5%	100.0%
West	Count	16696	21579	27229	26812	92316
	% within Region of hospital (STRATA)	18.1%	23.4%	29.5%	29.0%	100.0%
Total	Count	140987	130160	114555	100847	486549
	% within Region of hospital (STRATA)	29.0%	26.8%	23.5%	20.7%	100.0%

Table 12 Location (urban/rural) of hospital Discharge Records 2008-2010

Location	Frequency	Percent
Rural	62851	13%
Urban	432039	86%
Total	494890	99%
Missing	5110	1%
Total	**500000**	**100%**

Table 13 Location (urban/rural) of hospital Discharges by Median Household Income, 2008-2010

Location (urban/rural) of hospital		Median household income national quartile for patient ZIP Code				Total
		$1 - $38,999	$39,000 - $47,999	$48,000 - $62,999	$63,000 or more	
Rural	Count	32004	21610	6083	1216	60913
	% within Location (urban/rural) of hospital	52.5%	35.5%	10.0%	2.0%	100.0%
Urban	Count	107054	106883	107335	99338	420610
	% within Location (urban/rural) of hospital	25.5%	25.4%	25.5%	23.6%	100.0%
Total	Count	139058	128493	113418	100554	481523
	% within Location (urban/rural) of hospital	28.9%	26.7%	23.6%	20.9%	100.0%

Table 14 Median Household Income based on Discharges year from 2008 to 2010

Calendar year	Median household income national quartile for patient ZIP Code				Total
	$1 - $38,999	$39,000 - $47,999	$48,000 - $62,999	$63,000 or more	
2008	47,774	46,115	38,537	35,502	167,928
2009	45,321	42,927	37,987	33,157	159,392
2010	47,892	41,118	38,031	32,188	159,229
Total	140,987	130,160	114,555	100,847	486,549

4.3 Hospitalization Distribution by Median Household Income

In **Table 15**, we examined the distribution of median income, race and gender and found that hospitalization of white females was predominantly higher across the four categorical median incomes. The overall hospitalization across the income groups was predominantly white compared to other groups. Also, **Table 16** compares insurance coverage across the groups and found that generally a greater percentage of whites had health insurance coverage via Medicare, and this finding was observed across the median income groups. Blacks had a high percentage of Medicaid coverage, a public assistance health program. However, those that had self-pay and no charges were also significant compared to the ones with other types of insurance coverages. Both self-pay and no charge accounted for about 19% together.

Table 15 Frequency Distribution of Race and Median Household Income by Gender

Gender			Median household income national quartile for patient ZIP Code				Total
			$1 - $38,999	$39,000 - $47,999	$48,000 - $62,999	$63,000 or more	
male	Race (uniform)	white	28,277	32,348	29,083	28,837	118,545
		black	11,607	4,837	3,636	2,386	22,466
		Hispanic	5,705	3,401	3,565	2,073	14,744
		Asian or Pacific Islander	483	546	934	1,273	3,236
		Native American	447	331	152	108	1,038
		other	1,310	1,150	1,246	1,021	4,727
	Total		47,829	42,613	38,616	35,698	164,756
female	Race (uniform)	white	38,298	44,626	41,346	40,710	164,980
		black	17,864	7,847	6,143	4,174	36,028
		Hispanic	10,534	7,027	6,600	4,510	28,671
		Asian or Pacific Islander	889	1,148	1,811	2,965	6,813
		Native American	703	561	252	163	1,679
		other	1,953	1,813	2,001	1,680	7,447
	Total		70,241	63,022	58,153	54,202	245,618
Total	Race (uniform)	white	66,575	76,974	70,429	69,547	283,525
		black	29,471	12,684	9,779	6,560	58,494
		Hispanic	16,239	10,428	10,165	6,583	43,415
		Asian or Pacific Islander	1,372	1,694	2,745	4,238	10,049
		Native American	1,150	892	404	271	2,717
		other	3,263	2,963	3,247	2,701	12,174
	Total		118,070	105,635	96,769	89,900	410,374

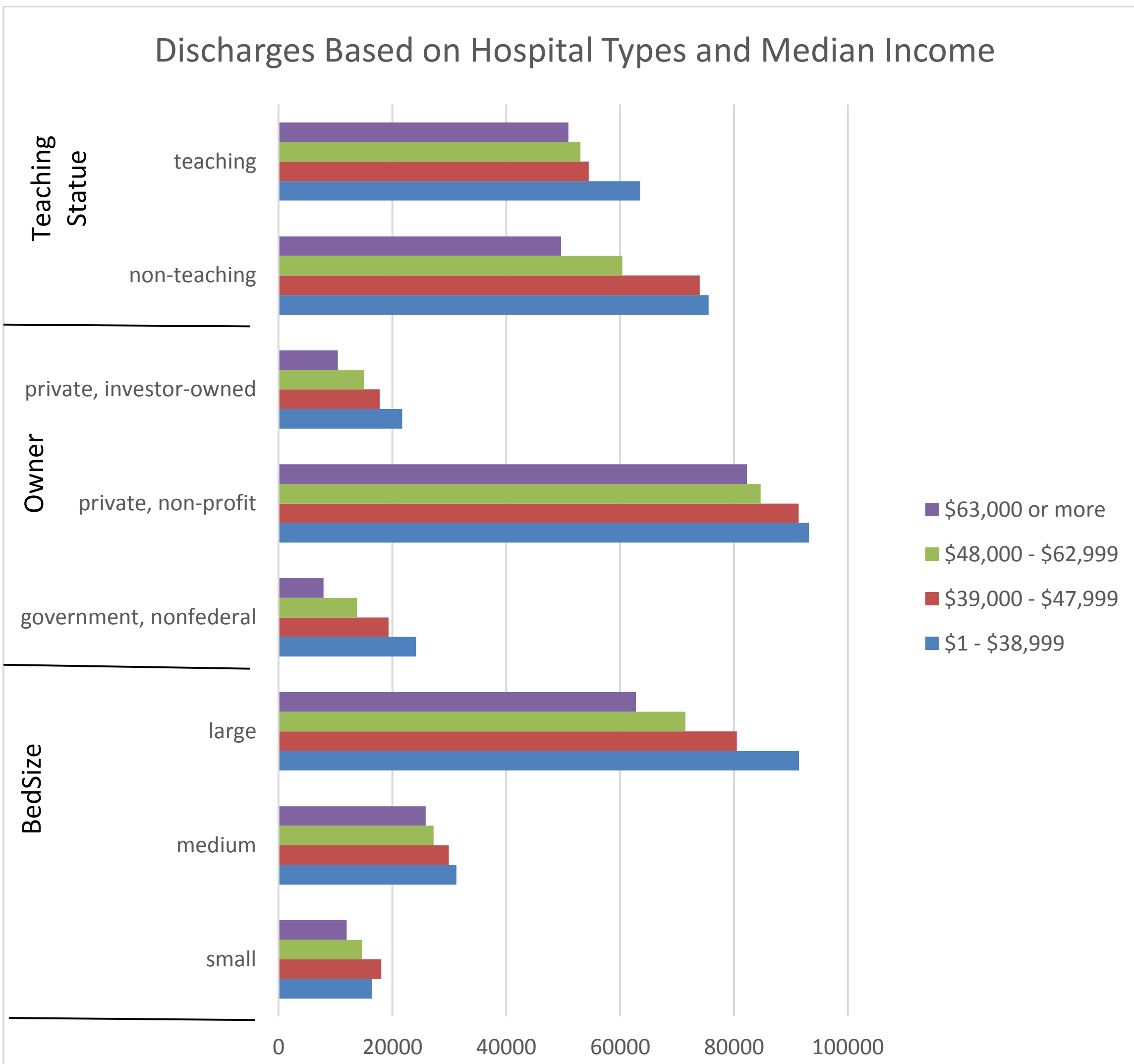
Table 16 Distribution of Insurance Types and Income, stratified by Race

Race			Median household income national quartile for patient ZIP Code				Total
			$1 - $38,999	$39,000 - $47,999	$48,000 - $62,999	$63,000 or more	
white	Primary expected payer (uniform)	Medicare	33956	37869	32636	31134	135595
			12.0%	13.4%	11.5%	11.0%	47.9%
		Medicaid	9678	8537	6238	3504	27957
			3.4%	3.0%	2.2%	1.2%	9.9%
		Private including HMO	16398	23729	25802	30797	96726
			5.8%	8.4%	9.1%	10.9%	34.2%
		Self-pay	3780	3632	2896	1975	12283
			1.3%	1.3%	1.0%	.7%	4.3%
		No charge	377	398	336	151	1262
			.1%	.1%	.1%	.1%	.4%
		Other	2243	2658	2373	1892	9166
			.8%	.9%	.8%	.7%	3.2%
	Total		66432	76823	70281	69453	282989
			23.5%	27.1%	24.8%	24.5%	100.0%
black	Primary expected payer (uniform)	Medicare	11159	4582	3310	2104	21155
			19.1%	7.9%	5.7%	3.6%	36.3%
		Medicaid	8804	3203	2128	1121	15256
			15.1%	5.5%	3.6%	1.9%	26.2%
		Private including HMO	5785	3214	3121	2581	14701
			9.9%	5.5%	5.4%	4.4%	25.2%
		Self-pay	2532	1040	727	427	4726
			4.3%	1.8%	1.2%	.7%	8.1%
		No charge	271	112	104	34	521
			.5%	.2%	.2%	.1%	.9%
		Other	827	499	360	274	1960
			1.4%	.9%	.6%	.5%	3.4%
	Total		29378	12650	9750	6541	58319
			50.4%	21.7%	16.7%	11.2%	100.0%
Hispanic	Primary expected payer (uniform)	Medicare	4500	2440	2651	1588	11179
			10.4%	5.6%	6.1%	3.7%	25.8%
		Medicaid	6032	3744	3168	1654	14598
			13.9%	8.6%	7.3%	3.8%	33.7%
		Private including HMO	3110	2532	2921	2520	11083
			7.2%	5.8%	6.7%	5.8%	25.6%
		Self-pay	1701	1034	800	457	3992
			3.9%	2.4%	1.8%	1.1%	9.2%
		No charge	301	190	169	64	724
			.7%	.4%	.4%	.1%	1.7%
		Other	581	470	443	288	1782
			1.3%	1.1%	1.0%	.7%	4.1%
	Total		16225	10410	10152	6571	43358
			37.4%	24.0%	23.4%	15.2%	100.0%
Asian or Pacific Islander	Primary expected payer (uniform)	Medicare	440	557	860	1237	3094
			4.4%	5.5%	8.6%	12.3%	30.8%
		Medicaid	387	375	499	494	1755
			3.9%	3.7%	5.0%	4.9%	17.5%
			398	584	1121	2260	4363

		Private including HMO	4.0%	5.8%	11.2%	22.5%	43.5%
		Self-pay	98	114	139	144	495
			1.0%	1.1%	1.4%	1.4%	4.9%
		No charge	4	8	11	6	29
			.0%	.1%	.1%	.1%	.3%
		Other	44	54	114	91	303
			.4%	.5%	1.1%	.9%	3.0%
	Total		1371	1692	2744	4232	10039
			13.7%	16.9%	27.3%	42.2%	100.0%
Native American	Primary expected payer (uniform)	Medicare	438	364	129	94	1025
			16.2%	13.4%	4.8%	3.5%	37.8%
		Medicaid	274	150	79	34	537
			10.1%	5.5%	2.9%	1.3%	19.8%
		Private including HMO	259	266	143	123	791
			9.6%	9.8%	5.3%	4.5%	29.2%
		Self-pay	68	40	18	14	140
			2.5%	1.5%	.7%	.5%	5.2%
		No charge	5	2	1	0	8
			.2%	.1%	.0%	0.0%	.3%
		Other	106	66	33	6	211
			3.9%	2.4%	1.2%	.2%	7.8%
	Total		1150	888	403	271	2712
			42.4%	32.7%	14.9%	10.0%	100.0%
other	Primary expected payer (uniform)	Medicare	1106	979	1071	728	3884
			9.1%	8.1%	8.8%	6.0%	32.0%
		Medicaid	821	684	600	389	2494
			6.8%	5.6%	4.9%	3.2%	20.5%
		Private including HMO	800	864	1193	1255	4112
			6.6%	7.1%	9.8%	10.3%	33.8%
		Self-pay	390	274	204	201	1069
			3.2%	2.3%	1.7%	1.7%	8.8%
		No charge	22	22	53	15	112
			.2%	.2%	.4%	.1%	.9%
		Other	115	136	122	109	482
			.9%	1.1%	1.0%	.9%	4.0%
	Total		3254	2959	3243	2697	12153
			26.8%	24.3%	26.7%	22.2%	100.0%
Total	Primary expected payer (uniform)	Medicare	51599	46791	40657	36885	175932
			12.6%	11.4%	9.9%	9.0%	43.0%
		Medicaid	25996	16693	12712	7196	62597
			6.3%	4.1%	3.1%	1.8%	15.3%
		Private including HMO	26750	31189	34301	39536	131776
			6.5%	7.6%	8.4%	9.7%	32.2%
		Self-pay	8569	6134	4784	3218	22705
			2.1%	1.5%	1.2%	.8%	5.5%
		No charge	980	732	674	270	2656
			.2%	.2%	.2%	.1%	.6%
		Other	3916	3883	3445	2660	13904
			1.0%	.9%	.8%	.6%	3.4%
	Total		117810	105422	96573	89765	409570
			28.8%	25.7%	23.6%	21.9%	100.0%

Subsequently, patients' insurance, median income, race were accessed and **Table 16** described the insurance distribution across the income levels. The result revealed that Medicare was mostly subscribed insurance among the median income groups compared to other types of insurance coverages. But, Hispanic relied more on Medicaid than other kinds of coverages. This could be due to the immigration issues and lack of employment.

4.3.1 Differences in Hospitalization Characteristics and Socioeconomic Status



Graph 4 Discharges Based on Hospital Types and Median Income, 2008 - 2010

Graph 4 above illustrates the number of discharges across the various hospital characteristics such as the teaching status, type, and bed size categorized by the median income levels. From the results, we can see that the non-teaching hospital, large size hospital, and private, non-profit organizations usually have high volume admissions compared to the other hospital types. Interestingly, those within these median income levels, \$1 – 38,999 and \$39,000 – 47, 999 had higher hospitalization rates than other median income groups. When we compared hospital characteristics with the region, we found that the south region had higher hospitalization rates, as shown in **Table 17** shows that Bed size had 190,827 admissions, owner, private, non-profit across the regions were 358,755, while non-teaching status were 265,457 discharges across the various regions. The south region tends to face health challenges, this could probably be due to other factors, which is beyond the scope of this study. Even though, this was not part of the study scope, but our assumption was that the south region are faced with health care access and transportation challenges and also probably due to their lifestyles.

Table 17 Discharges based on Hospital Characteristics in the 2008 -2010

Bed size by Region of hospital

		Region of hospital				Total
		Northeast	Midwest	South	West	
Bed size	small	13643	18019	21003	10003	62668
	medium	26245	22913	45814	22378	117350
	large	56973	71326	124010	62563	314872
Total		96861	112258	190827	94944	494890

Owner by Region of hospital

		Region of hospital				Total
		Northeast	Midwest	South	West	
Owner	government, nonfederal	7426	8263	36054	18039	69782
	private, non-profit	87135	97439	112937	61244	358755
	private, investor-owned	2300	6556	41836	15661	66353
Total		96861	112258	190827	94944	494890

Teaching status by Region of hospital

		Region of hospital				Total
		Northeast	Midwest	South	West	
Teaching status	non-teaching	35567	54933	113072	61885	265457
	teaching	61294	57325	77755	33059	229433
Total		96861	112258	190827	94944	494890

4.4 Diagnoses and Procedures that Affect Socioeconomic Status

Some comorbidities appeared more frequently across the median income groups. **Table 18** below presents the clinical conditions, comorbidities that impact the median income groups. Among the 4 groups, the lower SES groups were more likely than other groups from the middle to the highest SES to experience increased hospitalizations. Particularly, the lower income group had a high number of hospitalizations. Hypertension was more significant among the various income groups and Diabetes and Depression were similar in number across the groups. However, the first category of income group encountered higher prevalence than the other groups. The results suggests that adequate health programs are necessitated out the hospitals in order to combat these health issues.

Table 18 Comorbidities by Median household income national quartile for patient ZIP Code

		Median household income national quartile for patient ZIP Code				Total
		$1 - $38,999	$39,000 - $47,999	$48,000 - $62,999	$63,000 or more	
Comorbidities	Alcohol	2302	1954	1581	1237	7074
	Depression	4150	4667	4110	3705	16632
	Diabetes	7591	6703	5374	4051	23719
	Drug Abuse	4478	3156	2474	1701	11809
	Hypertension	66422	58969	50913	43607	219911
Total		84943	75449	64452	54301	279145

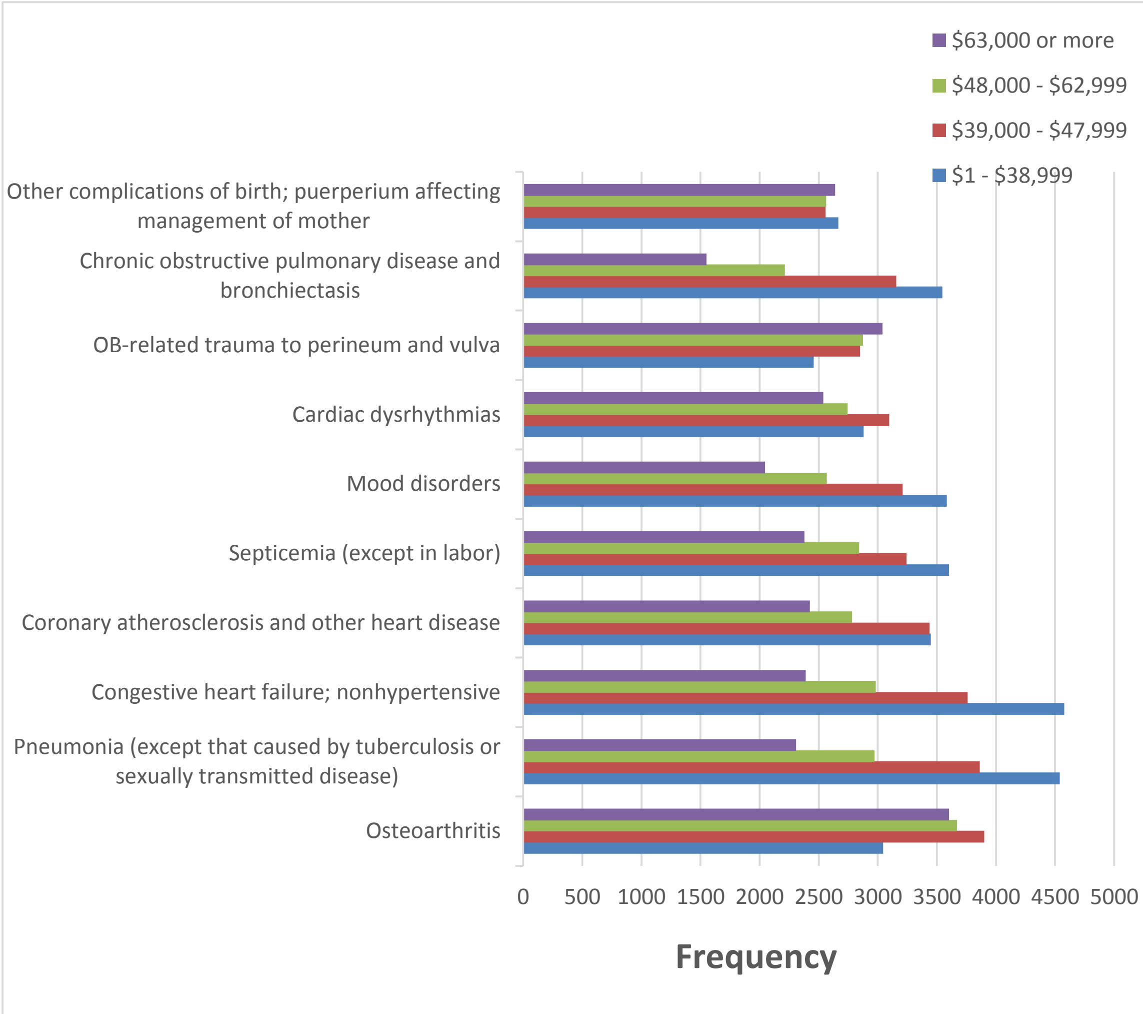


Figure 11 Top 10 Primary Diagnosis by Median Household Income, 2008-2010

Figure 11 above displays the top 10 primary diagnosis across the median income groups in the 2008-2010 data sets. The results indicated that the number of disease prevalence that were diagnosed on patients were significantly higher in the first income group compared to others. The analysis revealed that patients were diagnosed with higher rate of congestive heart failure; nonhypertensive more than other diseases that affected the first income group. Pneumonia was predominately high in the first income group as well.

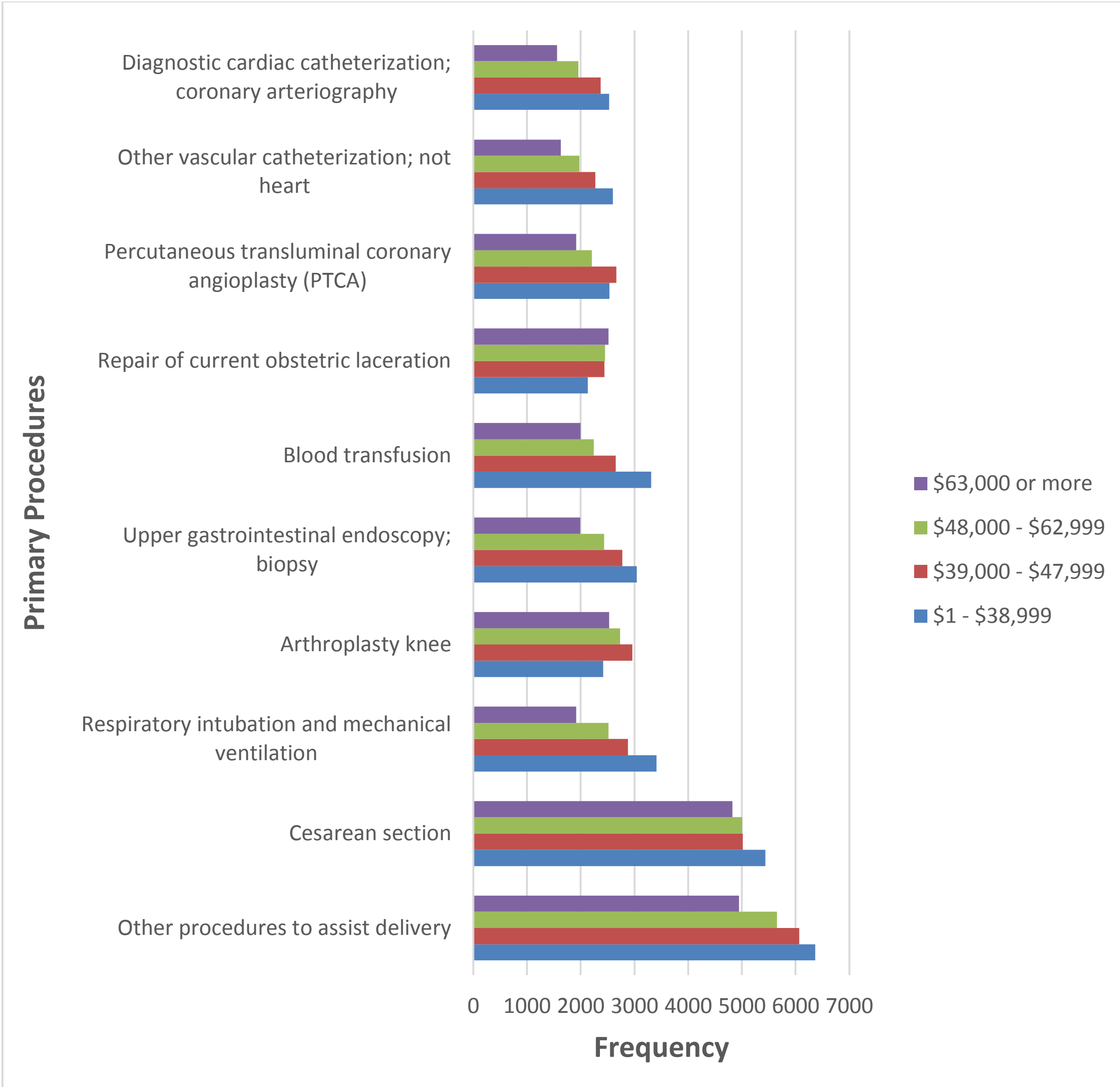


Figure 12 Top 10 Primary Procedures by Median Household Income 2008-2010

The results on **Figure 12** indicate that the most frequent procedures across the median income groups were other procedures to assist delivery and cesarean section and was higher than other procedures. These were more evident in the first income group compared to the other groups.

4.5 Socioeconomic Status and Mortality

Table 19 Mortality during Hospitalization by Median household income ZIP Code

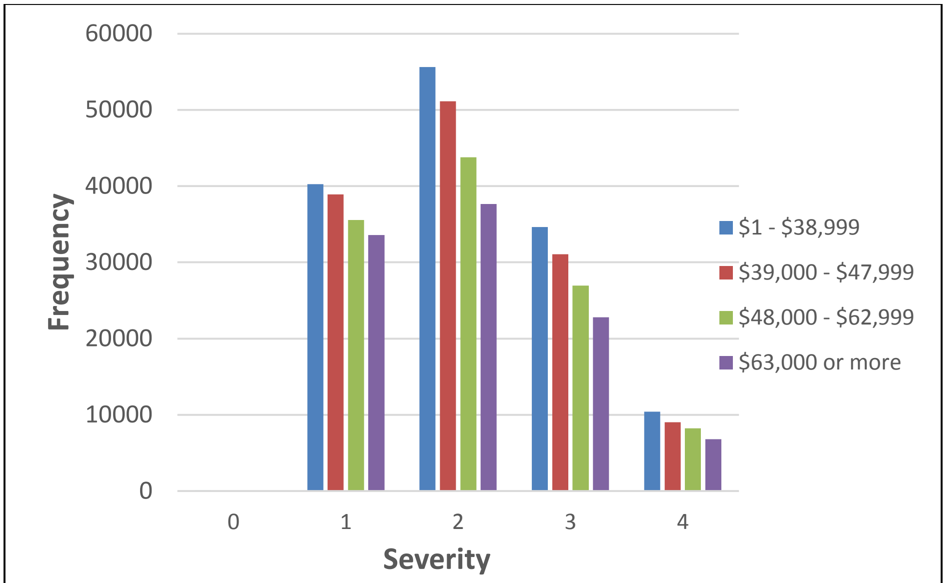
Calendar year	Median household income national quartile for patient ZIP Code				Total
	$1 - $38,999	$39,000 - $47,999	$48,000 - $62,999	$63,000 or more	
2008	1084	992	813	775	3664
2009	983	857	786	653	3279
2010	990	827	764	584	3165
Total	3057	2676	2363	2012	10108

Dramatic differences were observed in mortality among the median income groups in the U.S. during the 3 years study periods. The largest and most consistent mortality were generally observed for the first median income level (**Table 19 and Graph 5**) in 2008. This may have occurred due to the regression or the economic meltdown during this period. Although, the other median income groups also have significantly higher mortality rate in this time period compared to other years. When compared the mortality rate with age, we observed that the age groups 74 – 89 have higher mortality occurrences (n=5048) than other age groups (**Table 20**). **Graph 6** shows the severity of illness across the income levels. The graph indicated that the first income level, $1 – 38,999 encountered significant higher severity across the different severity levels compared to the other income groups. Severity level 2 was more evident across the groups.

Graph 5 Mortality during hospitalization by Median Income

Table 20 Mortality by Median household income for patient ZIP Code and AGE

AGE Group	Median household income national quartile for patient ZIP Code				Total
	$1 - $38,999	$39,000 - $47,999	$48,000 - $62,999	$63,000 or more	
18 - 28	61	38	34	20	153
29 - 43	138	117	88	62	405
44 - 58	533	369	360	252	1514
59 - 73	925	811	716	536	2988
74 - 89	1400	1341	1165	1142	5048
Total	3057	2676	2363	2012	10108



Severity	Median household income national quartile for patient ZIP Code				Total
	$1 - $38,999	$39,000 - $47,999	$48,000 - $62,999	$63,000 or more	
0	85	62	55	20	222
1	40,241	38,895	35,550	33,594	148,280
2	55,621	51,121	43,800	37,665	188,207
3	34,647	31,073	26,940	22,777	115,437
4	10,393	9009	8210	6791	34,403
Total	140,987	130,160	114,555	100,847	486,549

Graph 6 Severity of Illness by Median Household Income

4.6 Results of Research Questions

The previous section in this chapter described the patient, clinical diagnosis, procedures, and hospital characteristics in relation to socioeconomic status. In this section, we provide the results of the goals and objectives as they pertain to the questions and the defined hypotheses of this dissertation using the most appropriate statistical procedures and methods.

4.6.1 How does socioeconomic status, median household income vary by patient demographic characteristics, gender and age?

Using Pearson Chi-square test, we computed SES by income and gender. Gender is one of the elements that described the patient demographic characteristics. The test checks if the distribution of patient characteristics gender is due to chance or if there is a difference between the patient gender and SES by income. The hypotheses stated that:

- $H_0$1: There is no statistically significant difference between the four categories of the median income groups and hospitalized patients' gender and age in the 2008-2010 NIS data set.
- H_a1: There is statistically significant difference between the four categories of the median income groups and hospitalized patients' gender and age in the 2008-2010 NIS data set.

Notice that the output in the **Table 21** compares the differences between SES by Median income groups and gender, using the 4 x 2 chi-square, we found that P-value = 0.001 which is less than P-value = 0.05, degree of freedom = 3, and symmetric measure checked for violation and found no violation. In the 2008-2010 admission records, 29.3% of men were in the first income group, \$1 - \$38,999 compared with 28.8% of women as well as the second level, \$39,000 – 47,999 26.9% of men compared with 26.7% of women. In the other median household income categories, there more women than men. However, comparing the median household income by gender, the results suggest that there are differences in the median household income categories. But the differences was significant x^2=26.54, DF=3, p<0.001. Therefore we rejected the $H_0$1 (null hypotheses) and accept the H_a1 (alternative hypotheses) and concluded that there was statistically significant difference between the median income groups and hospitalized patient characteristic, gender.

Table 21 Variance in Median household income Group by Sex

			Median household income national quartile for patient ZIP Code				
			$1 - $38,999	$39,000 - $47,999	$48,000 - $62,999	$63,000 or more	Total
sex	male	Count	57124	52474	45724	39820	195142
		% within sex	29.3%	26.9%	23.4%	20.4%	100.0%
	female	Count	83769	77579	68657	60800	290805
		% within sex	28.8%	26.7%	23.6%	20.9%	100.0%
Total		Count	140893	130053	114381	100620	485947
		% within sex	29.0%	26.8%	23.5%	20.7%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	26.544[a]	3	.000
Likelihood Ratio	26.557	3	.000
Linear-by-Linear Association	25.936	1	.000
N of Valid Cases	485947		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 40406.03.

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	.007	.000
	Cramer's V	.007	.000
N of Valid Cases		485947	

Table 22 Age of Patients by Median Household Income

Median household income national quartile for patient ZIP Code	Mean	Median	N
$1 - $38,999	55.01	57.00	140,987
$39,000 - $47,999	56.09	58.00	130,160
$48,000 - $62,999	56.01	58.00	114,555
$63,000 or more	56.77	59.00	100,847
Total	55.90	58.00	486,549

ANOVA Table

			Sum of Squares	df	Mean Square	F	Sig.
Age Median household income national quartile for patient ZIP Code	Between Groups	(Combined)	193553.348	3	64517.783	159.470	.000
		Linearity	165094.800	1	165094.800	408.069	.000
		Deviation from Linearity	28458.548	2	14229.274	35.171	.000
	Within Groups		196844460.138	486545	404.576		
	Total		197038013.486	486548			

Measures of Association

	R	R Squared	Eta	Eta Squared
Age in years at admission * Median household income national quartile for patient ZIP Code	.029	.001	.031	.001

Table 23 Variance in Age of Patients by Median Household Income, Crosstab

			Median household income national quartile for patient ZIP Code				
			$1 - $38,999	$39,000 - $47,999	$48,000 - $62,999	$63,000 or more	Total
AGECAT	18 - 28	Count	19921	17417	14565	10466	62369
		% within AGECAT	31.9%	27.9%	23.4%	16.8%	100.0%
	29 - 43	Count	22801	20783	19880	19554	83018
		% within AGECAT	27.5%	25.0%	23.9%	23.6%	100.0%
	44 - 58	Count	32522	27569	23916	20324	104331
		% within AGECAT	31.2%	26.4%	22.9%	19.5%	100.0%
	59 - 73	Count	34475	32425	27982	24009	118891
		% within AGECAT	29.0%	27.3%	23.5%	20.2%	100.0%
	74 - 89	Count	31268	31966	28212	26494	117940
		% within AGECAT	26.5%	27.1%	23.9%	22.5%	100.0%
Total		Count	140987	130160	114555	100847	486549
		% within AGECAT	29.0%	26.8%	23.5%	20.7%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	1906.466[a]	12	.000
Likelihood Ratio	1922.635	12	.000
Linear-by-Linear Association	413.479	1	.000
N of Valid Cases	486549		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 12927.22.

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	.063	.000
	Cramer's V	.036	.000
	Contingency Coefficient	.062	.000
N of Valid Cases		486549	

In the **Table 22**, the average age of all median income groups was 55.90 and a total number of 486,549 hospital admission records were included. These 486,549 records were assigned to one of the four groups, based on their median household income. The average age does differ slightly across the groups. Patients that are in the first income category were younger than the patients in the other median income categories. Comparing the differences between the means, R^2 was about 29%, P-value = 0.001, which was less than P-value = 0.05, degree of freedom = 3 and the differences were significant and **Table 23** also indicated that there were significance differences across median income and age groups; therefore, the H_0 was rejected and the H_a1 was accepted.

We concluded that the outputs in **Table 22** and **Table 23** demonstrate that there were statistically significant differences between the means of median household income groups, by patient's gender and age in the 2008-2010 data set.

4.6.2 What are the groups of patients in the four categorical median household income and cost? Is there statistically significant difference between the four categorical median household income and hospital charge in the 2008-2010 discharge records?

- $H_0 2$: There is no statistically significant difference between the categorical median income groups and hospital charge in the 2008–2010 discharge records.
- $H_a 2$: There is statistically significant difference between the categorical median income groups and hospital charge in the 2008–2010 discharge records.

Table 24 Descriptive statistics for Total Charges

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
$1 - $38,999	139961	31379.73	51235.144	136.951	31111.31	31648.15	100	1434234
$39,000 - $47,999	128909	32224.21	50182.520	139.769	31950.27	32498.16	110	1414355
$48,000 - $62,999	112313	35558.93	55399.300	165.306	35234.93	35882.93	115	1497869
$63,000 or more	97545	38023.30	58396.870	186.977	37656.82	38389.77	133	1484145
Total	478728	33941.28	53545.816	77.389	33789.60	34092.96	100	1497869

Test of Homogeneity of Variances

Total Charges

Levene Statistic	df1	df2	Sig.
250.035	3	478724	.000

ANOVA

Total Charges

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3217704753413.617	3	1072568251137.872	374.965	.000
Within Groups	1369366515844917.000	478724	2860450940.093		
Total	1372584220598331.000	478727			

Post Hoc tests (Multiple Comparisons)

Total charge

Bonferroni

(I) Median household income national quartile for patient ZIP Code	(J) Median household income national quartile for patient ZIP Code	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
$1 - $38,999	$39,000 - $47,999	-844.484*	206.464	.000	-1389.19	-299.78
	$48,000 - $62,999	-4179.204*	214.257	.000	-4744.47	-3613.94
	$63,000 or more	-6643.568*	223.074	.000	-7232.10	-6055.04
$39,000 - $47,999	$1 - $38,999	844.484*	206.464	.000	299.78	1389.19
	$48,000 - $62,999	-3334.720*	218.308	.000	-3910.67	-2758.77
	$63,000 or more	-5799.084*	226.967	.000	-6397.88	-5200.28
$48,000 - $62,999	$1 - $38,999	4179.204*	214.257	.000	3613.94	4744.47
	$39,000 - $47,999	3334.720*	218.308	.000	2758.77	3910.67
	$63,000 or more	-2464.364*	234.079	.000	-3081.93	-1846.80
$63,000 or more	$1 - $38,999	6643.568*	223.074	.000	6055.04	7232.10
	$39,000 - $47,999	5799.084*	226.967	.000	5200.28	6397.88
	$48,000 - $62,999	2464.364*	234.079	.000	1846.80	3081.93

*. The mean difference is significant at the 0.05 level.

Total hospital charge

	Median household income national quartile for patient ZIP Code	N	Subset for alpha = 0.05			
			1	2	3	4
Tukey HSD[a,b]	$1 - $38,999	139961	31379.73			
	$39,000 - $47,999	128909		32224.21		
	$48,000 - $62,999	112313			35558.93	
	$63,000 or more	97545				38023.30
	Sig.		1.000	1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 117448.001.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

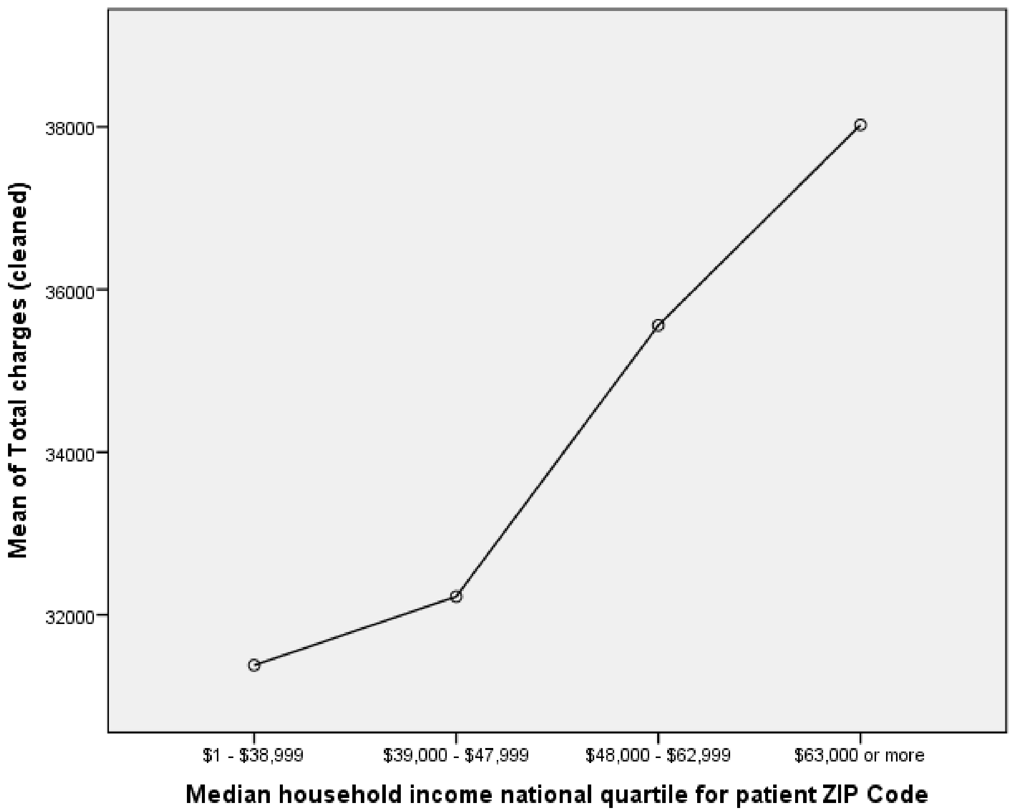


Table 24 above gave an overall descriptions of total hospital charges for each median income groups. The average (column mean) hospital cost for all income levels is $33,941.28; however, the average hospital cost ranges from a low of $31,379.73 for the patients in the lowest income group to a high of $38,023.30 for the patients in the highest income group. Also, the lowest variability (column Std. Deviation, $50,182.52) in hospital cost was for the median income group ($39,000 - $47,999), while the highest $58,396.87 was for the patients in the median income group ($63,000 or more), so also the lowest standard error was found in the median income group ($1 - $38,999), since they were the largest group. For each group there was 95% confidence interval for the population value of the average hospital cost, which indicated a 95% confidence that the true average hospital cost for those within lowest median income levels ($1 - $38,999) was between $31,111.31 and

$31,648.15. For those within the highest median income levels ($63,000 or more), there was a 95% confidence that the true average hospital cost was between $37,656.82 and $38,389.77.

Subsequently, the Analysis of Variance (ANOVA) test suggested that there were statistically significant differences between the average hospital costs among the 4 categories of median household income groups. These differences in means were significant (P-value = .001, DF = 374.97). The null hypothesis stated that the population means for all 4 categories of median income groups were the same. That is, there is no difference in the average hospital cost among the median income groups. The alternative hypothesis is that there is difference. From the ANOVA table above, we rejected the null hypothesis because it was unlikely that the average hospital cost for all median income groups were the same in the population. P-value = 0.001, revealed that there was a very strong evidence to reject the $H_0 2$ and accept the $H_a 2$. In other to determine the group means that were different from each other, we used the post hoc test-Bonferroni, multiple comparison procedures. The test indicated that the difference in cost between the groups were significantly different from one another. We saw that patients in the median income level, $63,000 or more incurred higher cost than each of the other median income groups. Patients in the median income group, $48,000 – $62,999 also incurred higher cost than the $1 - $38,999 and $39,000 - $47,999 median income groups.

4.6.3 Which clinical factors, such as the types of comorbidities and procedures that affect median income groups most prevalently? Is there significant difference among the median income groups?

$H_0 3$: There is no statistically significant difference in the prevalence of clinical factors, such as the types of comorbidities and procedure, among the median income groups.

$H_a 3$: There is statistically significant difference in the prevalence of clinical factors, such as the types of comorbidities and procedure, among the median income groups.

A nonparametric test, chi-square, was used to determine if there were statistical significant differences in the prevalence of clinical factors. We first computed the distributions across the combination of the comorbidities and the median household income variables, which are indicated in the appendix **Table 4**. From the results, we observed that there were differences in the variation of prevalence of comorbidities among the median income groups. We further explored the statistically significant differences by analyzing these comorbidities: Alcohol, Depression, Diabetes, Drug Abuse, and Hypertension. **Table 25** below indicated that there was strong evidence to reject the null hypothesis and accept the alternative hypothesis given that value = 808.02, degrees of freedom = 12, and the P-value = 0.000. From **Graph 7**, we concluded that hypertension affects the median income groups and not only did those at the bottom of the SES, median income level experience higher hypertension rates, but those above them also were straitened by hypertension. **Table 26 and Table 27** show the number of diagnoses and procedures that were performed across the median income groups.

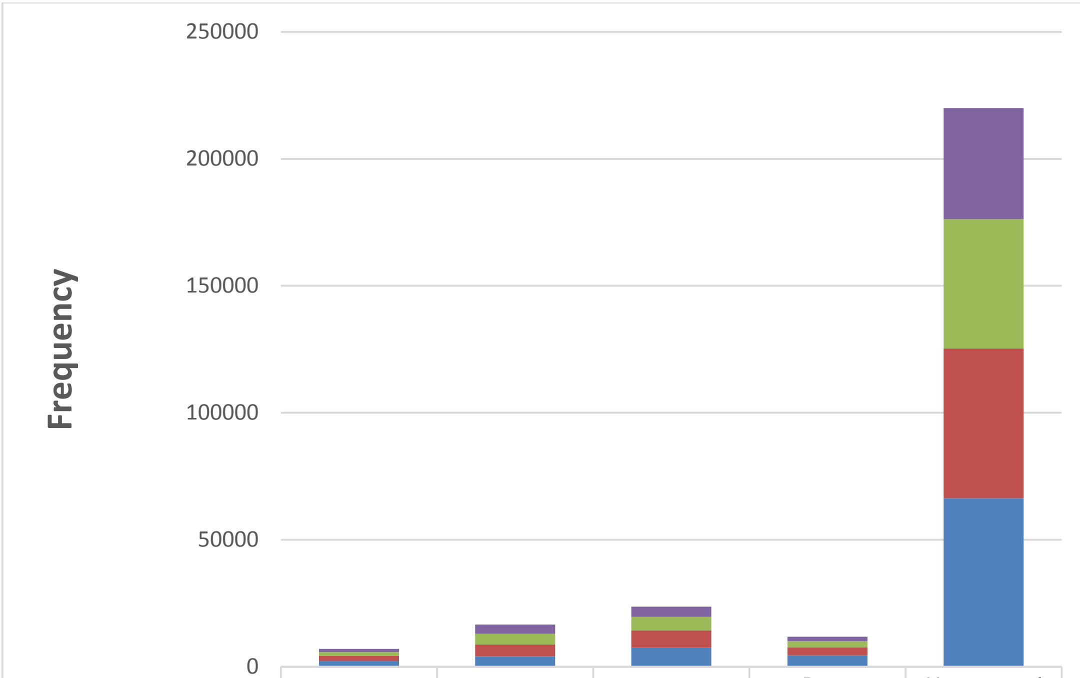
Table 25 Comorbidities by Median Income Groups for patient ZIP Code

Comorbidities			Median household income national quartile for patient ZIP Code: $1 - $38,999	$39,000 - $47,999	$48,000 - $62,999	$63,000 or more	Total
Comorbidities	Alcohol	Count	2302	1954	1581	1237	7074
		% within Comorbidities	32.5%	27.6%	22.3%	17.5%	100.0%
	Depression	Count	4150	4667	4110	3705	16632
		% within Comorbidities	25.0%	28.1%	24.7%	22.3%	100.0%
	Diabetes	Count	7591	6703	5374	4051	23719
		% within Comorbidities	32.0%	28.3%	22.7%	17.1%	100.0%
	Drug Abuse	Count	4478	3156	2474	1701	11809
		% within Comorbidities	37.9%	26.7%	21.0%	14.4%	100.0%
	Hypertension	Count	66422	58969	50913	43607	219911
		% within Comorbidities	30.2%	26.8%	23.2%	19.8%	100.0%
Total		Count	84943	75449	64452	54301	279145
		% within Comorbidities	30.4%	27.0%	23.1%	19.5%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	808.016[a]	12	.000
Likelihood Ratio	818.187	12	.000
Linear-by-Linear Association	3.255	1	.071
N of Valid Cases	279145		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 1376.08.



	Alcohol	Depression	Diabetes	Drug Abuse	Hypertension
$63,000 or more	1237	3705	4051	1701	43607
$48,000 - $62,999	1581	4110	5374	2474	50913
$39,000 - $47,999	1954	4667	6703	3156	58969
$1 - $38,999	2302	4150	7591	4478	66422

Graph 7 Frequency of Comorbidities by Median Household Income

Table 26 Diagnoses by Median Income Groups for patient ZIP Code in the 2008 - 2010 discharge records

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
$1 - $38,999	140987	8.53	4.913	.013	8.50	8.55	0	48
$39,000 - $47,999	130160	8.49	4.925	.014	8.46	8.51	0	33
$48,000 - $62,999	114555	8.59	5.147	.015	8.56	8.62	0	34
$63,000 or more	100847	8.40	5.184	.016	8.37	8.43	0	31
Total	486549	8.50	5.030	.007	8.49	8.52	0	48

Test of Homogeneity of Variances

Number of diagnoses

Levene Statistic	df1	df2	Sig.
217.183	3	486545	.000

ANOVA

Number of diagnoses

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2088.791	3	696.264	27.529	.000
Within Groups	12305603.044	486545	25.292		
Total	12307691.834	486548			

Post Hoc Tests

Multiple Comparisons

Dependent Variable: Number of diagnoses

Tukey HSD

(I) Median household income national quartile for patient ZIP Code	(J) Median household income national quartile for patient ZIP Code	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
$1 - $38,999	$39,000 - $47,999	.038	.019	.194	-.01	.09
	$48,000 - $62,999	-.065*	.020	.007	-.12	-.01
	$63,000 or more	.128*	.021	.000	.07	.18
$39,000 - $47,999	$1 - $38,999	-.038	.019	.194	-.09	.01
	$48,000 - $62,999	-.103*	.020	.000	-.16	-.05
	$63,000 or more	.089*	.021	.000	.04	.14
$48,000 - $62,999	$1 - $38,999	.065*	.020	.007	.01	.12
	$39,000 - $47,999	.103*	.020	.000	.05	.16
	$63,000 or more	.192*	.022	.000	.14	.25
$63,000 or more	$1 - $38,999	-.128*	.021	.000	-.18	-.07
	$39,000 - $47,999	-.089*	.021	.000	-.14	-.04
	$48,000 - $62,999	-.192*	.022	.000	-.25	-.14

*. The mean difference is significant at the 0.05 level.

Homogeneous Subsets

Number of diagnoses

Tukey HSD[a,b]

Median household income national quartile for patient ZIP Code	N	Subset for alpha = 0.05		
		1	2	3
$63,000 or more	100847	8.40		
$39,000 - $47,999	130160		8.49	
$1 - $38,999	140987		8.53	
$48,000 - $62,999	114555			8.59
Sig.		1.000	.243	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 119684.701.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

Table 27 Procedures by Median Income Groups for patient ZIP Code

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
$1 - $38,999	140987	1.63	2.128	.006	1.62	1.64	0	31
$39,000 - $47,999	130160	1.72	2.141	.006	1.71	1.73	0	31
$48,000 - $62,999	114555	1.79	2.197	.006	1.77	1.80	0	31
$63,000 or more	100847	1.91	2.220	.007	1.90	1.92	0	29
Total	486549	1.75	2.169	.003	1.74	1.76	0	31

Test of Homogeneity of Variances

Number of procedures

Levene Statistic	df1	df2	Sig.
21.387	3	486545	.000

ANOVA

Number of procedures

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4843.134	3	1614.378	343.760	.000
Within Groups	2284932.669	486545	4.696		
Total	2289775.803	486548			

Post Hoc Tests

Multiple Comparisons

Dependent Variable: Number of procedures

Tukey HSD

(I) Median household income national quartile for patient ZIP Code	(J) Median household income national quartile for patient ZIP Code	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
$1 - $38,999	$39,000 - $47,999	-.089*	.008	.000	-.11	-.07
	$48,000 - $62,999	-.155*	.009	.000	-.18	-.13
	$63,000 or more	-.279*	.009	.000	-.30	-.26
$39,000 - $47,999	$1 - $38,999	.089*	.008	.000	.07	.11
	$48,000 - $62,999	-.067*	.009	.000	-.09	-.04
	$63,000 or more	-.190*	.009	.000	-.21	-.17
$48,000 - $62,999	$1 - $38,999	.155*	.009	.000	.13	.18
	$39,000 - $47,999	.067*	.009	.000	.04	.09
	$63,000 or more	-.123*	.009	.000	-.15	-.10
$63,000 or more	$1 - $38,999	.279*	.009	.000	.26	.30
	$39,000 - $47,999	.190*	.009	.000	.17	.21
	$48,000 - $62,999	.123*	.009	.000	.10	.15

*. The mean difference is significant at the 0.05 level.

Homogeneous Subsets

Number of procedures

Tukey HSD[a,b]

Median household income national quartile for patient ZIP Code	N	Subset for alpha = 0.05			
		1	2	3	4
$1 - $38,999	140987	1.63			
$39,000 - $47,999	130160		1.72		
$48,000 - $62,999	114555			1.79	
$63,000 or more	100847				1.91
Sig.		1.000	1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 119684.701.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

4.6.4 What are the types of health insurance coverage the groups have?

$H_0$4: There is no statistically significant difference in the types of health insurance coverage the groups have.

H_a4: There is statistically significant difference in the types of health insurance coverage the groups have.

The results in **Table 28** and **Graph 8** show that there were significant differences between the health insurance coverage and the median income groups with 15 degree of freedom, P = 0.000. Specifically, we examined the relationship between SES, median income groups and insurance status. Regardless of the level, Medicare coverage was the primary insurance across all the median income groups. Those without insurance coverage were about 10% and were mostly within the low income group.

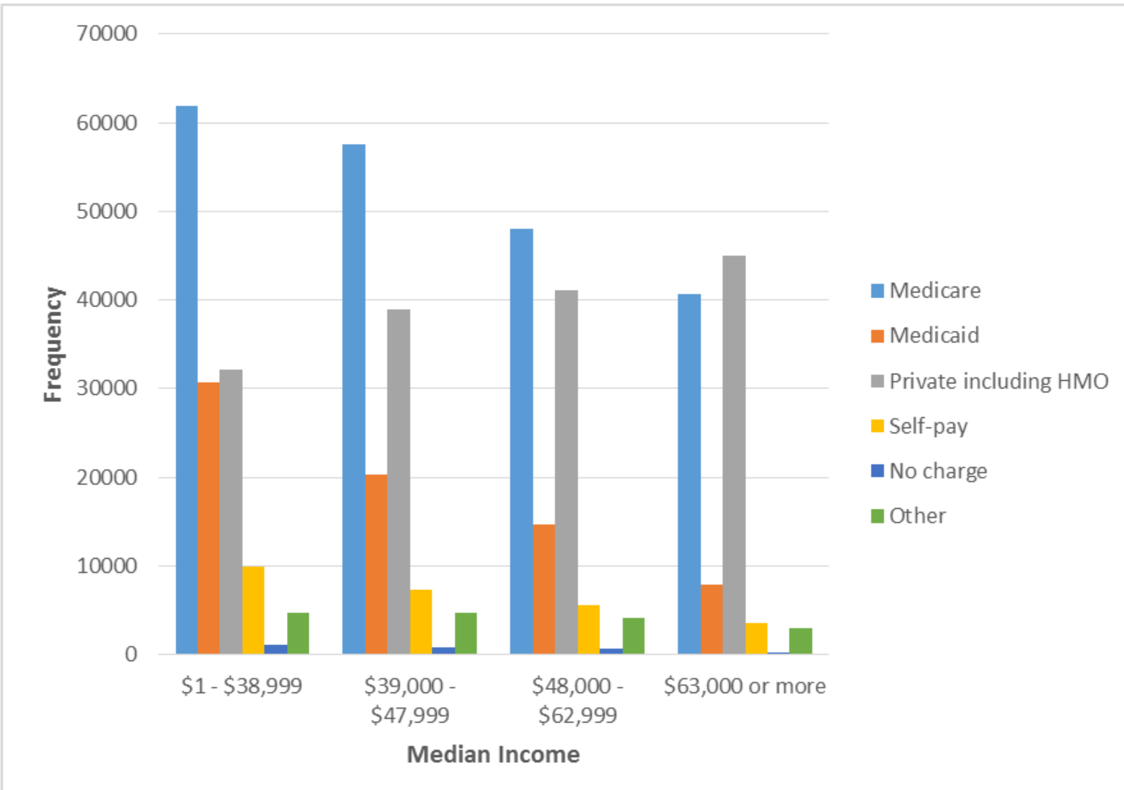
Table 28 Health Insurance Coverage by Median Household Income

			Median household income national quartile for patient ZIP Code				
			$1 - $38,999	$39,000 - $47,999	$48,000 - $62,999	$63,000 or more	Total
Primary expected payer (uniform)	Medicare	Count	61832	57587	48016	40721	208156
		% within Primary expected payer (uniform)	29.7%	27.7%	23.1%	19.6%	100.0%
	Medicaid	Count	30741	20394	14749	7954	73838
		% within Primary expected payer (uniform)	41.6%	27.6%	20.0%	10.8%	100.0%
	Private including HMO	Count	32220	38903	41136	45042	157301
		% within Primary expected payer (uniform)	20.5%	24.7%	26.2%	28.6%	100.0%
	Self-pay	Count	10004	7393	5541	3620	26558
		% within Primary expected payer (uniform)	37.7%	27.8%	20.9%	13.6%	100.0%
	No charge	Count	1078	794	735	297	2904
		% within Primary expected payer (uniform)	37.1%	27.3%	25.3%	10.2%	100.0%
	Other	Count	4757	4810	4127	3041	16735
		% within Primary expected payer (uniform)	28.4%	28.7%	24.7%	18.2%	100.0%
Total		Count	140632	129881	114304	100675	485492
		% within Primary expected payer (uniform)	29.0%	26.8%	23.5%	20.7%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	19389.385[a]	15	.000
Likelihood Ratio	19678.663	15	.000
Linear-by-Linear Association	1069.168	1	.000
N of Valid Cases	485492		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 602.19.



Graph 8 Median Household Income by Health Insurance Coverage by ZIP Code

4.6.5 Are there statistically significant difference in the cost and length of stay amongst the different median income groups?

$H_0$5: There is no statistically significant difference in the costs (TOTCHG) and length of stay (LOS) amongst the different median income groups.

H_a5: There is statistically significant difference in the costs (TOTCHG) and length of stay (LOS) amongst the different median income groups.

To examine the research question, a Multivariate Analysis of Variance (MANOVA) was conducted using the Tukey's Post Hoc test to determine if there is difference in mean LOS and TOTCHG values among the median income groups, and the corresponding significance value. Prior to conducting the MANOVA Pearson correlations were performed between all dependent variables in order to test the MANOVA assumption that the dependent variables would be correlated (multicollinearity) with each other. Other MANOVA assumptions such as normality and homogeneity of variance were met.[67,68] We also looked at the homogeneity of variance and concluded that variance between the groups is equal. After all assumptions were checked, it was determined that MANOVA was appropriate statistical analysis used to model the continuous dependent variables by the discrete independent variables. For this research question, the continuous dependent variables were LOS and TOTCHG; the independent, median income groups (1 = $1 to $37, 999; quartile 2 = $38, 000 to $47, 999; quartile 3 = $48, 000 to $62, 999; quartile 4 = $63, 000 or more.), have four levels. In addition, MANOVA assesses whether mean differences among the median income groups on a combination of dependent variables are likely to have occurred by chance. The results in the **Table 29** below show that the test is significant,

Table 29 Cost and Length of Stay by Median Income Group

	Median household income national quartile for patient ZIP Code	Mean	Std. Deviation	N
Length of stay	$1 - $38,999	4.94	6.801	139949
	$39,000 - $47,999	4.68	6.144	128909
	$48,000 - $62,999	4.64	5.951	112312
	$63,000 or more	4.62	6.201	97543
	Total	4.74	6.312	478713
Total charges	$1 - $38,999	31366.37	51074.389	139949
	$39,000 - $47,999	32224.21	50182.520	128909
	$48,000 - $62,999	35558.63	55399.452	112312
	$63,000 or more	38015.20	58355.164	97543
	Total	33935.70	53491.752	478713

Box's Test of Equality of Covariance Matrices[a]

Box's M	9927.649
F	1103.064
df1	9
df2	2004392502789.915
Sig.	.000

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

a. Design: Intercept + ZIPINC_QRTL

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.
Intercept	Pillai's Trace	.373	142188.007[b]	2.000	478708.000	.000
	Wilks' Lambda	.627	142188.007[b]	2.000	478708.000	.000
	Hotelling's Trace	.594	142188.007[b]	2.000	478708.000	.000
	Roy's Largest Root	.594	142188.007[b]	2.000	478708.000	.000
ZIPINC_QRTL	Pillai's Trace	.006	509.600	6.000	957418.000	.000
	Wilks' Lambda	.994	510.348[b]	6.000	957416.000	.000
	Hotelling's Trace	.006	511.096	6.000	957414.000	.000
	Roy's Largest Root	.006	1000.753[c]	3.000	478709.000	.000

a. Design: Intercept + ZIPINC_QRTL

b. Exact statistic

c. The statistic is an upper bound on F that yields a lower bound on the significance level.

Levene's Test of Equality of Error Variances[a]

	F	df1	df2	Sig.
Length of stay	58.564	3	478709	.000
Total charges	251.150	3	478709	.000

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + ZIPINC_QRTL

which means that differences between the median income groups exist. The difference in the cost and LOS was statistically significantly different among the median income groups, Pillais Trace = .37, F-values = 73.8, degree of freedom = 3, P-value = .000. Therefore, we reject the null hypothesis that there is no statistically significant difference in the population means. That is, the median income groups differ in their joint distribution of cost and LOS.

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	Length of stay	8821.766[a]	3	2940.589	73.830	.000
	Total charges	3220626159398.500[b]	3	1073542053132.833	376.067	.000
Intercept	Length of stay	10471394.106	1	10471394.106	262907.532	.000
	Total charges	552404245104388.750	1	552404245104388.750	193509.798	.000
ZIPINC_QRTL	Length of stay	8821.767	3	2940.589	73.830	.000
	Total charges	3220626159434.529	3	1073542053144.843	376.067	.000
Error	Length of stay	19066591.820	478709	39.829		
	Total charges	1366550358689585.000	478709	2854657753.854		
Total	Length of stay	29810918.000	478713			
	Total charges	1921072031722629.000	478713			
Corrected Total	Length of stay	19075413.587	478712			
	Total charges	1369770984848984.000	478712			

a. R Squared = .000 (Adjusted R Squared = .000)

b. R Squared = .002 (Adjusted R Squared = .002)

Finally, to find where the differences lie, a post-hoc test was conducted, see **Table 30**, contains multiple comparisons based on Tukey test. This process tells us which individual median income group means are different from one another, or that only some of the groups differ. The results revealed that for LOS there are significant differences between the groups except for median income groups 2 and 3 (mean difference = .04 and P-value =

.494); groups 2 and 4 (mean difference = .05, P-value = .203); groups 3 and 4 (mean difference = .04 and P-value = .936), that is to say that the length of stay of these categories varies in the rate of hospitalization, while the cost had no individual mean difference across the median income groups.

Table 30 Post Hoc - Multiple Comparisons

Tukey HSD

Dependent Variable	(I) Median household income national quartile for patient ZIP Code	(J) Median household income national quartile for patient ZIP Code	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Length of stay	$1 - $38,999	$39,000 - $47,999	.27*	.024	.000	.21	.33
		$48,000 - $62,999	.30*	.025	.000	.24	.37
		$63,000 or more	.32*	.026	.000	.25	.39
	$39,000 - $47,999	$1 - $38,999	-.27*	.024	.000	-.33	-.21
		$48,000 - $62,999	.04	.026	.494	-.03	.10
		$63,000 or more	.05	.027	.203	-.02	.12
	$48,000 - $62,999	$1 - $38,999	-.30*	.025	.000	-.37	-.24
		$39,000 - $47,999	-.04	.026	.494	-.10	.03
		$63,000 or more	.02	.028	.936	-.05	.09
	$63,000 or more	$1 - $38,999	-.32*	.026	.000	-.39	-.25
		$39,000 - $47,999	-.05	.027	.203	-.12	.02
		$48,000 - $62,999	-.02	.028	.936	-.09	.05
Total charges	$1 - $38,999	$39,000 - $47,999	-857.84*	206.259	.000	-1387.73	-327.96
		$48,000 - $62,999	-4192.26*	214.045	.000	-4742.15	-3642.37
		$63,000 or more	-6648.83*	222.853	.000	-7221.35	-6076.31
	$39,000 - $47,999	$1 - $38,999	857.84*	206.259	.000	327.96	1387.73
		$48,000 - $62,999	-3334.42*	218.087	.000	-3894.69	-2774.14
		$63,000 or more	-5790.99*	226.739	.000	-6373.49	-5208.49
	$48,000 - $62,999	$1 - $38,999	4192.26*	214.045	.000	3642.37	4742.15
		$39,000 - $47,999	3334.42*	218.087	.000	2774.14	3894.69
		$63,000 or more	-2456.57*	233.844	.000	-3057.33	-1855.82
	$63,000 or more	$1 - $38,999	6648.83*	222.853	.000	6076.31	7221.35
		$39,000 - $47,999	5790.99*	226.739	.000	5208.49	6373.49
		$48,000 - $62,999	2456.57*	233.844	.000	1855.82	3057.33

Based on observed means.

The error term is Mean Square(Error) = 2854657753.854.

*. The mean difference is significant at the .05 level.

4.7 Results of Logistic Regression Analysis

Finally, a series of logistic regression analysis as outlined in section 3.9 were performed to examine whether elective hospitalization are correlated with patients' socioeconomic position. First, we will conduct a simple logistic regression using hospital elective admission and income that predicts that relationship between a patient's income and hospital admission. The objective is to answer the question of whether income significantly predicts hospital elective admission. The null hypothesis states that there is no supported relationship between hospital elective admission (Y) and median income (X). Using Equation 1 in section 3.9. Second, we will perform a multiple logistic regression analysis that tests the relationship between age, race, gender, method of payment, comorbidities, hospital region, and admission.

4.7.1 Simple Logistic Regression

Table 31 Simple Logistic Regression with Median Income as the Independent Variable

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	485357	97.1
	Missing Cases	14643	2.9
	Total	500000	100.0
Unselected Cases		0	.0
Total		500000	100.0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
non-elective	0
elective	1

Categorical Variables Codings

		Frequency	Parameter coding (1)	(2)	(3)
Median household income national quartile for patient ZIP Code	$1 - $38,999	140560	.000	.000	.000
	$39,000 - $47,999	129835	1.000	.000	.000
	$48,000 - $62,999	114364	.000	1.000	.000
	$63,000 or more	100598	.000	.000	1.000

Classification Table[a,b]

	Observed		Predicted: Elective versus non-elective admission: non-elective	elective	Percentage Correct
Step 0	Elective versus non-elective admission	non-elective	352122	0	100.0
		elective	133235	0	.0
	Overall Percentage				72.5

a. Constant is included in the model.

b. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-.972	.003	91297.836	1	.000	.378

Variables not in the Equation

			Score	df	Sig.
Step 0	Variables	ZIPINC_QRTL	531.504	3	.000
		ZIPINC_QRTL(1)	3.276	1	.070
		ZIPINC_QRTL(2)	15.773	1	.000
		ZIPINC_QRTL(3)	278.974	1	.000
	Overall Statistics		531.504	3	.000

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	532.915	3	.000
	Block	532.915	3	.000
	Model	532.915	3	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	569948.391[a]	.001	.002

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	.000	2	1.000

Contingency Table for Hosmer and Lemeshow Test

		Elective versus non-elective admission = non-elective		Elective versus non-elective admission = elective		Total
		Observed	Expected	Observed	Expected	
Step 1	1	104853	104853.000	35707	35707.000	140560
	2	93945	93945.000	35890	35890.000	129835
	3	82446	82446.000	31918	31918.000	114364
	4	70878	70878.000	29720	29720.000	100598

Classification Table[a]

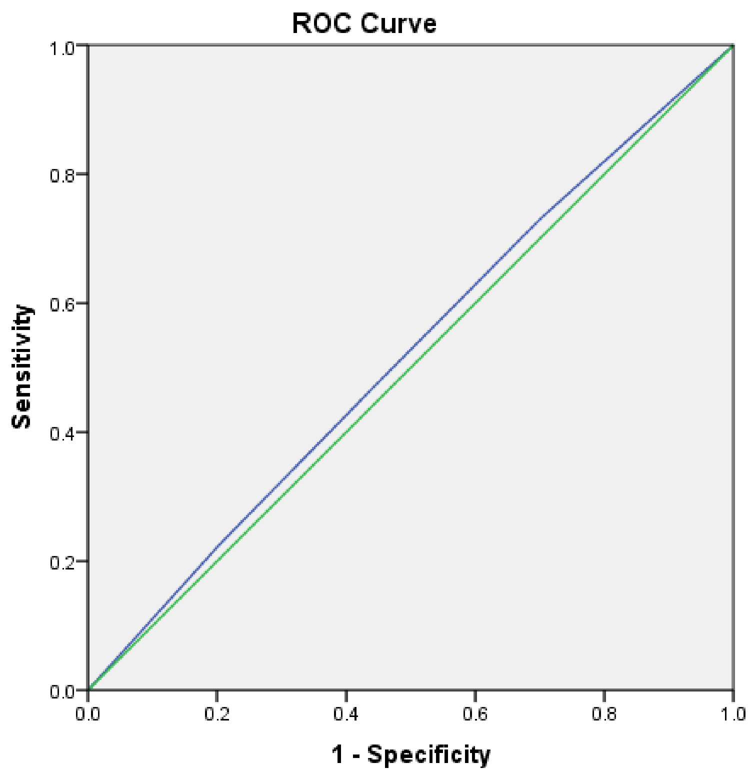
	Observed		Predicted: Elective versus non-elective admission: non-elective	Predicted: Elective versus non-elective admission: elective	Predicted: Percentage Correct
Step 1	Elective versus non-elective admission	non-elective	352122	0	100.0
		elective	133235	0	.0
	Overall Percentage				72.5

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	ZIPINC_QRTL			530.883	3	.000			
	ZIPINC_QRTL(1)	.115	.009	173.783	1	.000	1.122	1.103	1.141
	ZIPINC_QRTL(2)	.128	.009	203.023	1	.000	1.137	1.117	1.157
	ZIPINC_QRTL(3)	.208	.009	507.559	1	.000	1.231	1.209	1.254
	Constant	-1.077	.006	30908.316	1	.000	.341		

a. Variable(s) entered on step 1: ZIPINC_QRTL.



Diagonal segments are produced by ties.

Area Under the Curve

Test Result Variable(s): Predicted probability

Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
.520	.001	.000	.518	.522

The test result variable(s): Predicted probability has at least one tie between the positive actual state group and the negative actual state group. Statistics may be biased.

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

The results in the **Table 31** illustrates the simple logistic regression analysis, which shows that the model has the evidence that income predicts individual's hospital elective admission. The coefficient significance level for all of the coefficients, P-value is less than alpha (.05), we say that the model is statistically significant because df (3), R^2= 532.915, the -2Log Likelihood statistics = 569948.391 and P-value = .000. The model indicates that we can classify our yes cases 72.5% of the time. Using regression equation 1 in section 3.9 to predict values for hospital elective admission based on median income. H_0 states that there is no supported relationship between hospital elective admission and income, (b = 0). Since the coefficients significance level is less than 0.05, we can reject the null hypotheses and concluded that the result is statistical significant evidence that there is a relationship between patients' income and hospital elective admission. The 95% confidence interval for the population slope ranges from 1.103 to 1.254 across the groups. There was evidence that the population slope for elective admission with median income was nonzero. However, we used the logistic regression equation for predicting the dependent variable from the independent variable, we test the differences in hospital elective admission between the median income groups. Using the defined dummies variables in the Table Categorical Variables Coding.

Model: ELECTIVE = β_0 + β_1 median income (1) + β_2 median income (2) + β_3 median income (3)

Where β_0 = Constant; β_1 = Slope, the Table Variable in the Equation shows that, the slope represents how much will a dependent variable change for every unit change in independent variable, median income. From the individual categorical comparisons Exp(B), we can conclude that compared to those in the baseline median income \$1 – 38,999 group, the \$39,000 - \$47,999 group have 1.122 times higher odds (95%CI = 1.103 to 1.141), \$48,000 - \$62,999 group have 1.137 times higher odds (95%CI = 1.117 to 1.157), and \$63,000 or more group have 1.231 times higher odds (95%CI = 1.209 to 1.254) for hospital elective admission to occur. We observed that as the median income increases the likelihood for hospital elective admission increases. Therefore, the simple logistic equation and predicted probability of occurrence are as follows:

$Y = \beta_0 + \beta_1 x_1$

Predicted Logit of (ELECTIVE) = -1.077 + (.115) * Median income Equation 7

From the given equation, the odds of patients' hospital elective admission, ELECTIVE = 1 increases with the increase in median income, which indicates that the higher the median income the more likely that patients will be admitted. The predicted probability indicate the likelihood of y =1. If the predicted probability is greater than 0.5 we can predict that y=1, otherwise y =0. In other words, an increase in x makes the outcome of y more or less likely. The odds ratio of x means that the outcome y= 1 is x as likely as the outcome of y=0. We can now use the model to predict the odds ratio of an ELECTIVE admission, which increase by 1.112 times, that is, odds = $e^{\beta 0 + \beta 1 x 1}$. If median income is

\$39,000 - \$47,999, then the intercept-only model is ln(odds) = .115. If we exponentiate both sides of the expression we find that our predicted odds [(Exp(B)] = ($e^{.115}$ = 1.122). That is to say, the predicted odds of deciding that y = 1 is 1.122 chances. For example, the odds ratio for a 5 change income yielded, $e^{\beta 1*5}$ = ($e^{(.115)*5}$ = 1.777 times). In accordance with other previous results, the statistical significance of the coefficient shows a positive influences on the relationship between income and Elective hospital admission. The overall percentage of the model shows 72.5% of the time the model will predict positively. A measure of goodness-of-fit often used to evaluate the fit of a logistic regression model is based on the simultaneous measure of sensitivity (True positive) and specificity (True negative) for possible cutoff points. The receive operating characteristic (ROC) curve plot a positive or sensitivity versus false-positive in terms of true-positive rate (sensitivity) and false-positive rate (1-specificity) for discriminating correctly at a percentage rate of 72.5%. The area under the curve is .520 with 95% confidence interval (.518, .522). Also, the area under the curve is significantly difference from 0.5 since p-value is 0.000 meaning that the logistic regression classifies the group significantly better than by chance.

4.7.2 Multiple Logistic Regression

As shown in **Table 32**, a multiple logistic regression analysis was employed to predict the probability that a participant would be admitted by elective admission in the hospital. The validity of the model was performed with the assumption that for unit of change from one categorical variable to another, the odds ratio that elective admission was made will equal the odds ratio of the coefficients, considering all other variables are unchanged of the multiple logistic regression analysis used to predict the likelihood of hospital elective admission, ELECTIVE = 1. The predictor variables were participant's age, gender, race, method of payment, median income, hospital region and elective as the outcome variable. A test of the model was statistically significant, $X^2(6, N= 500,000) = 8164.78$, $p < 0.05$. The model correctly classified the outcome for an overall rate at 73.1% and the model shows goodness of fit. The Table variables in the equation shows the logistic regression coefficient, Wald test, and Exp(B), odds ratio for each of the predictors. A .05 criterion of statistical significance was used and all the predictors had a p-value < .000. The odds ratio for female indicates that when holding all other variables constant, there was 1.419 times more likely to be accepted through elective admission compared to others. Since the observed significance level was less than 0.005, we can reject the null hypothesis that there is no linear relationship between hospital elective admission and the predictor variables.

Using the equation 5 in section 3.9, we can write the multiple regression equation as

Predicted Logit of (ELECTIVE) = -.754 -.011*AGE + .350*FEMALE -.084*RACE + .042*PAY1 + .009*HOSP_REGION + .056*ZIPINC_QRTL

A logistic regression model allows us to establish a relationship between a binary outcome variable and a group of predictor variables. It models the logit transformed probability as a linear relationship with the predictor variables. Below is the logit model:

$$\text{Logit (p)} = \ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 x_1 + \ldots + \beta_k * xk \ \text{ antilog } \rightarrow \left(\frac{P}{1-P}\right) = e^{\beta 0 + \beta 1x1} \quad \text{Equation 8}$$

Apply such a model to our dataset, each estimated coefficient is the expected change in the log odds of an elective admission for a unit increase in the corresponding predictor variable holding the other predictor variables constant at certain value. More so, each exponentiated coefficient is the ratio of two odds, or the change in the multiplicative scale for a unit increase in the corresponding predictor variable holding other variables at certain value. For example.

Predicted hospital elective admission =

Logit(p) = log(p/(1-p))= $\beta_0 + \beta_1$*Age + β_2 Female + β_3*Race + β_4*Pay1 + β_5 *Hosp_Region + β_6*ZIPINC_QRTL

The fitted model indicates that, holding Age, Race, Pay1, Hosp_region and ZIPINC_QRTL at a fixed value, the odds of elective admission for Female (female = 1) over the odds of elective admission for males (Female = 0) is exp (.350) = 1.419. The ROC curve validated the good fit of the model in the data sets 2008 – 2010.

Table 32 Multiple Linear Regression with Age, Insurance, and Region as Predictors

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	408491	81.7
	Missing Cases	91509	18.3
	Total	500000	100.0
Unselected Cases		0	.0
Total		500000	100.0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
non-elective	0
elective	1

Classification Table[a,b]

	Observed		Predicted: Elective versus non-elective admission: non-elective	Predicted: Elective versus non-elective admission: elective	Predicted: Percentage Correct
Step 0	Elective versus non-elective admission	non-elective	298459	0	100.0
		elective	110032	0	.0
	Overall Percentage				73.1

a. Constant is included in the model.

b. The cut value is .500

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-.998	.004	80050.046	1	.000	.369

Variables not in the Equation

			Score	df	Sig.
Step 0	Variables	AGE	4918.543	1	.000
		FEMALE	3056.117	1	.000
		RACE	178.363	1	.000
		PAY1	1859.723	1	.000
		HOSP_REGION	23.191	1	.000
		ZIPINC_QRTL	324.718	1	.000
	Overall Statistics		8110.288	6	.000

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	8164.779	6	.000
	Block	8164.779	6	.000
	Model	8164.779	6	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	467828.014[a]	.020	.029

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	1209.243	8	.000

Contingency Table for Hosmer and Lemeshow Test

		Elective versus non-elective admission = non-elective		Elective versus non-elective admission = elective		
		Observed	Expected	Observed	Expected	Total
Step 1	1	32800	33695.622	8018	7122.378	40818
	2	31283	32598.056	9521	8205.944	40804
	3	31863	31777.099	8986	9071.901	40849
	4	32099	31110.962	8753	9741.038	40852
	5	31475	30589.000	9447	10333.000	40922
	6	30846	29855.692	9981	10971.308	40827
	7	30033	28982.047	10818	11868.953	40851
	8	28322	27894.319	12514	12941.681	40836
	9	26093	26740.834	14757	14109.166	40850
	10	23645	25215.368	17237	15666.632	40882

Classification Table[a]

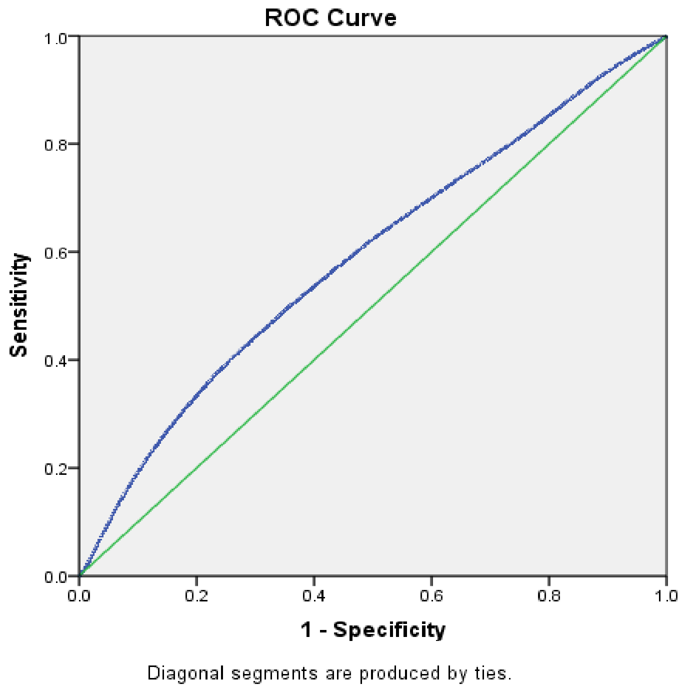
			Predicted		
			Elective versus non-elective admission		Percentage Correct
	Observed		non-elective	elective	
Step 1	Elective versus non-elective admission	non-elective	298459	0	100.0
		elective	110032	0	.0
	Overall Percentage				73.1

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B) Lower	95% C.I.for EXP(B) Upper
Step 1[a]	AGE	-.011	.000	2580.036	1	.000	.990	.989	.990
	FEMALE	.350	.008	2137.534	1	.000	1.419	1.398	1.440
	RACE	-.084	.003	624.281	1	.000	.920	.914	.926
	PAY1	.042	.003	159.258	1	.000	1.042	1.036	1.049
	HOSP_REGION	.009	.003	7.576	1	.006	1.009	1.003	1.016
	ZIPINC_QRTL	.056	.003	304.840	1	.000	1.057	1.051	1.064
	Constant	-.754	.022	1172.421	1	.000	.471		

a. Variable(s) entered on step 1: AGE, FEMALE, RACE, PAY1, HOSP_REGION, ZIPINC_QRTL.



Area Under the Curve

Test Result Variable(s): Predicted probability

Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval Lower Bound	Asymptotic 95% Confidence Interval Upper Bound
.590	.001	.000	.588	.592

The test result variable(s): Predicted probability has at least one tie between the positive actual state group and the negative actual state group. Statistics may be biased.

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

CHAPTER V DISCUSSION AND LIMITATIONS

5.1 Introduction

This chapter provides a brief overview of the study, including a statement of the problem. The majority of this chapter is devoted to a summary and discussion of the study hypotheses and to a discussion of the pertinence of the results that examine the association between socioeconomic status by income and hospitalization characteristics in the U.S. population.

5.1.1 Summary of the Study Problem

Health and economic disparities are factors that consistently determine the life quality of Americans. Factors such as education, income, and occupation are suggested to be the measures that influence SES. Although many articles have argued that health disparities have decreased substantially in the U.S, there is substantial evidence from diverse studies that affirm that variations in SES, which is determined by income, education, occupation, and other indirect factors, are associated with increasing health issues. As SES continue to impact health disparities, urgent efforts and remediation are being initiated to reduce health inequity in the U.S. Due to this problem, higher hospitalization rates in the United States, especially among racial and ethnic minorities with low income continue to be a challenge. Recent research has demonstrated that social risk factors, such as substance abuse, missed clinic visits, multiple address changes, and excessive emergency department use, are associated with higher risk of readmission,[81] which ultimately impacts high hospitalization rate in the U.S especially among the low SES populations.

In order to curb this health issue, many health care programs have been implemented. The most recent initiative is the Delivery System Reform Incentive Payment (DSRIP) Program, a federal approved initiative that is geared toward pay-for-performance, designated to public hospitals that receive essential Medicaid dollars, which supports provision of health services to low income patients. In spite of initiatives like DSRIP, health disparity caused by SES remains a major problem in the U.S. To date, researchers have been unable to fully understand the factors that influence socioeconomic status.

The overall purpose of this study was to identify and further understand income factors that may contribute to SES and hospitalization. In this manner, the study sought to fill the gap in the research on how SES by income impacts hospitalization in the U.S. The assumption of this study was that a better understanding of income influences hospitalization could provide a key input into policy decisions that will reduce health inequalities in the U.S.

Acknowledging that the integration of models and theories is generally lacking, a theoretical framework, a simplified model of pathways from socioeconomic status to health was used as basis for the inquiry into these variables. Other variables that were identified were heavily based on an extensive review of the available literature.

This study was conducted using the NIS survey data set from 2008 to 2010, with sample of (N=500000) discharge data. A quantitative research design was used to examine the relationship between SES by income and hospitalization outcomes. We used the median household income by zip code as our measure for socioeconomic status. In addition, the hypotheses outlined in this study provided insights that permit us to use statistical procedures to measure the health disparities between the socioeconomic groups in the U.S. Five hypotheses were formulated for this study.

5.2.1 Socioeconomic Status and Demographic Characteristics – Hypothesis 1

The first hypothesis postulates significant differences between the four categories of the median income groups and hospitalized patients' demographic, gender and age. These patient characteristics confirm to be some of the predictors when it comes to health. Our findings confirm that SES and patient's demographic characteristics such as age and gender have statistically significant relationship with hospitalization across all groups. The results presented that women were faced with higher hospitalization rate across the income levels compared to men, as has expected. In addition, age group 59-73 also encountered increased hospitalization. This finding was contrary to the other studies that reported that older age groups are more likely to face ill-health and therefore encounter higher hospitalization rates.[70] Based on our data sets, patient between the ages of 74 and 89 had less hospitalization compared to patients between the ages of 59 - 73. Demographic characteristics predicted that hospitalization of the different median income groups progressively decline among patients' residential zip codes. This suggests that there are considerable variations in the association across median income and demographic characteristics such as gender and age.

SES has different influences according to gender and age group, and studies examining the relationship between SES and hospitalization have reported that the relationship varied depending on gender, race, and the degree of development.[71] A study has explained that this occurs because the impact of household income and education are different in men and women.[72 73]

5.2.2 Disparities in Hospital Cost among the Median Income Groups–Hypothesis 2

The second hypothesis supports that there is statistically significant difference between the categorical median income groups and hospital charges in the 2008–2010 discharge records. There is systematic cost differences in health status between the different socioeconomic groups. Based on our study results, there was a 95% confidence interval that patients in the higher income group tend to pay more hospital cost compared to the patients in the lower income spectrum. Another explanation to this findings could be that genetic and constitutional variations that is hereditary and unfavorable laws ensure these disparities exist among individuals, just like their physical characteristics.

The prevalence of ill health differs between different income groups, age groups, and the type of insurance coverage which is more evident in older than younger people due to the natural ageing process.[74] Successively, the hypotheses outlined in our study provided insights that allowed us to explain with statistical processes the impact of income levels on hospitalization and when tested with ANOVA, using the post hoc test Bonferroni, it was suggested that there were statistically significant differences in the hospital cost among the four categorical income levels and p-value $<.001$. Particularly, this difference was apparent in the high income group. Individuals' socioeconomic position is associated with high hospitalization rate. In addition, gender, age, race, cost, insurance type, and comorbidities, these factors most significantly influence ill-health in the United States.

Other studies that have examined these differences have found low income groups tend to pay more hospital cost than others. This study makes a contribution to the field by providing strong support that the cost of health care have variance among different income positions. From a policy perspective the results of the present study suggest that concerted effort should be made to initiate programs that will not only concentrate on the low income but also for all levels of income groups and to evaluate the cost of health care in general.

5.2.3 Types of Clinical Conditions that Affect the Median Income Groups – Hypothesis 3

The third hypothesis of the study was to establish the differences in the prevalence of clinical conditions, such as the types of comorbidities and procedure, among the median income groups. The results show a dramatic increase in the prevalence of hypertension across the different median income groups. SES has been shown to be an important predictor that influences an individual's access to resources, knowledge of quality of care, and care utilization. Therefore, this health disparity is not fixed or permanent, and generally, if social processes generate these variations then the differences should be amenable or even eradicated. Mortality and morbidity increase with declining social position. This finding is consistent with previous results in other studies. [41 42]

In spite of different health initiatives, the literature reviews supported by our findings indicate that clinical conditions, obstructive pulmonary and bronchiectasis, cardiac, coronary, congestive heart failure, pneumonia, mental and substance abuse, and septicemia diseases remain prevalent among the median income groups. Generally, the comparisons across these clinical conditions show significant differences. Our results demonstrate that in 2008 these clinical conditions and procedures occurred more frequently compared to 2009 and 2010.

Consistent with Healthy People 2010, health disparities across the socioeconomic status are challenges that need immediate resolution in the U.S. According to other reports, although some disparities are diminishing, many opportunities for improvement can still be found. In addition, the ability to monitor and track improvements in health disparities is critical and community-based solutions are key for obtaining the elimination of health care disparities. Also, due to lack of insurance, access to quality care, transportation, and inadequate communication that affect racial and underserved populations,[75] it is critical to put together adequate plans which will eliminate these health disparities.

5.2.4 Socioeconomic Status and Health Insurance – Hypothesis 4

Hypothesis four examined the types of health insurance coverage the median income groups have and if there is a statistically significant difference. This study has shown that there is statistical significant difference in the types of insurance coverages. Our findings show that regardless of race or income, those covered by Medicare and private insurance were more frequently hospitalized compared to other insurance types across the median income groups. But patients with private insurance were more frequently found in the high SES group and were more likely to be whites compared to other races. On the other hand, patients with Medicaid and no insurance coverage were more frequently found in the low SES group and more likely to be blacks and Hispanics. This confirms the report that approximately 47 million people in the U.S. 16% of the estimated 300 million population do not have health insurance, and the lack of health insurance disproportionately affects Hispanics, who accounts for the most uninsured population.[76]

Overall our bivariate and multivariable analysis indicate that indeed there are disparities across the socioeconomic groups based on the types of insurance used. This implies that insurance disparities affect hospitalization. Limited access to insurance is ultimately linked to limited access to health facilities in the urban settings and hospitals with better bed-size capacity.[77] [78] Such arrays of care have shown to result in suboptimal outcomes; therefore, these access-related disparities should be addressed to ensure optimal hospitalization outcomes which will improve resource utilization depending on the residential zip code.

5.2.5 Socioeconomic Status, Cost and Length of Stay – Hypothesis 5

This study hypothesized that there is statistically significant difference in the cost and length of stay among the different median income groups. The increasing cost in care, inequalities in income and a falling standard of living for a large segment of the U.S. population have been reported. The results indicate that the cost of hospitalization of the patients from the lower SES quartiles was less as compared to those from higher SES quartiles. This may have been due to the type of procedures that insurance entities pay for the lower SES and due to the type of insurance coverage they have. However, whether this represents an underutilization of resources for patients from lower SES zip codes or an overutilization of resources for patients from higher SES zip codes is not completely clear from the analysis. Also, the differences in SES have been consistently associated with variations in morbidity.[79]

Our study suggests that with respect to SES and LOS, there is significant difference in the LOS across the median income groups, p-value <0.001. The mean average cost was $33,935.70 and mean length of stay was 4.7 days. The lower income group's ($1 - $38,999) average length of stay was 4.9 days, which was slightly higher than the other median income groups. This was in line with other reports that have demonstrated that low income population experience more frequent hospitalization and a longer mean stays. Fortunately, the average length of stay has drastically deceased in recent time. In addition, the predictive analysis that was conducted using regression indicated there were statistically significant association between patient's demographic such as Age, Gender, Race, Insurance, Hospital region and Median household income and elective admission outcomes in 72.5% overall chances. The ROC curve also confirmed that the regression models were a good fit for the 2008 – 2010 NIS data sets.

5.4 Limitation

Our study encountered several limitations that need be outlined. First, NIS is an administrative database, which may be subject to errors in data collection, procedures and disease classification and coding. The data were not collected for the purpose of this specific study. The use of median household income by zip code for socioeconomic data to individual income leads to possible ecological fallacy,[66] a concept that is associated with a modifiable area unit problem (MAUP). MAUP occurs when arbitrarily defined boundaries are used for the measurement and reporting of spatial phenomena. The ecological fallacy occurs when we make conclusions about individuals based only on analyses of group data. In this stance, the SES of the zip code region may not be similar to that of the individual level analysis. If the data set was provided with street addresses or specific identifiers for individuals, ecological fallacy would have been minimized. However, zip code happens to be the lowest patient's location identifier available in the NIS dataset.

Some other limitations were the lack of educational and occupational data, which are part of the core predictors of SES. We also faced the challenge of comparing the living standard by states due to lack of data. The use of other datasets in conjunction with the NIS data set would have provided more accurate and valid results to our analyses, which would have helped to strengthen our findings and assumptions. Another limitation was that the data set comprised of only community hospitals data and rehabilitation and long-term acute care hospital data were excluded from the NIS data sets. This inadequacy of essential health data contributed to these limitations. Such limitations hinder the adequate distribution of resources to those that really need it. Also, this is a retrospective study, which may be subject to traditional biases such as selection bias. Finally, limitation might result from the fact that we utilized median household income of the entire zip code to compute the SES for each hospitalized patient instead of patient level income information. The fact that the data have been de-identified and there is a difference between looking at individual admissions verses individual patients.

In conclusion, we have demonstrated that there are significant differences in socioeconomic status and health using median income levels. Future research should determine if SES affects other essential hospitalization outcomes in both children and adults, including more defined SES measures such as education, occupation, and individual income.

5.5 Suggestions for Future Research

The socioeconomic status of each patient was determined at a neighborhood level based on postal code of residence, and might not necessarily correspond to socioeconomic status at the individual level. We suggest that a study should be conducted with patient's level income data as well as other socioeconomic position predictors such as education, occupation, environmental factors, biobehavioral, and behavioral data. These would further investigate the probable causes of health disparities impacted by socioeconomic status.

CHAPTER VI CONCLUSIONS AND RECOMMENDATIONS

6.1 Introduction

The final chapter contains the summary and recommendations of the data analysis based on the purpose of the study. The recommendations will provide useful solutions that will help to reduce health disparities.

6.2 Conclusion

The main goal of the study was to examine whether socioeconomic status and income was associated with hospitalization outcomes in the U.S., using the NIS data sets from 2008 to 2010. The reasons and motivation for this project was the assumption that inequalities in SES are responsible for health disparities and have created social imbalance in individual well-being. The general theoretical literature on this subject in the U.S. and globally is inconclusive on several vital questions. In order to accomplish our study purpose, we devised five questions as mentioned in previous chapters.

Over the past decades, research on the socioeconomic determinants of health has increased significantly, and yet the mechanisms by which SES impacts health are yet very unclear. The study findings have shown that SES as determined by income influences hospitalization outcomes and affirmed that health is a reflection of the social inequalities and inequities that exist in the nation. There is evidence that many components of socioeconomic status such as age, gender, income, community, and insurance are related to high hospitalization rates, especially among the low median income group. Even though there have been tremendous improvements in the U.S.A's indicators for health, education, income, wealth accumulation, and employments in the past decades, there is also much room for improvement with new innovations, high growth rates, and influx of immigrants with diverse cultural backgrounds.

Our results clearly indicate that socioeconomic position is a contributing factor that determines where we live, work, learn, and play, which ultimately may have a bigger role than medical care in determining how healthy we are. We have established that high hospitalization rate are prevalent in the regions like the South, thus incorporating more health care programs in these regions is essential in order to reduce hospitalization. Health awareness such as physical activities, diets, regular checkup and follow up appointments with primary care providers can be done through health centers, community organizations and social counselors. These can assist with social issues such as income, evaluation of living conditions and adequate follow up after discharge. Furthermore, this signifies that determinants of good health start before we need medical care. It begins with mechanisms that impact social inequalities. More so, as SES decreases, individuals are exposed to more demands and have fewer resources with which to address them.

As our study depicts, insurance coverage, comorbidities, and LOS ultimately influence health. An effective approach to prevention is well warranted and desired, and in this regard, well-defined policies and health initiatives are very necessary, which will help to eliminate health disparities and improve people's health in the United States. In addition, our study noted that difference in hospital costs were significant across the median income groups and therefore necessitate measures that will reduce health care costs. One particular program that is being implemented is the Delivery System Reform Incentive Payment (DSRIP) Program, a federal approved initiative that is geared toward pay-for-performance, designated to public hospitals that receive essential Medicaid dollars. The fundamental purpose is to restructure the health care delivery system by reinvesting the Medicaid program with the primary goal of reducing avoidable hospital use by 25% over 5 years. A total of $6.42 billion is allocated to this program.[69 80] DSRIP is adopted in New York State. This initiative is aimed at reducing hospital use and to improving public health care delivery system; however, there are many challenges that are hinder the success of the program. These are not limited to conflicts among the local partners that form the Performing Provider System (PPS), disparity in clinical workflows, interoperability, governance, disparity in the Electronic Medical Records System, and sensitive patient information. If these challenges are not combated they will hinder the success of the DSRIP initiative.

6.3 Recommendations

The study found that there are considerable variations in socioeconomic status and hospitalization outcomes across the income levels. It provided evidences that support a strong relationship between income and hospitalization. Based on the study findings, it is recommended that this health program (DSRIP) need to be restructured in order to generate achievable results. From our results, the low-income families are particularly vulnerable to change in social welfare policy in that they do not have the personal resources to compensate for changes in health coverage and hospital costs. The study also suggests that there is need for further assessment of the initiative.

Despite the fact that DSRIP conflicts with other federally funded programs such as Patient Protection Care Act, commonly known as Affordable Care Act (ACA) and Regional Health Information Organization (RHIO), which are essentially designed to improve patient care via: Safety, Care quality, Satisfaction, Clinical outcomes, and Delivery-system efficiencies, it can contribute to the reduction of hospitalization if effectively implemented.

Based on our study results, it is evident that increased hospitalization continues to be a major challenge that the U.S. population face. If implemented successfully, DSRIP can help to reduce avoidable hospital use and therefore can impact health care cost, which eventually will affect SES. In order for this to work, it is necessary to hold providers accountable to patients' health and health costs and linking payments to outcomes. With this economic incentive, providers will incorporate social interventions into care, as well as address clinical needs. This will ultimately impacts SES and hospitalization rates. However, before this can be achieved, we recommend that a well-defined health policy and standardized clinical workflows are essential to the success of the DSRIP program. Currently, most of the policies and workflows inhibit effective patient care. Due to differences in clinical workflows across the care settings, it has become a challenge to the DSRIP initiative in that up to date there are no defined standards in the clinical workflows. As we know, standard is useful because it permits the disassociated people to collaborate. Therefore, having a standardized clinical workflow will assist in achieving the DSRIP goals.

Another challenge is interoperability, there is need for defined data sharing. Consideration that the promise of interoperability is yet to be fulfilled for integrated health delivery services, it contribute to the issues that inhibit the success of DSRIP. It is clear that gathering data and making the data available to providers are central to the health care process. The ability for different information technology systems and software applications to communicate, exchange data, and use the exchanged information is paramount to the success of DSRIP. So, it is important to define how data will be shared within and across the partners in order to advance the effective delivery of healthcare. Without having access to adequate clinical data, providers will be unable to make informed clinical decisions or to deliver optimal care. To meet this need, patient consent is needed in that there are regulations and compliances that prohibit providers to have access to patients' data without being part of the patients' care team. Therefore, patients need to authorize providers to have access to their clinical data even if the provider is not part of the care team. Furthermore, apart from data sharing and patient health information, sensitive patient information such as substance abuse and HIV data are necessary to be shared across clinicians in order to advance the delivery of care. More so, patient's consent is essentially important to achieve DSRIP objectives and to accomplish this, patients need to be aware of the benefits of sharing their health data between clinicians without restrictions.

The fact that in today's world of clinical settings, there is no approved Electronic Medical Records (EMR) System is a challenge to the DSRIP program because data-collection process will be undermined, so by approving an EMR system will help to meet the DSRIP goals. In addition, having local partners to form a PPS is a challenge because there will always be bottlenecks due to undefined governance structure. This lack of governance has become an inhibitor to the DSRIP initiative, therefore, by introducing good governance and well-defined policies, the conflict among PPS will be resolved.

Also, policymakers need to review inpatient reduction, that is, capacity planning. For instance, if a facility does not meet the 25% readmission reduction, what happens to that facility? Who will make the decision to either close or merge the facility with another hospital? Thus, adjustment to the DSRIP models will address the aforementioned challenges and mitigate the entrenched problems in health care delivery system and ultimately improve patient satisfaction. One way to go about changing social inequality and eliminating disparity is through shifting public policy and remodeling several health care services, insurance, and social services. For decision makers, these results emphasize the major effort required to optimize the medical management of high hospitalization rate, improve management of chronic diseases and detection and correction of disorders or disabilities which can be a positive factor for improvement of the socioeconomic groups. In conclusion, preventive, social, economic, and behavioral actions must also be promoted to achieve a better health outcome and improve decision quality in the U.S.

REFERENCES

1. Fox AJ. Longitudinal studies based on vital registration records. Rev Epidemiol Sante Publique. 1989; 37(5–6):443–448.

2. Krieger N, Fee E. Measuring social inequalities in health in the United States: a historical review, 1900–1950. Int J Health Serv. 1996;26(3):391–418.

3. Liberatos P, Link BG, Kelsey JL. The measurement of social class in epidemiology. Epidemiol Rev. 1988; 10:87–121.

4. Susser M, WatsonW, Hopper K. Sociology in Medicine. 3rd ed. New York, NY: Oxford University Press; 1985.

5. Hobfoll SE. The influence of culture, community, and the nested-self in the stress process: Advancing Conservation of Resources theory. Applied Psychology. 2001; 50:337–370.

6. Gallo, L.C., & Matthews, K.A. (2003). Understanding the association between socioeconomic status and physical health: Do negative emotions play a role? Psychological Bulletin, 129, 10–51.

7. American Psychological Association. http://www.apa.org/pi/ses/resources/publications/factsheet-erm.aspx. Accessed 12/06/2014.

8. Mueller CW, Parcel TL. Measures of socioeconomic status: alternatives and recommendations. Chil Dev. 1981;52:13-30.

9. Green L. Manual for scoring socioeconomic status for research on health behavior. Public Health ReCp. 1970;85:815-827.

10. Almond, D. (2006). 'Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 U.S. population'. Journal of Political Economy 114: 562-712.

11. National Center for Health Statistics, Health, United States, 2006: With Chartbook on Trends in the Health of Americans. 2006.

12. J.P. Bunker, "Surgical Manpower: A Comparison of Operations and Surgeons in the United States and in England and Wales," The New England Journal of Medicine 282 (1970): 135-144; J.E. Wennberg and A. Gittelsohn, "Small Area Variations in Health Care Delivery," Science 182 (1973): 1102-1108; M.R. Chassin et al., "Variations in the Use of Medical and Surgical Services by the Medicare Population,"

The New England Journal of Medicine 314 (1986): 285-290; and M.R. Chassin et al., “Does Inappropriate Use Explain Variations in the Use of Health Care Services?” Journal of the American Medical Association 258 (1987): 2533-2537.

13. Pappas, G. et al., “The Increasing Disparity in Mortality between Socioeconomic Groups in the United States, 1960 and 1986,” The New England Journal of Medicine, Vol. 392 No. 2 (1993).

14. Healthy People 2020, http://www.healthypeople.gov/2020/about/foundation-health-measures/Disparities. Access 12/06/2014

15. Franz, M., Lensche, H., & Schmitz, N. (2002). Psychological distress and socioeconomic status in single mothers and their children in a German city. Social Psychiatry and Psychiatric Epidemiology, 38, 59–68.

16. Farr,W. 1975. Vital statistics: amemorial volume of selections from the reports and writings of William Farr. Introduction by Mervyn Susser and Abraham Adelstein. Scarecrow Press. Metuchen, NJ.

17. Braveman P, Gruskin S. 2003. Defining equity in health. J. Epidemiol. Community Health 57:254–58.

18. Braveman P, Krieger N, Lynch J. 2000. Health inequalities and social inequalities in health. Bull. WHO 78:232–34.

19. Braveman P, Starfield B, Geiger HJ. 2001. World Health Report 2000: How it removes equity from the agenda for public health monitoring and policy. BMJ 323:678–81.

20. Andresen FM, Miller DK. The future (history) of socioeconomic measurement and implications for improving health outcome among African Americans. J Gerontol A Biol Sci Med Sci. 2005;60:1345-350.

21. Krieger N, Williams DR, Moss NE. Measuring social class in US public health research: concepts, methodologies, and guidelines. Annu Rev Public Health 1997;18:341–78.

22. Lynch J, Kaplan G. Socioeconomic position. In: Berkman LF, Kawachi I, eds. Social epidemiology. 1st ed. Oxford: Oxford University Press, 2000:13–35.

23. Galobardes B, Shaw M, Lawlor DA, Lynch JW, Davey Smith G (2006) Indicators of socioeconomic position (part 1). J Epidemiol Community Health 60: 7–12.

24. Lillie-Blanton, M. et al., “Reducing Racial Ethnic, and Socioeconomic Disparities in Health Care: Opportunities in National Health Reform,” Health Care Winter 2008

25. Adler NE, Newman K (2002) Socioeconomic disparities in health: pathways and policies. Health Aff (Millwood) 21: 60–76.

26. Miech RA, Hauser RM. Socioeconomic status (SES) and health at midlife; a comparison of educational attainment with occupation-based indicators. Ann Epidemiol. 2001;11:75-84.

27. Andersen, R., Newman J.F.,(1969) "Societal and Individual Determinants of Medical Care Utilization in the United States"

28. Anderson, O. W., Feldman J. J., (1956). "Family Medical Costs and Voluntary Health Insurance: A Nationwide Survey". New York: McGraw-Hill.

29. Lleras-Muney, A. (2005), "The relationship between education and adult mortality in the United States", Review of Economic Studies 72(1): pp. 189-221.

30. Oreopoulos, P. (2006), "Estimating average and local average treatment effects of education when compulsory schooling laws really matter", American Economic Review, 96(1): pp. 152-175

31. Silles, M. (2009), "The causal effect of education on health: evidence from the United Kingdom", Economics of Education Review, 28(1): pp. 122-128.

32. Smith, J.P. (2007), "The impact of socioeconomic status on health over the life course", Journal of Human Resources 42(4): pp. 739-764.

33. Cutler, D.M., Lleras-Muney, A. and Vogl, T. (2010), "Socioeconomic Status and Health: Dimensions and Mechanisms", Oxford Handbook of Health Economics, forthcoming.

34. Mojtabai R, Olfson M: Medication costs, adherence, and health outcomes among Medicare beneficiaries. Health Aff (Millwood) 2003, 22:220-229.

35. Goldman, D.P. and Smith, J.P. (2002), "Can Patient Self-Management Help Explain the SES Health Gradient?", Proceedings of the National Academy of Science 99: pp. 10929-10934.

36. Backlund E, Sorlie PD, Johnson NJ. The shape of the relationship between income and mortality in the United States. Evidence from the national longitudinal mortality study. Ann Epidemiol 1996;6:12–20.

37. Ecob R, Davey Smith G. Income and health: what is the nature of the relationship? Soc Sci Med 1999;48:693–705.

38. Hollingshead AB. The indiscriminate state of social class measurement. Soc Forces.1971;49:563-567.

39. Duncan OD. Asocioeconomic indexfor all occupations. In: Reiss AJ Jr, Duncan OD,Hatt PK, North CC, eds. Occupation and Social Status. New York: The Free Press of Glencoe; 1&61:109-138.

40. Philbin EF, Dec GW, Jenkins PL, Di Slavo TG: Socioeconomic status as an independent risk factor for hospital readadmission for health failure. Am J Cardiol 2001, 87:1367-1371.

41. C.E. Ross and J. Mirovsky: Does Unemployment Affect Health? Journal of Health and Social Behavior 36, no. 3 (1995): 230-243; and S.H. Wilson and G.M. Walker, "Unemployment and Health: A Review," Public Health 107, no. 3 (1993): 153-162.

42. Rodriguez E: "Keeping the Unemployed Healthy: The Effect of Means-Tested and Entitlement Benefits in Britain, Germany, and the United States," American Journal of Public Health 91, no. 9 (2001): 1403-1411.

43. Case, A., Deaton, A. (2005), "Broken Down by Work and Sex: How our Health Declines", in: David A. Wise (ed.) (2005), Analyses in the Economics of Aging, The University of Chicago Press, Chicago, pp. 185-212.

44. Viscusi, K.W. (1978), "Wealth Effects and Earnings Premiums for Job Hazards", Review of Economics and Statistics 60: pp. 408-416.

45. Viscusi, K.W. (1979), Employment hazards: An investigation of market performance, Cambridge: Harvard U. Press.

46. http://jama.jamanetwork.com/article.aspx?articleid=187597. Accessed on 1/28/15.

47. Warner DF, Hayward MD. 2002. Race disparities in men's mortality: The role of childhood social conditions in a process of cumulative disadvantage. Paper presented at the annual meetings of the Population Association of America, Atlanta, Jan. 9-15.

48. Espino DV, Burge SK, Moreno CA. 1991. The prevalence of selected chronic diseases among the Mexican-American elderly: Data from the 1982-1984 Hispanic Health and Nutrtion Examination Survey. Journal of the American Board of Family Practice 4:217-222.

49. Jasso G, Massey DS, Rosenzweig MR, Smith JP. 2000. The New Immigrant Survey Pilot (NIS-P): Overview and new findings about U.S. legal immigrants at admission. Demography 37:127-138.

50. LaVeist TA, Wallace JM Jr. 2000. Health risk and inequitable distribution of liquor stores in African American neighborhoods. Social Science and Medicine 51:613-617.

51. Morrison RS, Wallenstein S, Natale DK, Senzel RS, Huang L-L. 2000. "We don't carry that"— ailure of pharmacies in predominantly nonwhite neighborhoods to stock opiod analgesics. New England Journal of Medicine 342:1023-1026.

52. The Institute of Medicine (IOM) report 2003. Unequal Treatment: Confronting Racial and Ethnic Disparities in Healthcare.

53. Cooper, R. (1994). A case study in the use of race and ethnicity in public health surveillance. Public Health Reports, 109, 46-52.

54. Oman, R. F, et al., (2002). An adolescent age group approach to examining youth risk behaviors. American Journal of Health Promotion, 16(3), 167-716.

55. Rogers RG, Hummer RA, Nam CB. 2000. Living and Dying in the USA. San Diego, CA: Academic Press.

56. Adler, N. E., Boyce, T., Cheseney, M. A., Cohen, S., Folkman, S., Kahn, R. L., et al. (1994). Socioeconomic status and health: The challenge of the gradient. American Psychologist, 49, 15–24.

57. Barnett, E., Casper, M.L., Halverson, J.A., Elmes, G.A., Braham, V.E., Majeed, Z.A., Bloom, A.S., and Stanley, S. (2001). Men and heart disease: An atlas of racial and ethnic disparities in mortality (1st ed.). Office for Social Environment and Health Research. Morgantown: West Virginia University.

58. Institute of Medicine. 2003. Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care. Smedley BD, Stith AY, Nelson AR, eds. Committee on Understanding and Eliminating Racial and Ethnic Disparities in Health Care, Board on Health Sciences Policy. Washington, DC: The National Academies Press.

59. Blendon R, Aiken L, Freeman H, Corey C. 1989. Access to medical care for black and white Americans. Journal of the American Medical Association 261:278-281.

60. Marmot, M.G., Shipley, M., and Rose, G. Inequalities in death—Specific explanations of a general pattern. Lancet (May 5, 1984) 1:1003-6.

61. Rose, G., and Marmot, M.G. Social class and coronary heart disease. British Heart Journal (1981) 45,1:13-19.

62. U.S. Department of Health and Human Services. Health status of minorities and low-income groups. 3rd ed. Washington, DC: U.S. Government Printing Office, 1991.

63. Tashakkori, A., & Teddlie, C. (Eds.). (2003). Handbook on mixed methods in the behavioral and social sciences. Thousand Oaks, CA: Sage Publications.

64. Agency for Healthcare Research and Quality, Healthcare Cost and Utilization Project. Nationwide Inpatient Sample (NIS), 2008-2010 [cited 2015 February 15]. http;//www.hcup-us.ahrq.gov/nisoverview.jsp

65. Nancy E. Adler and Alana Conner Snibbe. The Role of Psychosocial Processes in Explaining the Gradient Between Socioeconomic Status and Health Current Directions in Psychological Science August 2003 12: 119-123.

66. Schwartz, S. "The Fallacy of the Ecological Fallacy: The Potential Misuse of a Concept and the Consequences." American Journal of Public Health 84.5 (1994): 819–824.

67. Tabachnick, B. G. & Fidell, L. S. (2012). Using multivariate statistics (6th ed.). Boston, MA: Pearson

68. Meyers, L.S., Gamst, G., & Guarino, A. (2006). Applied multivariate research: Design and interpretation. Thousand Oaks, CA: Sage Publishers.

69. http://www.health.ny.gov/health_care/medicaid/redesign/dsrip/overview.htm

70. Hauch A. et. Al. (2014) Association of Socioeconomic Status, Race, and Ethnicity with Outcomes of Patient Undergoing Thyroid Surgery. JAMA. Access on Feb. 2015.

71. Lillie-Blanton M., et at (2008) Reducing Racial, Ethnic, and Socioeconomic Disparities in Health Care. JLME. Accessed Mar. 2015.

72. Robbins JM, Vaccarino V, Zhang H, Kasl SV (2001) Socioeconomic status and type 2 diabetes in African American and non-Hispanic white women and men: evidence from the Third National Health and Nutrition Examination Survey. Am J Public Health 91: 76–83. PMID: 11189829

73. Rabi DM, Edwards AL, Southern DA, Svenson LW, Sargious PM, et al. (2006) Association of socioeconomic status with diabetes prevalence and utilization of diabetes care services. BMC Health Serv Res 6: 124. PMID: 17018153

74. Trivedi A., et al. (2005) Trends in the Quality of Care and Racial Disparities in Medicare Managed Care. NEJM. Accessed on Dec. 2014.

75. Sun M.. et al. (2014) Disparities in selective referral for cancer surgeries: implications for the current healthcare delivery system. BMJ Open. Accessed on Jan. 2015.

76. Pfizer, A profile of uninsured persons in the United States. www.pfizer.com .access on Feb. 2015.

77. Khuri SF, Henderson WG. The case against volume as a measure of quality of surgical care. World J Surg 2005;29:1222–9.

78. Keeler EB, Rubenstein LV, Kahn KL, et al. Hospital characteristics and quality of care. JAMA 1992;268:1709–14.

79. Jia H, Lubetkin EI. The statewide burden of obesity, smoking, low income and chronic diseases in the United States. J Public Health 2009;31:496–505.

80. http://www.health.ny.gov/health_care/medicaid/redesign/dsrip/overview.htm

81. Amarasingham R, et all. An automated model to identify heart failure patients at risk for 30-day readmission or death using electronic medical record data. Med Care. 2010;48:981–988.

APPENDICES

APPENDIX A: Table 2 Data elements in the 2008-2010 NIS file

Data Element	Name	Description
Admission information		
Admission day	AWEEKEND	Admission on weekend: (0) admission on Monday-Friday, (1) admission on Saturday-Sunday
	ADAYWK	*Admission day of week: (1) Sunday, (2) Monday, (3) Tuesday, (4) Wednesday, etc.*
Admission month	AMONTH	Admission month coded from (1) January to (12) December
Admission source	ASOURCE	Admission source, uniform coding: (1) ER, (2) another hospital, (3) another facility including long-term care, (4) court/law enforcement, (5) routine/birth/other
Admission type	ELECTIVE	Indicates elective admission: (1) elective, (0) non-elective admission
Patient demographic location information		
Age at admission	AGE	Age in years coded 0-124 years
	AGEDAY	Age in days coded 0-365 only when the age in years is less than 1
Sex of patient	FEMALE	Indicates gender for NIS beginning in 1998: (0) male, (1) female
	SEX	*Indicates gender for NIS prior to 1998: (1) male, (2) female*
Race of patient	RACE	Race, uniform coding: (1) white, (2) black, (3) Hispanic, (4) Asian or Pacific Islander, (5) Native American, (6) other
Location of patient's residence	PL_NCHS2006	Patient Location: NCHS Urban-Rural Code (V2006). This is a six-category urban-rural classification scheme for U.S. counties: (1) "Central" counties of metro areas of >=1 million population,(2) "Fringe" counties of metro areas of >=1 million population,(3) Counties in metro areas of 250,000-999,999 population,(4) Counties in metro areas of 50,000-249,999 population,(5) Micropolitan counties,(6) Not metropolitan or micropolitan counties
	PL_UR_CAT4	*Urbanâ€"rural designation for patient's county of residence: (1) large metropolitan, (2) small metropolitan, (3) micropolitan, (4) non-metropolitan or micropolitan*

Median household income for patient's ZIP Code	ZIPINC_QRTL	Median household income quartiles for patient's ZIP Code. For 2008, the median income quartiles are defined as: (1) $1 - $38,999; (2) $39,000 - $47,999; (3) $48,000 - 62,999; and (4) $63,000 or more.
Payer information		
Primary expected payer	PAY1	Expected primary payer, uniform: (1) Medicare, (2) Medicaid, (3) private including HMO, (4) self-pay, (5) no charge, (6) other
Secondary expected payer	PAY2	Expected secondary payer, uniform: (1) Medicare, (2) Medicaid, (3) private including HMO, (4) self-pay, (5) no charge, (6) other
Diagnosis and procedure information		
ICD-9-CM diagnoses	DX1 - DX25	Diagnoses, principal and secondary (ICD-9-CM). Beginning in 2003, the diagnosis array does not include any external cause of injury codes. These codes have been stored in a separate array ECODEn. Beginning in 2009, the diagnosis array was increased from 15 to 25.
	NDX	Number of diagnoses coded on the original record
	DSNDX	*Number of diagnosis fields provided by the data source*
	DXSYS	*Diagnosis coding system (ICD-9-CM)*
	DXV1 - DXV15	*Diagnosis validity flags*
ICD-9-CM procedures	PR1 - PR15	Procedures, principal and secondary (ICD-9-CM)
	NPR	Number of procedures coded on the original record
	DSNPR	*Number of procedure fields in this data source*
	PRSYS	*Procedure system (1) ICD-9-CM, (2) CPT-4, (3) HCPCS/CPT-4*
	PRV1 - PRV15	*Procedure validity flag: (0) Indicates a valid and consistent procedure code, (1) Indicates an invalid code for the discharge date*
	PRDAY1	Number of days from admission to principal procedure.
	PRDAY2 - PRDAY15	Number of days from ad mission to secondary procedures
DRG information		
Diagnosis Related Group (DRG)	DRG	DRG in use on discharge date
	DRG_NoPOA	DRG in use on discharge date, calculated without Present On Admission (POA) indicators
	DRGVER	Grouper version in use on discharge date
	DRG10	*DRG Version 10 (effective October 1992 - September 1993)*

	DRG18	*DRG Version 18 (effective October 2000 - September 2001)*
	DRG24	DRG Version 24 (effective October 2006 - September 2007)
Major Diagnosis Category (MDC)	MDC	MDC in use on discharge date
	MDC_noPOA	MDC in use on discharge date, calculated without Present on Admission (POA) indicators
	MDC10	*MDC Version 10 (effective October 1992 - September 1993)*
	MDC18	*MDC Version 18 (effective October 2000 - September 2001)*
	MDC24	MDC Version 24 (effective October 2006 - September 2007)
Data Elements in the NIS Disease Severity Measures File and, Data Elements in the NIS Diagnosis and Procedures Groups File		
Clinical Classifications Software (CCS) category	DXCCS1 - DXCCS25	Clinical Classifications Software (CCS) category for all diagnoses for NIS beginning in 1998. Beginning in 2009, the diagnosis array was increased from 15 to 25.
	DCCHPR1	*CCS category for principal diagnosis for NIS prior to 1998. CCS was formerly called the Clinical Classifications for Health Policy Research (CCHPR).*
	E_CCS1 - E_CCS4	CCS category for the external cause of injury and poisoning codes
	PRCCS1 - PRCCS15	CCS category for all procedures for NIS beginning in 1998
	PCCHPR1	*CCS category for principal procedure for NIS prior to 1998. CCS was formerly called the Clinical Classifications for Health Policy Research (CCHPR)*
Number of chronic conditions	NCHRONIC	Count of chronic conditions in the diagnosis vector
Operating room procedure indicator	ORPROC	Major operating room procedure indicator for the record: (0) no major operating room procedure, (1) major operating room procedure
Resource use information		
Total charges	TOTCHG	Total charges, edited
	TOTCHG_X	Total charges, as received from data source
Length of stay	LOS	Length of stay, edited
	LOS_X	Length of stay, as received from data source
Discharge information		
Discharge quarter	DQTR	Coded: (1) First quarter, Jan - Mar, (2) Second quarter, Apr - Jun, (3) Third quarter, Jul - Sep, (4) Fourth quarter, Oct - Dec

	DQTR_X	Discharge quarter, as received from data source
Discharge year	YEAR	
Disposition of patient (discharge status)	*DISP*	*Disposition of patient, uniform coding used prior to 1998: (1) routine, (2) short-term hospital, (3) skilled nursing facility, (4) intermediate care facility, (5) another type of facility, (6) home health care, (7) against medical advice, (20) died*
	DIED	Indicates in-hospital death: (0) did not die during hospitalization, (1) died during hospitalization
	DISPUNIFORM	Disposition of patient, uniform coding used beginning in 1998: (1) routine, (2) transfer to short-term hospital, (5) other transfers, including skilled nursing facility, intermediate care, and another type of facility, (6) home health care, (7) against medical advice, (20) died in hospital, (99) discharged alive, destination unknown
	TRAN_OUT	Transfer Out Indicator: (0) not a transfer, (1) transferred out to a different acute care hospital, (2) transferred out to another type of health facility
Weights (to calculate national estimates)		
Discharge weights	DISCWT	Discharge weight on Core file and Hospital Weights file for NIS beginning in 1998. In all data years except 2000, this weight is used to create national estimates for all analyses. In 2000 only, this weight is used to create national estimates for all analyses, excluding those that involve total charges.
(Weights for 1988-1993 are on Hospital Weights file)	*DISCWT_U*	*Discharge weight on Core file and Hospital Weights file for NIS prior to 1998*
	DISCWTcharge	*Discharge weight for national estimates of total charges. In 2000 only, this weight is used to create national estimates for analyses that involve total charges*.
Hospital information		
Hospital identifiers (encrypted)	DSHOSPID	Hospital number as received from the data source
	HOSPID	HCUP hospital number (links to Hospital Weights file)
Hospital location	HOSPST	State postal code for the hospital (e.g., AZ for Arizona)

	HOSPSTCO	*Modified Federal Information Processing Standards (FIPS) State/county code for the hospital links to Area Resource File (available from the Bureau of Health Professions, Health Resources and Services Administration). Beginning in 2003, this data element is available only on the Hospital Weights file*.
Hospital stratifier	NIS_STRATUM	Stratum used to sample hospitals, based on geographic region, control, location/teaching status, and bed size. Stratum information is also contained in the Hospital Weights file.
Other identifiers		
Data source information	*DSNUM*	*Data source number*
	DSTYPE	*Data source type: (1) State data organization, (2) Hospital association, (3) Consortia*
Record identifier, synthetic	KEY	Unique record number for file beginning in 1998
Discharge counts	N_DISC_U	Number of AHA universe discharges in the stratum
	S_DISC_U	Number of sampled discharges in the sampling stratum (NIS_STRATUM or STRATUM)
	S_DISC_S	*Number of sampled discharges in the stratum STRAT_ST*
	N_DISC_F	*Number of frame discharges in the stratum*
	N_DISC_S	*Number of State's discharges in the stratum*
	TOTAL_DISC	Total number of discharges from this hospital in the NIS
	TOTDSCHG	*Total number of discharges from this hospital in the NIS*
Discharge weights	DISCWT	Discharge weight used in the NIS beginning in 1998. In all data years except 2000, this weight is used to create national estimates for all analyses. In 2000 only, this weight is used to create national estimates for all analyses, excluding those that involve total charges.
Discharge Year	YEAR	Discharge year
Hospital counts	*N_HOSP_F*	*Number of frame hospitals in the stratum*
	N_HOSP_S	*Number of State's hospitals in the stratum*
	N_HOSP_U	Number of AHA universe hospitals in the stratum
	S_HOSP_S	*Number of sampled hospitals in STRAT_ST*
	S_HOSP_U	Number of sampled hospitals in the stratum (NIS_STRATUM or STRATUM)
Hospital identifiers	HOSPID	HCUP hospital number (links to Inpatient Core files)

	AHAID	AHA hospital identifier that matches AHA Annual Survey Database (not available for all States)
	IDNUMBER	AHA hospital identifier without the leading 6 (not available for all States)
	HOSPNAME	Hospital name from AHA Annual Survey Database (not available for all States)
Hospital location	HOSPCITY	Hospital city from AHA Annual Survey Database (not available for all States)
	HOSPZIP	Hospital ZIP Code from AHA Annual Survey Database (not available for all States)
Hospital characteristics	HOSP_BEDSIZE	Bed size of hospital (STRATA): (1) small, (2) medium, (3) large
	HOSP_CONTROL	Control/ownership of hospital collapsed (STRATA): (0) government or private, collapsed category, (1) government, nonfederal, public, (2) private, non-profit, voluntary, (3) private, invest-own, (4) private, collapsed category
	H_CONTRL	Control/ownership of hospital: (1) government, nonfederal (2) private, non-profit (3) private, investor-owned
	ST_OWNER	*Control/ownership of hospital: (1) public (2) private, non-profit (3) private for profit*
	HOSP_LOCATION	Location: (0) rural, (1) urban
	H_LOC	*Location: (0) rural, (1) urban*
	HOSP_LOCTEACH	Location/teaching status of hospital (STRATA): (1) rural, (2) urban non-teaching, (3) urban teaching
	HOSP_MHSMEMBER	Multi-hospital system membership: (0) non-member, (1) member
	HOSP_MHSCLUSTER	Multi-hospital system cluster code: (1) centralized health system, (2) centralized physician/insurance health system, (3) moderately centralized health system, (4) decentralized health system, (5) independent hospital system, (6) unassigned
	H_LOCTCH	*Location/teaching status of hospital: (1) rural, (2) urban non-teaching, (3) urban teaching*
	LOCTEACH	*Location/teaching status of hospital: (1) rural, (2) urban non-teaching, (3) urban teaching*
	HOSP_REGION	Region of hospital (STRATA): (1) Northeast, (2) Midwest, (3) South, (4) West
	H_REGION	*Region of hospital: (1) Northeast, (2) Midwest, (3) South, (4) West*
	ST_REG	*Region of hospital: (1) Northeast, (2) Midwest, (3) South, (4) West*
	HOSP_TEACH	Teaching status of hospital: (0) non-teaching, (1) teaching

	H_TCH	*Teaching status of hospital: (0) non-teaching, (1) teaching*
	NIS_STRATUM	Stratum used to sample hospitals beginning in 1998; includes geographic region, control, location/teaching status, and bed size
	STRATUM	*Stratum used to sample hospitals prior to 1998; includes geographic region, control, location/teaching status, and bed size*
	STRAT_ST	*Stratum for State-specific weights*
Hospital weights	HOSPWT	Weight to hospitals in AHA universe (i.e., total U.S.) beginning in 1998
	HOSPWT_U	*Weight to hospitals in AHA universe (i.e., total U.S.) prior to 1998*
	HOSPWT_F	*Weight to hospitals in the sample frame*
	HOSPWT_S	*Weight to hospitals in the State*
Table 3. Data Elements in the NIS Disease Severity Measures Files		
AHRQ Comorbidity Software (AHRQ)	CM_AIDS	AHRQ comorbidity measure: Acquired immune deficiency syndrome: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_ALCOHOL	AHRQ comorbidity measure: Alcohol abuse: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_ANEMDEF	AHRQ comorbidity measure: Deficiency anemias: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_ARTH	AHRQ comorbidity measure: Rheumatoid arthritis/collagen vascular diseases: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_BLDLOSS	AHRQ comorbidity measure: Chronic blood loss anemia: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_CHF	AHRQ comorbidity measure: Congestive heart failure: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_CHRNLUNG	AHRQ comorbidity measure: Chronic pulmonary disease: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_COAG	AHRQ comorbidity measure: Coagulopathy: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_DEPRESS	AHRQ comorbidity measure: Depression: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_DM	AHRQ comorbidity measure: Diabetes, uncomplicated: (0) Comorbidity is not present, (1) Comorbidity is present

	CM_DMCX	AHRQ comorbidity measure: Diabetes with chronic complications: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_DRUG	AHRQ comorbidity measure: Drug abuse: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_HTN_C	AHRQ comorbidity measure: Hypertension, (combine uncomplicated and complicated): (0) Comorbidity is not present, (1) Comorbidity is present
	CM_HYPOTHY	AHRQ comorbidity measure: Hypothyroidism: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_LIVER	AHRQ comorbidity measure: Liver disease: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_LYMPH	AHRQ comorbidity measure: Lymphoma: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_LYTES	AHRQ comorbidity measure: Fluid and electrolyte disorders: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_METS	AHRQ comorbidity measure: Metastatic cancer: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_NEURO	AHRQ comorbidity measure: Other neurological disorders: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_OBESE	AHRQ comorbidity measure: Obesity: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_PARA	AHRQ comorbidity measure: Paralysis: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_PERIVASC	AHRQ comorbidity measure: Peripheral vascular disorders: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_PSYCH	AHRQ comorbidity measure: Psychoses: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_PULMCIRC	AHRQ comorbidity measure: Pulmonary circulation disorders: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_RENLFAIL	AHRQ comorbidity measure: Renal failure: (0) Comorbidity is not present, (1) Comorbidity is present

	CM_TUMOR	AHRQ comorbidity measure: Solid tumor without metastasis: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_ULCER	AHRQ comorbidity measure: Peptic ulcer disease excluding bleeding: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_VALVE	AHRQ comorbidity measure: Valvular disease: (0) Comorbidity is not present, (1) Comorbidity is present
	CM_WGHTLOSS	AHRQ comorbidity measure: Weight loss: (0) Comorbidity is not present, (1) Comorbidity is present
All Patient Refined DRG (3M)	APRDRG	All Patient Refined DRG
	APRDRG_Risk_Mortality	All Patient Refined DRG: Risk of Mortality Subclass: (0) No class specified, (1) Minor likelihood of dying, (2) Moderate likelihood of dying, (3) Major likelihood of dying, (4) Extreme likelihood of dying
	APRDRG_Severity	All Patient Refined DRG: Severity of Illness Subclass: (0) No class specified, (1) Minor loss of function (includes cases with no comorbidity or complications), (2) Moderate loss of function, (3) Major loss of function, (4) Extreme loss of function
All-Payer Severity-adjusted DRG (Optum Insight)	*APSDRG*	*All-Payer Severity-adjusted DRG*
	APSDRG_Mortality_Weight	*All-Payer Severity-adjusted DRG: Mortality Weight*
	APSDRG_LOS_Weight	*All-Payer Severity-adjusted DRG: Length of Stay Weight*
	APSDRG_Charge_Weight	*All-Payer Severity-adjusted DRG: Charge Weight*
Disease Staging (Thomson Reuters)	DS_DX_Category1	Disease Staging: Principal Disease Category
	DS_Stage1	Disease Staging: Stage of Principal Disease Category
	DS_LOS_Level	*Disease Staging: Length of Stay Level: (1) Very low (less than 5% of patients), (2) Low (5 - 25% of patients), (3) Medium (25 - 75% of patients), (4) High (75 - 95% of patients), (5) Very high (greater than 95% of patients)*
	DS_LOS_Scale	*Disease Staging: Length of Stay Scale*
	DS_Mrt_Level	*Disease Staging: Mortality Level: (0) Extremely low - excluded from percentile calculation (mortality probability less than .0001), (1) Very low (less than 5% of patients), (2) Low (5 - 25% of patients), (3) Medium (25 - 75% of patients), (4) High (75 - 95% of patients), (5) Very high (greater than 95% of patients)*

	DS_Mrt_Scale	*Disease Staging: Mortality Scale*
	DS_RD_Level	*Disease Staging: Resource Demand Level: (1) Very low (less than 5% of patients), (2) Low (5 - 25% of patients), (3) Medium (25 - 75% of patients), (4) High (75 - 95% of patients), (5) Very high (greater than 95% of patients)*
	DS_RD_Scale	*Disease Staging: Resource Demand Scale*
Linkage Data Elements	HOSPID	HCUP hospital identification number
	KEY	HCUP record identifier

APPENDIX B: Table 3 List of State organizations partners with HCUP for 2008-210 NIS

State	Data Organization
AK	Alaska State Hospital and Nursing Home Association
AR	Arkansas Department of Health
AZ	Arizona Department of Health Services
CA	Office of Statewide Health Planning & Development
CO	Colorado Hospital Association
CT	Connecticut Hospital Association
FL	Florida Agency for Health Care Administration
GA	Georgia Hospital Association
HI	Hawaii Health Information Corporation
IA	Iowa Hospital Association
IL	Illinois Department of Public Health
IN	Indiana Hospital Association
KS	Kansas Hospital Association
KY	Kentucky Cabinet for Health and Family Services
LA	Louisiana Department of Health and Hospitals
MA	Division of Health Care Finance and Policy
MD	Health Services Cost Review Commission
ME	Maine Health Data Organization
MI	Michigan Health & Hospital Association
MN	Minnesota Hospital Association
MO	Hospital Industry Data Institute
MS	Mississippi Department of Health
MT	MHA - An Association of Montana Health Care Providers
NC	North Carolina Department of Health and Human Services
NE	Nebraska Hospital Association
NJ	New Jersey Department of Health

NM	New Mexico Department of Health
NV	Nevada Department of Health and Human Services
NY	New York State Department of Health
OH	Ohio Hospital Association
OK	Oklahoma State Department of Health
OR	Oregon Association of Hospitals and Health Systems
PA	Pennsylvania Health Care Cost Containment Council
RI	Rhode Island Department of Health
SC	South Carolina State Budget & Control Board
SD	South Dakota Association of Healthcare Organizations
TN	Tennessee Hospital Association
TX	Texas Department of State Health Services
UT	Utah Department of Health
VT	Vermont Association of Hospitals and Health Systems
VA	Virginia Health Information
WA	Washington State Department of Health
WI	Wisconsin Department of Health Services
WV	West Virginia Health Care Authority
WY	Wyoming Hospital Association
*New Hampshire data was not available in time to be included in the 2010 NIS.	

APPENDIX C: Table 4 Comorbidities by Median Household Income

Comorbidities		Median household income national quartile for patient ZIP Code				Total
		$1 - $38,999	$39,000 - $47,999	$48,000 - $62,999	$63,000 or more	
Acquired immune deficiency syndrome	Count	458	240	182	90	970
	% within ZIPINC_QRTL	.4%	.2%	.2%	.1%	
	% of Total	.1%	.1%	.0%	.0%	.3%
Alcohol abuse	Count	6,952	5,542	4,500	3,507	20,501
	% within ZIPINC_QRTL	6.2%	5.4%	5.1%	4.7%	
	% of Total	1.8%	1.5%	1.2%	.9%	5.5%
Deficiency anemias	Count	22,867	19,215	17,193	14,767	74,042
	% within ZIPINC_QRTL	20.4%	18.9%	19.6%	19.8%	

	% of Total	6.1%	5.1%	4.6%	3.9%	19.7%
Rheumatoid arthritis/collagen vascular diseases	Count	3,144	2,978	2,694	2,488	11,304
	% within ZIPINC_QRTL	2.8%	2.9%	3.1%	3.3%	
	% of Total	.8%	.8%	.7%	.7%	3.0%
Chronic blood loss anemia	Count	3,734	3,179	2,884	2,269	12,066
	% within ZIPINC_QRTL	3.3%	3.1%	3.3%	3.0%	
	% of Total	1.0%	.8%	.8%	.6%	3.2%
Congestive heart failure	Count	11,195	9,529	7,762	6,198	34,684
	% within ZIPINC_QRTL	10.0%	9.4%	8.9%	8.3%	
	% of Total	3.0%	2.5%	2.1%	1.6%	9.2%
Chronic pulmonary disease	Count	25,943	22,713	18,577	14,324	81,557
	% within ZIPINC_QRTL	23.1%	22.3%	21.2%	19.2%	
	% of Total	6.9%	6.0%	4.9%	3.8%	21.7%
Coagulopathy	Count	4,811	4,478	4,266	3,933	17,488
	% within ZIPINC_QRTL	4.3%	4.4%	4.9%	5.3%	
	% of Total	1.3%	1.2%	1.1%	1.0%	4.6%
Depression	Count	12,755	12,847	11,102	9,278	45,982
	% within ZIPINC_QRTL	11.4%	12.6%	12.7%	12.4%	
	% of Total	3.4%	3.4%	3.0%	2.5%	12.2%
Diabetes, uncomplicated	Count	28,463	23,735	19,838	14,806	86,842
	% within ZIPINC_QRTL	25.3%	23.3%	22.7%	19.9%	
	% of Total	7.6%	6.3%	5.3%	3.9%	23.1%
Diabetes with chronic complications	Count	6,050	5,181	4,406	3,363	19,000
	% within ZIPINC_QRTL	5.4%	5.1%	5.0%	4.5%	
	% of Total	1.6%	1.4%	1.2%	.9%	5.1%
Drug abuse	Count	6,944	4,436	3,496	2,337	17,213
	% within ZIPINC_QRTL	6.2%	4.4%	4.0%	3.1%	
	% of Total	1.8%	1.2%	.9%	.6%	4.6%
Hypertension (combine uncomplicated and	Count	66,422	58,969	50,913	43,607	219,911
	% within ZIPINC_QRTL	59.1%	58.0%	58.2%	58.5%	
	% of Total	17.7%	15.7%	13.5%	11.6%	58.5%

complicated)						
Hypothyroidism	Count	11,969	12,365	11,468	10,529	46,331
	% within ZIPINC_QRTL	10.7%	12.2%	13.1%	14.1%	
	% of Total	3.2%	3.3%	3.0%	2.8%	12.3%
Liver disease	Count	3,621	3,089	2,735	2,204	11,649
	% within ZIPINC_QRTL	3.2%	3.0%	3.1%	3.0%	
	% of Total	1.0%	.8%	.7%	.6%	3.1%
Lymphoma	Count	860	874	846	875	3,455
	% within ZIPINC_QRTL	.8%	.9%	1.0%	1.2%	
	% of Total	.2%	.2%	.2%	.2%	.9%
Fluid and electrolyte disorders	Count	27,893	24,817	21,927	18,454	93,091
	% within ZIPINC_QRTL	24.8%	24.4%	25.1%	24.7%	
	% of Total	7.4%	6.6%	5.8%	4.9%	24.8%
Metastatic cancer	Count	2,571	2,682	2,552	2,487	10,292
	% within ZIPINC_QRTL	2.3%	2.6%	2.9%	3.3%	
	% of Total	.7%	.7%	.7%	.7%	2.7%
Other neurological disorders	Count	9,837	8,443	7,317	6,208	31,805
	% within ZIPINC_QRTL	8.8%	8.3%	8.4%	8.3%	
	% of Total	2.6%	2.2%	1.9%	1.7%	8.5%
Obesity	Count	13,879	12,588	10,796	7,858	45,121
	% within ZIPINC_QRTL	12.4%	12.4%	12.3%	10.5%	
	% of Total	3.7%	3.3%	2.9%	2.1%	12.0%
Paralysis	Count	3,396	2,798	2,629	2,165	10,988
	% within ZIPINC_QRTL	3.0%	2.8%	3.0%	2.9%	
	% of Total	.9%	.7%	.7%	.6%	2.9%
Peripheral vascular disorders	Count	7,349	6,974	5,750	4,803	24,876
	% within ZIPINC_QRTL	6.5%	6.9%	6.6%	6.4%	
	% of Total	2.0%	1.9%	1.5%	1.3%	6.6%
Psychoses	Count	6,607	5,246	4,542	3,405	19,800
	% within ZIPINC_QRTL	5.9%	5.2%	5.2%	4.6%	
	% of Total	1.8%	1.4%	1.2%	.9%	5.3%
Pulmonary	Count	2,122	2,009	1,957	1,696	7,784

circulation disorders	% within ZIPINC_QRTL	1.9%	2.0%	2.2%	2.3%	
	% of Total	.6%	.5%	.5%	.5%	2.1%
Renal failure	Count	14,788	12,380	10,609	8,476	46,253
	% within ZIPINC_QRTL	13.2%	12.2%	12.1%	11.4%	
	% of Total	3.9%	3.3%	2.8%	2.3%	12.3%
Solid tumor without metastasis	Count	2,499	2,326	2,085	1,975	8,885
	% within ZIPINC_QRTL	2.2%	2.3%	2.4%	2.6%	
	% of Total	.7%	.6%	.6%	.5%	2.4%
Peptic ulcer disease excluding bleeding	Count	58	35	43	28	164
	% within ZIPINC_QRTL	.1%	.0%	.0%	.0%	
	% of Total	.0%	.0%	.0%	.0%	.0%
Valvular disease	Count	3,537	3,630	3,596	3,640	14,403
	% within ZIPINC_QRTL	3.1%	3.6%	4.1%	4.9%	
	% of Total	.9%	1.0%	1.0%	1.0%	3.8%
Weight loss	Count	5,666	4,790	3,945	3,058	17,459
	% within ZIPINC_QRTL	5.0%	4.7%	4.5%	4.1%	
	% of Total	1.5%	1.3%	1.0%	.8%	4.6%
Total	Count	112,325	101,696	87,503	74,572	376,096
	% of Total	29.9%	27.0%	23.3%	19.8%	100.0%

Percentages and totals are based on respondents.

a. Dichotomy group tabulated at value 1.

APPENDIX D: IRB Approval

RUTGERS
THE STATE UNIVERSITY OF NEW JERSEY

Newark Campus

School of Health Related Professions

Department of Health Informatics

Project Title: SOCIOECONOMICS STATUS AND HOSPITALIZATION CHARACTERISTICS IN THE UNITED STATES: A RETROSPECTIVE STUDY

Degree for Doctor of Philosophy in Biomedical Informatics

Principal Investigator: Adiebonye Eunice Jumbo

Co-Investigator/Supervisor: Dr. Shankar Srinivasan

Funding Source(s): None

Date: November 24, 2014

Version: #1

RUTGERS APPROVED
THE STATE UNIVERSITY OF NEW JERSEY
IRB ID: Pro20140001035
Approval Date: 12/11/2014
Expiration Date: