A Study on Racial, Ethnic and Socioeconomic Disparities with In-Hospital Mortality and Length of Stay of Patients with Congestive Heart Failure

Chuanwei Wang

Graduate School of Public Health

Yonsei University

Department of Global Health

A Study on Racial, Ethnic and Socioeconomic Disparities with In-Hospital Mortality and Length of Stay of Patients with Congestive Heart Failure

A Master Thesis
Submitted to the Department of Global Health
and the Graduate School of Yonsei University
in partial fulfillment of the
requirements for the degree of
Master of Public Health

Chuanwei Wang

January 2013

This certifies that the Dissertation of
Wang Chuanwei is approved.

_____

Thesis Supervisor: Tae Hyun Kim

_____

Thesis Committee Member : So Hee Park

_____

Thesis Committee Member : Hee Jin Kim

Graduate School of Public Health

Yonsei University

Department of Global Health

Acknowledgements:

I'd like to start off by thanking Dr. Kim Tae Hyun for guiding me through this rewarding process of writing this dissertation. Without his help, I would have been forever lost in the complicated technical jargon and specifics of research. I'd like to thank the entirety of Yonsei University's Graduate School of Public Health, specifically Dr. Myongsei Sohn , Dr Young Moon Chae and the entire secretary staff from 2011-2013 for their help in guiding me through the process of admissions, scholarships and vital work experience that will prepare me for the future, without their help, none of this would've ever been possible. I'd like to thank the rest of my dissertation review committee of Dr. Park Sohee, and Dr. Kimm Hee Jin for their continued input in my dissertation. Having an outside look really taught me how true professionals approached thesis writing. I'd like to thank Shin Bo Kyung, Elliot Kwong, Rebecca Cho, and Pwo Heejin for being great classmates and friends that I know I could turn to in any situation, school or not. I am eternally grateful to Steven Huh and Kim Hee Jung to give me the means to live in Korea. I want to say thank you to my family, including my father and mother Dr Jizhong Wang, and Lingyun Gong for giving me the support to finish a wild dream of mine. Finally, I want to say a special thank you to Oh Eunji, my girlfriend that's been with me from the start. Without your support, I would never have been able to reach the finish line. I love you.

Table of Contents

List of Figures

List of Tables

Abstract:

Each year, an estimated 17 million people worldwide die from cardiovascular disease. As one of the leading killers in the United States (1 in 4 die from this disease), it's shown that over $109 billion dollars have been spent on healthcare costs alone. (WHO) This, combined with the fact that over half of the American population is at risk of cardiovascular disease is an important reason why the study of this disease is so vital. (CDC) This dissertation uses the Healthcare Cost and Utilization Project (HCUP) to show the growing trends of treatment among healthcare providers with regards to factors like race, location of hospital, gender among others. We will show that while there are many non-medical factors that may influence the spread of this disease, there are also factors that may be controlled by the hospital itself in order to properly predict possible mortality among cardiovascular disease patients. What we have found is that there is are many non-medical factors that can increase and reduce the risk of in-hospital mortality. However, we have found that creating a viable model for length of stay with different independent variables was not as successful as the R^2 of the simple linear regression maxed out at 0.103. What this means is that our model does not fit perfectly. It can imply that some may be variables that don't work well with each other, and other underlying reasons that may cause have caused the model to not fit perfectly. Additional research will have to be done to see how to make a model fit with simple linear regression of length of stay.

Chapter 1: Introduction

1.1 Cardiovascular Disease

Cardiovascular disease is a disease that affects the heart and it's related blood vessels due to build up of plague within. This can lead to heart attacks, and congestive heart failure and ultimately, death. Over 17 million people every year are affected by this disease due to a variety of factors that are both medical, and non medical. (WHO, 2013) Also, as cardiovascular disease has symptoms that are easily overlooked, this can lead to people and doctors to misdiagnose it. (Waterloo, 2013) As such, over $109 billion dollars are spent on the treatment of this disease alone. (CDC, 2013) Cardiovascular disease most commonly takes the form of atherosclerosis, a build up of plague in the arteries. As the buildup of the plague grows, blood increasing meets higher resistance and causes blockage of blood within the body. This can lead to heart attacks, strokes and other litany of problems like congestive heart failure. (American Heart Clinic, 2013)

1.2 Congestive Heart Failure (ICD-9 Code 428.0)

Congestive Heart Failure is a type of cardiovascular disease that is caused by when the heart cannot pump blood, as it normally should. (Mayo Clinic, 2013) The danger of it is that it can develop suddenly and cause problems in areas of the body where it may not seem like that's the cause. Also, congestive heart failure often leads to the heart growing weaker and weaker which can lead to more severe problems like frequent heart attacks and ultimately death down the road. Common symptoms include coughing, fatigue, weight gain, shortness of breath and swollen appendages among others. (MedlinePlus, 2013) Because these symptoms are so common and often are related to other diseases, congestive heart failure is very easily misdiagnosed as something else, which can cost vital time that the patient might not have. Cardiovascular diseases as a whole can take a lot to recover from, often requiring both medication, and changes in life style. This culminates into a situation where the patient very likely has to stay within the hospital to recover, and where the doctor can do further checkups to make sure the patient is fine. Studying both the medical and the non-medical aspects of this disease is important as we can find relationships between variables about a patient even before they come in.

1.3 Healthcare Cost and Utilization Project (HCUP)

The Healthcare Cost and Utilization project, or HCUP is the largest longitudinal collection of data on nationwide hospitals. These datasets are published once a year and are released with information on patients, participating hospitals and disease severity levels. This data is one of the most complete tools to show a more complete relationship between different policies of different states, and a broad view on what separates hospitals from each other. As this data encompasses most of the United States (there are some exceptions in which some years a few states decide to opt out of this project, but it differs year from year), it is relatively not burdensome to compare not just different hospitals, but different regions with each other as well. An important thing to note is that this data cannot be merged together through different years, as there is a difference in what states decide to partake in this program. (HCUP, 2013) Because of this data is a national data view; it can be used to do research on a number of Public Health, Health Policy, and Hospital Administration views. Using this data to develop long-term health policies and health education among the general populace and at-risk populations means less waste in spending of valuable money, and to inform people on how to lead a healthier and more balanced lifestyle.

1.4 Diagnosis-Related Group

Diagnosis-Related Group, or DRG is a way for hospitals to see how much resources patients will need based on different diseases that they may have. (DRG, 2003) DRG can also be used to determine payment for hospitals, which is dependent on the severity of disease and the resources used by the hospital. There are also a few different types of DRG, including All-Payer Severity-adjusted DRG. This is simply a DRG that is adjusted for severity, and findings have found that it explains more than other DRG measurements. (Wynn & Scott, 2007). The main focus for DRG's used though is on how much capital will the hospital expect a patient to use and thus how much to charge.

1.5 Disease Staging Level

Hospitals often use stages to represent the level that a disease is at. While DRG is most often used to represent the weight of a patient or disease has on the hospital, disease staging level is focused on the patient. (4) It is defined as

Stage 1: minimal severity

Stage 2: problems in one organ or system

Stage 3: multiple organ or system disease or failure

Stage 4: Death

1.6 Objective

We want to find the relationship between different factors such as race and ethnicity, median income level, payer information and in-hospital mortality that may not have a direct relationship with the patient's physical health. We believe that we will find a relationship, and a model that can accurately predict dependent variables like length of stay.

Chapter 2: Literature Review

2.1 Disparities in Race/Ethnicity and Gender in In-Hospital Mortality rates for Coronary Artery Bypass Surgery Patients (Becker & Rahimi, 2006)

In hospital mortality rates is something that is often looked at to determine how effective a particular treatment; characteristic or surgery is to a large group of patients. Specifically in this case, coronary artery bypass surgery patients are studied to show that mortality rates between patients of different gender and race are different.

Coronary artery bypass surgery is a surgery that is designed to improve the blood flow to the heart by grafting new blood vessels like arteries and veins from other parts of the body to the heart in order to provide improved blood movement throughout the body. This procedure is often referred to as a single, double, triple, quadruple, or quintuple bypass surgery. The number of bypass is defined as the number of coronary arteries that are bypassed because of the surgery. (National Heart, Lung and Blood Institute, 2012)

Despite the fact that the number of patients that need this surgery is going up, “national data from the Healthcare Cost and Utilization Project (HCUP) reports that an in-hospital mortality decline from 3.33% to 2.30%, a 30.9% decline over a 10 year period.” However the in-hospital mortality decline has not been standard across the board across all ethnicities and genders. Using a five year sample of the HCUP data, and multivariate analysis – logistic regression, among the ethnicities that were seen within the data (white, black, Hispanic and Asian/pacific islander), the in hospital mortality rates were much higher for black and Hispanic compared to Asian/pacific islander and whites. Other variables like comorbidities and risk factors were also tested to see a relationship between mortality rates after surgery. What this shows is that there is a clear link between mortality rates of patient and medical factors like smoking, obesity and ethnicity. It is noteworthy to point out while this study focused primarily on medical factors; there is the issue of the non-medical factors to take into consideration.

Taking these into account, we see that cardiovascular diseases and its subsequent treatment options and recovery are governed by more than just a simple one lined diagnosis. Treatments and recovery could differ based on simple things as race, ethnicity, gender, and outside habits like smoking, drugs and alcohol abuse. Hospitals in at-risk areas can also start preventive measures to reduce in-hospital mortality.

2.2 Health Care Disparity In the Care of the Vascular Patient (Kirksey, 2011)

Vascular disease is a disease of the blood vessels that can cause blockage of blood from the heart to the rest of the body. (Medline, 2013) While similar to cardiovascular disease, the difference is that vascular diseases focus more on the blood vessels, while cardiovascular diseases focus more on the entire system of the heart and blood vessels. But while in-hospital mortality rate of patients has gone down, we see that it's possible to have non-medical factors influencing the survival rate. For example, "1 in 5 American households speak a primary language other than English at home. The communication gap presents an obstacle to the efficient delivery of medical care when family members seek health care." (Kirksey, 2011) As such, problems that stem from a culture background can lead to true medical problems. Specifically

1. "Increased medical errors jeopardizing patient safety;

2. Increased health expenditures due to inefficiently delivered health care;

3. Worse outcomes for the management of chronic medical conditions and procedural interventions. " (Kirksey, 2011)

The problem that stems from a lack or difficulty communicating from healthcare professionals to patients is one that can be avoided if properly prepared for. However if the doctor or nurse cannot communicate with the patient adequately, medication could be taken wrongly, and lifestyle changes that need to happen, might never happen. However there is resistance within the medical community for this to happen. As doctors feel, "compassion is at the core of personality traits that characterize health care providers. As an extension of this premise, some suggest that the possession of compassion insulates physicians and other health care providers from many of the potential pitfalls of America's growing cultural diversity." (Kirksey, 2011) What must be done from both a medical and public health standpoint is to promote culture diversity within the medical community in so that at-risk and minority populations receive the same chances as those who's primary language is

English and primary culture is “American”. The possible benefits to this change is

1. “Reducing the overall health care costs which are increased by poor compliance, medical errors;

2. Improving patient outcomes for specific chronic disease states like vascular disease.” (Kirksey, 2011)

2.3 Racial and ethnic disparities in cardiovascular medication use among older adults in the United States (Qato, Lindau, Conti, Schumm, Alexander &, 2010)

Cardiovascular disease among the elderly is one of the leading causes of death in America. In a research done by Dr. Barnhart, it was found that that there were also non-medical factors that could relate to disparities among cardiovascular disease. Not only that, cardiovascular diseases is the leading cause of death among patients of hospitalization over 60 years of age, and often can lead to death while in a fragile state. (Barnhart, Cohen, Wright & Wylie-Rosett, 2006) Also, the prevalence of prescription and over the counter medication has lead Americans to treatments at home, which can further cause disparities in cardiovascular health.

As stated above, cardiovascular diseases are hard to diagnose and often result from more than a single factor, which can cause these disparities. Using a sample of 4400 “community dwelling person, 57 to 84 years of age” and were white (non-Hispanic), black (non-Hispanic) and Hispanic (not-black), it was found that 35% (or 1066 people) were at a high cardiovascular risk, 35% (or 977 people) were at moderate cardiovascular risk and 30% (or 812 people) were classified as low risk. It was also found that older adults were more likely to use “preventive therapies than those at moderate or low risk; nearly half (48%) regularly used a statin and 41% regularly used aspirin.” (Qato, Lindau, Conti, Schumm, Alexander &, 2010) What’s surprising is that the report says that there is a “substantial underuse of both statins and aspirin among older adults despite growing evidence of primary and secondary benefits”. (Qato, Lindau, Conti, Schumm, Alexander &, 2010) Also older African Americans and Hispanic tend to use less preventive measures. It was said that this could be related to the location of pharmacies from lower socioeconomic areas.

What we have seen is that people of different ethnicities, race and socioeconomic background behave in different ways that could affect the prevalence of cardiovascular diseases among the elderly. This effect could explain the higher mortality rates of African American’s and Hispanics compared to White’s and Asian/Pacific Islanders. Moreover, the underlying issues of different culture,

socioeconomics and location are factors that should be looked at.

2.4. Racial and Ethnic Disparities in Mortality Following Congenital Heart Surgery (Benavidez & Jenkins, 2006)

Congenital Heart surgery is a type of surgery that is given to children born with some type of heart defect, more commonly called a congenital heart disease. (Congenital heart defect - corrective surgery) The objective of this study was to determine racial and ethnic disparities after these surgeries. Similar to our study, their dataset was obtained from HCUP KID, a subset of the HCUP that is specially designed to find a random sample of children patients. The study was then divided into six different racial groups, white, black, Hispanic, Asian, other and missing. The data was also divided into categories like premature, risk category, geographic location, median zip code income, and insurance type. What was found is that there was indeed racial and ethnic disparities in mortality following congenital heart surgery, however it was not clearly defined as what was causing it. The researchers found that there was consistent reports that black and Hispanic children were dying much more often than any other racial group, but they found it wasn't consistent per state. For example, they said "blacks had a higher risk of dying in Massachusetts and New York but a lower risk of death in Pennsylvania. Hispanics had similar findings, with the higher death rates in New York and a risk of death in Pennsylvania not significantly different from Whites." (Benavidez & Jenkins, 2006) What's interesting about this finding is that it seems to point away from a biological reasoning on the disparities in mortality after surgery.

2.5 Racial/Ethnic and Socioeconomic Differences in Multiple Risk Factors for Heart Disease and Stroke in Women: Behavioral Risk Factor Surveillance System, 2003 (Hayes, Denny, Keenan, Croft, Sundaram & Greenlund, 2006)

Many of the studies that have been done on heart diseases are done on large populations, usually without an expressed interest among the differences between genders. What this study has found is that among women, there are very important risk factors that can be linked with heart disease and stroke. Controlling for factors such as age, income, education, and health coverage, it was found that women that were black, had a higher odds for risk factors than their white counterparts. It was also seen that as age went up, medical conditions such as high blood pressure and diabetes were seen more, while in younger women, risk factors such as smoking were also seen. (Hayes, Denny, Keenan, Croft, Sundaram & Greenlund, 2006) This shows that risk factors that are often characterized with heart disease can possibly explain why the prevalence of heart disease is so high among the black or Hispanic population.

2.6 The relationship between severity of illness and hospital length of stay and mortality (Horn, Sharkey, Buckle, Backofen, Averill & Horn, 1991)

One of the ways for hospitals to determine pricing is by a method called Diagnosis-related group, or DRG. (DRG, 2003) This is to help the hospital have an idea about the quality of care needed for patients, while adjusting for payment that the hospital will later charge. The study asked if it was a relationship between mortality and length of stay with a severity adjusted variable, called Computerized Severity Index, or CSI. One of the biggest reasons for this study was an attempt to discover if the DRG should be changed in order to adjust for severity. In doing so, it would easier to compare hospitals of different regions and patients, as well as provide the hospital with a means to determine the financial burden of high severity patients. Using data from three university teaching hospitals, and two community teaching hospitals, they had found that for length of stay, with CSI and DRG, they found that the levels of variation are much smaller for patients that had lower levels of severity. However, for the incredibly ill, or dead patients, the variance grew much larger. For mortality, they found a strong relationship between severity and death. What could disrupt this data is that there were only a few hospitals that were used, and that they were all teaching hospitals. This means that there is more variability within those hospitals, and thus, the actual influence of severity adjustment could be conservative if the study was redone with a larger set of hospitals. (Horn, Sharkey, Buckle, Backofen, Averill & Horn, 1991)

2.7 Summary

The review of previous literature tells us a few things.

1) Ethnicity among adults has played a large role in determining in-hospital mortality among patients with cardiovascular disease. We saw in the literature that among the ethnicities tested, white, black, Hispanic, Asian/pacific islander, that mortality rates were seen as higher for black and Hispanic people compared to white and Asian/pacific islanders. We think that this is a very important distinction that's made and that it's one of the most common stated differences among patients with cardiovascular disease.

2) Healthcare disparity is a real thing and can affect mortality. While this may seem like a very obvious statement, we see that healthcare disparity doesn't simply just mean medical issues, but also non-medical issues like language. As we wanted to study a wide group of population in the United States, we need to take this into account as the regions of where people come from could play a very big difference in how people are treated.

3) Age is also a risk factor of cardiovascular disease. What we noticed in the literature review is that people of elderly age were at risk for cardiovascular disease and was one of the leading killers in that age bracket. Furthermore, it was more likely that patients would not get adequate treatment, either due to the

over-prevalence of prescription drugs, or simply a lifestyle choice through the many years.

4) There is a similar ethnic disparity trend in children. We saw in a literature review that showed that children that got cardiovascular disease tend to be from black and Hispanic families. However, this was countered by that between different states, it was not a consistent observation that held. What this could mean is that there is simply many more factors affect cardiovascular disease than simply just ethnicity.

5) Women also have seen similar disparities with heart disease. This literature told us that in women, we see that the trends that we had expected, that blacks would have higher incidence of heart disease, held true.

6) Severity adjusted data was a good predictor of mortality and length of stay. What the literature told us was that among hospitals, severity adjusted data would be able to predict mortality and length of stay better than non-severity adjusted data. This is important for the hospital because they can predict the type of care that they can give for the patients, and prices that they set.

Chapter 3: Methodology

3.1 Participants

The participants of the study are from the Healthcare Cost and Utilization Project (HCUP) in 2003. The Nationwide Inpatient Sample (NIS) is part of the HCUP, and is a database of hospital inpatient stays. These datasets are published once a year and are released with information on patients, participating hospitals and disease severity levels. This data is one of the most complete tools to show a more complete relationship between different policies of different states, and a broad view on what separates hospitals from each other. Because of this data is a national data view; it can be used to do research on a number of Public Health, Health Policy, and Hospital Administration views. As this data encompasses most of the United States (there are some exceptions in which some years a few states decide to opt out of this project, but it differs year from year), it is relatively not burdensome to compare not just different hospitals, but different regions with each other as well. An important thing to note is that this data cannot be merged together through different years, as there is a difference in what states decide to partake in this program. There is three different datasets that were used, Inpatient, Hospital and Severity.

Inpatient dataset: The inpatient participants are numbered around 5 million to 8 million patients total, and 442,673 patients with cardiovascular disease. Further categories like gender (male/female),

race (white/black/Hispanic/Asian+pacific islander/Native American/other), length of stay, admission source (ER/another hospital/another facility including long term care/court/law enforcement/routine) and age.

Hospital dataset: The NIS data is supposed to be a 20% of representation of U.S hospitals, all non-federal, short-term, general, public and teaching and other specialty hospitals. There are also some community hospitals, which are specialists that include OBGYN, ear-nose-throat, short-term rehabilitation, orthopedic, and pediatric institutions. The hospital dataset includes basic information like hospital location, zip code, bed size, control of hospital (government/private/public), location (rural/urban), region (northeast/Midwest/South/West) and if the hospital is a teaching hospital or not.

Severity dataset: The severity dataset is a dataset that lists different diseases like depression, diabetes uncomplicated/chronic complications, liver disease, congestive heart failure and peripheral vascular disorders among others.

3.2 Methodology

The goal of this study is to find the relationship between in hospital mortality of congestive heart failure patients and other factors like race, location, bed size of hospital, length of stay, age, gender and income level. . Using SPSS version 22 and the data from HCUP 2003 dataset, We used Bivariate Analysis to generate data in order to see if there was any empirical relationship between different factors. This is to determine to see if simple causality. We also used statistical analysis like logistic regression to determine the odds ratio of different categories and in-hospital mortality. This is to determine if someone was more likely to die because of certain conditions. The conditions that we're most interested in are the ones that are non-medical in nature and more related to the management of hospitals. This means we want to look at factors such as if a teaching status of a hospital makes a difference in the mortality of congestive heart failure patients. We used a bivariate analysis to generate data in order to see if there was any empirical relationship between different factors. This is to determine to see if simple causality.

3.3 Statistical Analysis

Logistic regression is a type of analysis that is used when we want to predict the dependent variable given a certain independent variable. (Joensson & Bhandari, 2008) In this case, we will be using in-hospital mortality, or death, as the dependent variable. This is used because we want to know the in-hospital mortality rates and how independent factors like payer information, or location of hospital will affect it.

Linear regression is a type of analysis when we want to see how different variables can fit to a linear model of increase given a dependent variable with multiple independent variables. We will be using length of stay as the dependent variable, and a variety of independent variables such as disease severity, patient location and elective vs. non-elective admission. In doing so, we will obtain a model with the independent variables and see how well it will fit. We also took the natural log of the length of stay in order to create a more normal distribution for the data.

Odds ratio is a measure that tries to explain the association between two different values, the probability of something happening over another. (Szumilas, 2010) However, there are many risks to using odds ratio, as it can be very easily misunderstood. In an article by Dr. Davies, it's said that odds ratio can be easily misinterpreted as relative risk, which can be different in certain

situations. “If the odds ratio is interpreted as a relative risk it will always overstate any effect size: the odds ratio is smaller than the relative risk for odds ratios of less than one, and bigger than the relative risk for odds ratios of greater than one”. (Davies & Crombie, 1998)

The dataset from HCUP was received in a raw form and an SPSS version 22 run-program was written to see the database properly. We then filtered out all the patients that were not diagnosed with congestive heart failure and divided the dataset into three distinct groups; male, female and total. Furthermore, every category that was tested was re-done into categorical type, rather than numerical type. As a category is most often given a number to indicate a certain status (example: 0 for alive, 1 for death), SPSS originally takes it in as a constant numerical value. To get SPSS to see that it was only discrete values as 0 and 1, we created dummy variables that represented the answers for every category that we test.

Chapter 4: Results

4.1 Table 1: Descriptive Statistics

Descriptive Statistics				
Variables	**Male (N)**	**Male (%)**	**Female (N)**	**Female (%)**
Death				
Died	17437	9.5	20649	8.1
Did Not Die	166725	90.5	234660	91.9
Payer Information				
Medicare	146927	79.3	213866	83.2
Medicaid	9877	5.3	16290	6.3
Private/HMO	22539	12.2	22146	8.6
Self-pay	2481	1.3	2263	.9
No-charge	350	.2	256	.1
Other	3058	1.7	2204	.9
Location of Patient				
Large Metropolitan	95644	51.8	133925	52.3
Small Metropolitan	50052	27.1	68921	26.9
Micropolitan	21398	11.6	28515	11.1
Non-core	17401	9.4	24939	9.7
Ethnicity				
White	98877	53.3	133724	73.3
Black	17299	9.3	27345	15.0
Hispanic	11409	6.2	14443	7.9
Asian/Pacific Islander	2681	1.4	3008	1.6
Native American	269	.1	317	.2
Other	2724	1.5	3632	2.0
Median Income Level				
$1-$35,999	51452	27.7	74932	29.7
$36.000-$44,999	50163	27.1	69547	27.6
$45,000-$59,999	44607	24.1	60852	24.1
$60,000+	34865	18.8	46703	18.5
Elective vs Non-elective				
Non-elective	153675	82.9	215244	84.1
Elective	30917	16.7	40704	15.9

We see that in terms of the difference between males and females, there was not much difference between in-hospital mortality. There is a difference between payer information, and it shows that more females are on programs like Medicare, while males pay more with Private/HMO information. As expected, the distribution of location of the patients is relatively similar. However it is interesting to note that there is a very big difference between the numbers of white

males vs. white females, with more white females going to the hospital.

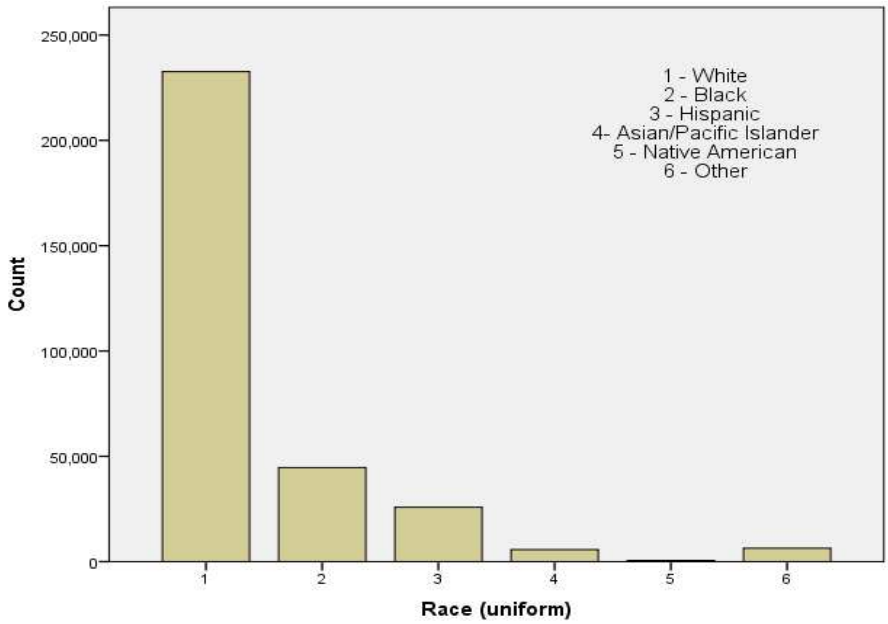
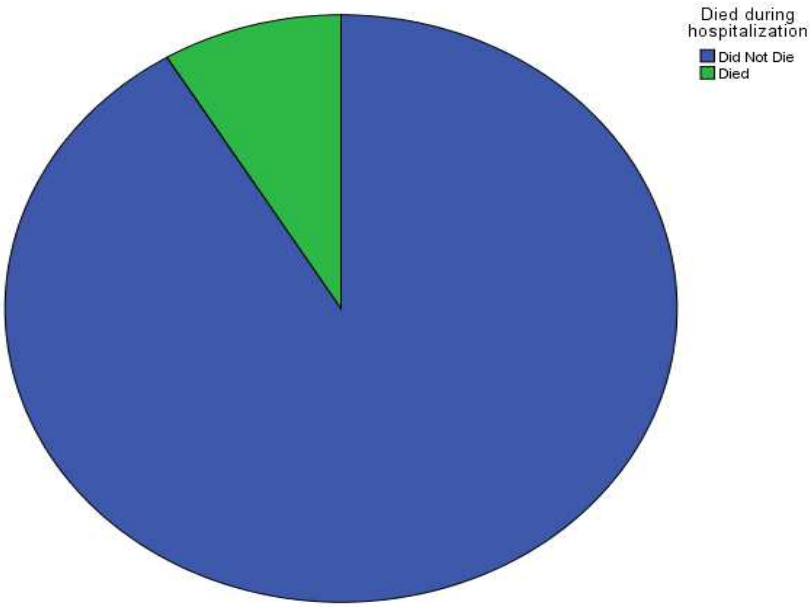


Figure 1: Race vs. population graph

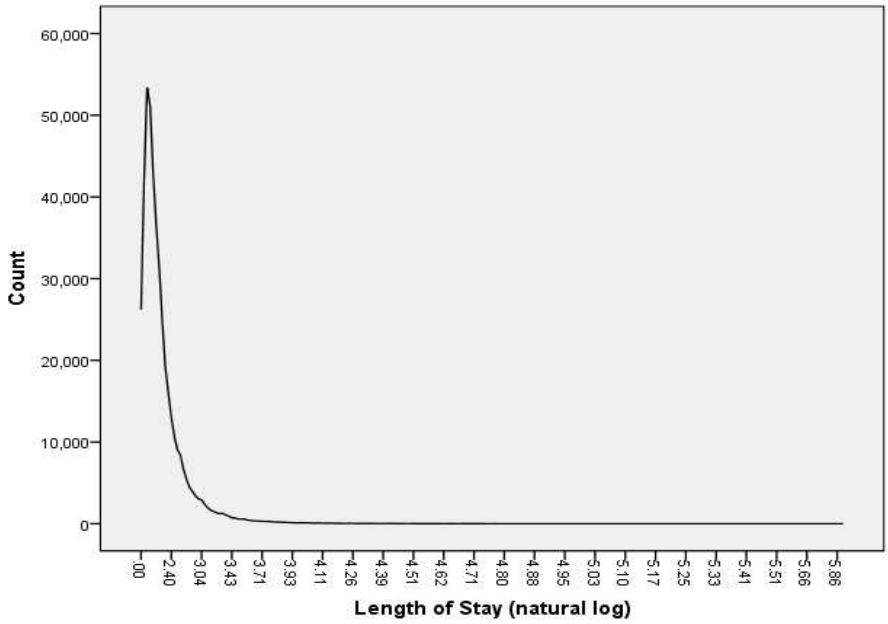
When looking at the overall data, we see that the vast majority of patients are white, with black and Hispanic following behind, and Asian/Pacific Islanders, Native American's being the last.

Figure 2: In-Hospital Mortality pie chart



Similar to what we expect as well, we see that the vast majority of the patients did not die due to congestive heart failure. This is in line to the literature review that we did where we saw that the majority of patients did not die from cardiovascular disease in the hospital.

Figure 3: Natural Log of Length of Stay graph



We see that in length of stay, that the majority of the patients stayed a relatively short amount of time, even length of stay having a minimum of 0 days, and maximum of 358 days. The median was 6 days, and the mean was 7.95 days.

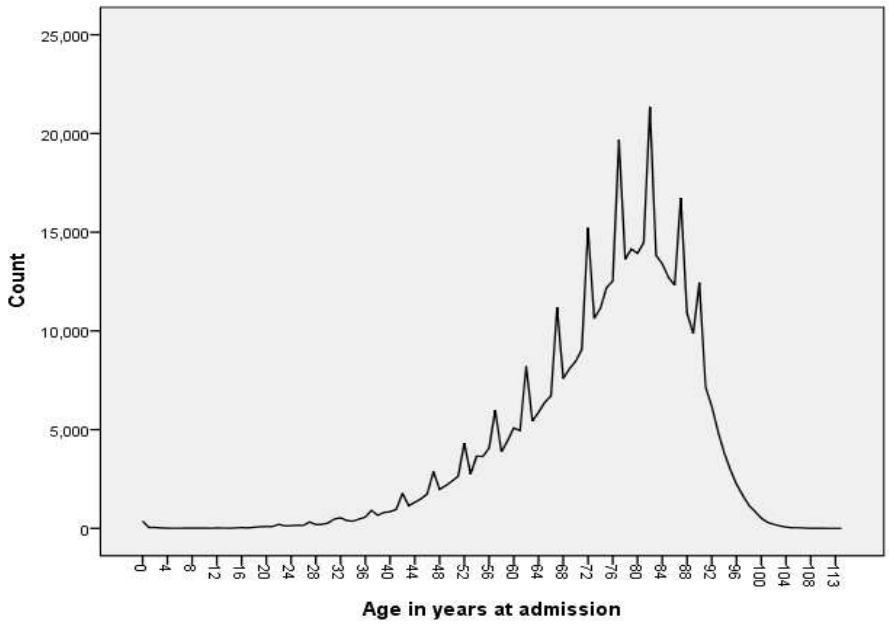


Figure 4: Age vs. population graph

What we can see here is that the number of patients goes up as age goes up as well. There were a few outliers at 0 and 113 years of age.

Logistic Regresion Variables	With Male OR | 95% CI	Death Female OR | 95% CI	As Total OR | 95% CI	Dependent	Variable
Ethnicity					**P Value**
White	1.00	1.00	1.00		* = <0.0001
					** = <0.001
Black	1.30	1.22	1.26		*** = <0.01
	1.22-1.38*	1.16-1.38*	1.22-1.31*		**** = >0.01
Hispanic	1.12	1.03	1.07		
	1.05-1.20**	0.97-1.20**	1.02-1.12***		
Asian/Pacific Islander	0.72	0.73	0.72		
	0.64-0.80*	0.65-0.80*	0.66-0.76*		
Native American	0.87	1.21	1.01		
	0.60-1.27****	0.79-1.27****	0.76-1.34****		
Other	1.09	1.02	1.05		
	0.96-1.24****	0.91-1.24****	0.96-1.24****		
Patient Location					
NorthEast	1.00	1.00	1.00		
Midwest	1.04	1.07	1.05		
	1.01-1.08***	1.03-1.10*	1.03-1.08*		
South	1.15	1.13	1.14		
	1.09-1.22*	1.08-1.19*	1.10-1.18*		
West	1.24	1.24	1.24		
	1.17-1.32*	1.18-1.32*	1.19-1.29*		

Table 2: Logistic Regression – Ethnicity and Patient Location

With	Death Male OR | 95% CI	As Female OR | 95% CI	Dependent Total OR | 95% CI	Variable
Disease Severity Level				**P Value**
Extremely Low (<25%)	1.00	1.00	1.00	* = <0.0001 ** = <0.001
Medium (25%-75%)	2.34 1.03-2.7*	2.71 2.34-3.13*	2.56 2.34-3.13*	*** = <0.01 **** = >0.01
High (75%-95%)	11.00 9.26-13.02*	14.28 12.33-16.39*	12.82 12.33-16.39*	
Very High (95%+)	50.42 42.55-59.88*	66.67 56.49-75.19*	58.82 56.49-75.19*	
Median Income Level				
$1-$35,999	1.00	1.00	1.00	
$36.000-$44,999	0.97 0.93-1.02****	0.99 0.95-1.02****	0.98 0.95-1.01*	
$45,000-$59,999	0.95 0.91-0.99****	0.93 0.90-0.97**	0.94 0.91-0.97*	
$60,000+	0.87 0.83-0.91*	0.86 0.82-0.89*	0.86 0.83-0.88*	

Table 3: Logistic Regression – Disease Severity and Median Income Level

With	Death Male OR | 95% CI	As Female OR | 95% CI	Dependent Total OR | 95% CI	Variable P Value
Payer Information				* = <0.0001
Medicare	1.00	1.00	1.00	** = <0.001
Medicaid	1.38 1.28-1.49*	1.53 1.43-1.64*	1.47 1.40-1.55*	*** = <0.01
Private/HMO	1.20 1.14-1.26*	1.14 1.08-1.20*	1.15 1.11-1.19*	**** = >0.01
Self-pay	1.39 1.19-1.62*	1.16 0.99-1.36***	1.25 1.12-1.40*	
No-charge	3.36 1.84-6.13*	2.48 1.28-4.83**	2.87 1.84-4.48*	
Other	0.91 0.81-1.03****	0.65 0.58-0.74*	0.77 0.71-0.84*	
Location of Patient				
Large Metropolitan	1.00	1.00	1.00	
Small Metropolitan	1.04 1.01-1.08*	1.07 1.03-1.10*	1.06 1.03-1.08*	
Micropolitan	1.15 1.10-1.22*	1.13 1.08-1.19*	1.14 1.10-1.18*	
Non-core	1.24 1.17-1.32*	1.24 1.18-1.31*	1.24 1.19-1.29*	

Table 4: Logistic Regression – Payer Information and Location of Patient

4.2 Logistic Regression

Within ethnicity, using white as the reference, we find that the odds to stay **alive**, compared to white, is greatest as an Asian/Pacific Islander, and least as Black. We see this general trend continue among females. However we should notice that among Native American females, the odds ratio drops dramatically.

During Patient Location, using the Northeast as the reference, we see that the Midwest and Northeast have similar levels of in-hospital mortality, while in the South and Western regions of the United States, it quickly drops off and in-hospital mortality goes **up**. This general trend is seen in women as well.

When we take a look at disease severity level, using an extremely low disease severity level as our reference, it makes sense that as diseases severity goes up, the odds of **dying** in the hospital also go up as well. This makes sense as a higher level of disease means

In the median household income, we see that as income level increases, the chance of survival also increases.

Regarding payer information, we see that Medicare and other patients have the lowest odds of in-hospital mortality, while those with other forms of payment tend to have higher in-hospital mortality.

Finally we see that the location (metropolitan vs. micropolitan vs. non-core) matters as well with large metropolitan areas having the lowest odds of mortality.

4.3 Linear Regression

We did the linear regression modeling, we found a common list of independent variables that gave the highest R^2, or fit for male, female and combined. The independent variables were, disease severity, urban-rural location (metropolitan vs. micropolitan vs. noncore), race, and median income level.

Linear Regression Variables	Natural Log of Length of Stay Male	As Female	Dependent Combined
R^2	0.096	0.079	0.086
Disease Severity Level	0.317	0.273	0.291
Urban-Rural Location	-0.062	-0.064	-0.063
Median Income Level	-0.021	-0.017	-0.019
Race	0.023	0.018	0.02

P Value
All P value are <0.0001

Table 5: Linear Regression with Disease Severity, Urban and Rural location, Median Income Level and Race

Chapter 5: Discussion

What we have found in our data is that there wasn't too much difference compared to what we had read in our literature review. For example, when doing logistic regression, we saw that white people traditionally had a higher odds of survival compared to someone that was black or of Hispanic descent. Even though there was a study that had talked about how for children, the trend held true even if it wasn't consistently that in every state, it could be explained through the United States 2010 census's regional distribution for race and ethnicity. For example, we see that in our region, the South and West part of America had some of the lowest odds for survival. This could be explained by the fact that the southern and western parts of the United States have a larger black and Hispanic population. (USA, 2010) Furthermore, it is these populations that traditionally have a lower socioeconomic condition, which can cause inadequate healthcare, poor lifestyle choices, improper knowledge of risk factors and more. Due all these factors combined, we can see why black and Hispanic people have a higher rate of mortality than white and Asian/Pacific Islanders.

Another one of the expect results was that for disease severity, as disease severity went up, so did in-hospital mortality. One of the possible limiting factors however for this was that it was not as evenly distributed, as we would've liked. There were much less of the people that had high disease severity compared to those that had lower disease severity. Due to the immense difference in sample size, a few outliers could cause massive

deviations, however we think that the general trend would stay the same.

Two of the more interesting variables that were tested were median income level and payer information. We had suspected that median income level was going to have an inverse relationship with in-hospital mortality and that as median income level went up, in-hospital mortality went down. What was surprising is how little of an effect that this had on the actual odds of in-hospital mortality. We had thought that having a higher income would have much higher odds of survival. What we think this means is that since doctors don't know about payment methods beforehand, that treatment of people, no matter of what income level is, is relatively the same. The difference between the odds ratio in income level could be explained by the lack of some external factors that may plague those of lower socioeconomic condition, compared to those of higher socioeconomic condition. For payer information, it was interesting to see that patients that paid with Medicare had lower in-hospital mortality than those of any other payer information. Medicare is a type of healthcare insurance plan by the United States government that benefits those that are 65 years of age and older. Medicaid however, is a type of government assistance program that helps when people have lower levels of income and are from lower socioeconomic backgrounds. (U.S Department of Health and Human Services, 2012) Because Medicare primarily benefits those that are of 65 years and older, we expected that this data would show Medicare not having the lowest in-hospital mortality. We think that this discrepancy may be caused by the fact

that the data is has many more Medicare patients than any other, and that there is some sufficient outliers for people who are paying with private/HMO's and self-pay.

The other interesting variable that we saw was that of people living in urban vs. rural areas. We saw that people who lived in large metropolitan areas had lower odds of in-hospital mortality than any of the other regions. We think that this difference could be explained by the fact that larger hospitals tend to be in busier cities, rather than smaller towns. Combined with the fact that larger hospitals tend to have bigger and more advanced facilities, it is not an unexpected result that there is a lower rate of in-hospital mortality in larger areas. Because of these reasons, we feel that the trend of as urban becomes more rural, the odds of in-hospital mortality increases.

However, despite in-hospital mortality data that fits well in the logistic regression, we found that we could not find a good model of length of stay with linear regression. Before running the experiments, we saw that with scatter-plot data that there was a definite trend in the natural log of length of stay and the independent variables that we chose. However, after running numerous experiments, we found that the R^2 value of our linear regression only topped out at 0.096 for males. We believe that there might be some underlying factors that may have influenced some of the variables, thus creating a very low R^2. We believe that further studies should be done to determine if there are any hidden factors within these variables that could cause these variables not to be a good model for length of stay.

Chapter 6: Conclusion

We found significant relationships between in-hospital mortality and some independent variables that are not directly related to one's physical well-being. One of the factors, income level, was directly related to socioeconomic well-being. It was interesting to see that it did not play as big of a part as other factors such as city size. The other factor that was strongly considered, race; we saw strong disparities between race and in-hospital mortality. Specifically among people of African-American or Hispanic backgrounds, we noticed that the rates of in-hospital mortality were significantly higher compared to people of Asian or White backgrounds. However, we were not able to determine a satisfactory model for length of stay and the independent factors. Further research will need to be done in order to determine if any of the variables are causing interference with each other and other possible hidden interactions. Moreover, we should look at the effects of education and other seemingly non-related factors to in-hospital mortality. We believe that education provides the basis of good overall preventive healthcare and would like to see if the effects of education play a statistically significant role in prevention of in-hospital mortality. We see that evidence of this from regional data from HCUP 2003, and the state data from the 2010 United States Census. As such, it may be more accurate to create a state level analysis, compared to ones that analyze from a national level.

Chapter 7: Reference

American Heart Association. (2013, August 30). *What is cardiovascular disease (heart disease)?*. Retrieved from http://www.heart.org/HEARTORG/Caregiver/Resources/WhatisCardiovascularDisease/What-is-Cardiovascular-Disease_UCM_301852_Article.jsp

Barnhart, J. M., Cohen, O., Wright, N., & Wylie-Rosett, J. (2006). Can non-medical factors contribute to disparities in coronary heart disease treatments?.*Journal of Health Care for the Poor and Underserved*, *17*(3), 559.

Becker, E. R., & Rahimi, A. (2006). Disparities in race/ethnicity and gender in in-hospital mortality rates for coronary artery bypass surgery patients.*Journal of National Medical Association*, *98*(11), 1729.

Benavidez, O. J., & Jenkins, K. J. (2006). Racial and ethnic disparities in mortality following congenital heart surgery. *Pediatric Cardiology*, (27), 321-328.

CDC. (2013, August 28). *America's heart disease burden*. Retrieved from http://www.cdc.gov/heartdisease/facts.htm

Congenital heart defect - corrective surgery. (n.d.).*Congenital heart defect - corrective surgery*. Retrieved from http://www.nlm.nih.gov/medlineplus/ency/article/002948.htm

Davies, H. T. O., & Crombie, I. K. (1998). When can odds ratios mislead?. *BMJ*, (316), 989. Retrieved from http://www.bmj.com/content/316/7136/989

DRG, H. (2003). Definitions manual for all-payer severity-adjusted drg (aps-drgs®) assignment. In HCUP (Ed.), *Definitions Manual for All-Payer Severity-adjusted DRG (APS-DRGs®) Assignment*. Retrieved from http://www.hcup-us.ahrq.gov/db/nation/nis/APS-DRGsV20DefinitionsManual.pdf

HAYES, D. K., DENNY, C. H., KEENAN, N. L., CROFT, J. B., SUNDARAM, A. A., & GREENLUND, K. J. (2006). Racial/ethnic and socioeconomic differences in multiple risk factors for heart disease and stroke in women: Behavioral risk factor surveillance system, 2003. *JOURNAL OF WOMEN'S HEALTH*, *15*(9), 1000-1008.

HCUP. (2013, July). *Healthcare cost and utilization project (hcup)*. Retrieved from http://www.ahrq.gov/research/data/hcup/index.html

Horn, S. D., Sharkey, P. D., Buckle, J. M., Backofen, J. E., Averill, R. F., & Horn, R. A. (1991). The relationship between severity of illness and hospital length of stay and mortality. *Medical Care*, *29*(4), 305-317.

Joensson, A., & Bhandari, M. (2008). *Clinical research for surgeons*. (p. 293).

Kirksey, L. (2011). Health care disparity in the care of the vascular patient. *Vascular and Endovascular Surgery*, *45*(418), Retrieved from http://ves.sagepub.com/content/45/5/418

Mayo Clinic. (2013, July 13). *Heart failure*. Retrieved from http://www.mayoclinic.com/health/heart-failure/DS00061/TAB=indepth

Medline. (2013, August 20). *Vascular diseases*. Retrieved from http://www.nlm.nih.gov/medlineplus/vasculardiseases.html
National Heart, Lung and Blood Institute. (2012, Feb 23).*What is coronary artery bypass grafting?*. Retrieved from What Is Coronary Artery Bypass Grafting?

Overview of Disease Severity Measures. Healthcare Cost and Utilization Project (HCUP). March 2006. Agency for Healthcare Research and Quality, Rockville, MD. www.hcup-us.ahrq.gov/db/nation/nis/severity_overview.jsp.

Qato, D. M., Lindau, S. T., Conti, R. M., Schumm, L. P., Alexander, G. C., & , (2010). racial and ethnic disparities in cardiovascular medication use among older adults in the united states.*pharmacoepidemiology and drug safety*, (19), 834-842.2

Szumilas, M. (2010). Explaining odds ratios. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, *19*(3), 227-229. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2938757/

U.S Department of Health and Human Services. (2012, Decemb 20). *What is the difference between medicare and medicaid?*. Retrieved from http://answers.hhs.gov/questions/3094

USA, C. (2010). *Us 2010 census*. Retrieved from http://www.census.gov/2010census/popmap/

Waterloo, R. R. (2013, June 12). *Heart disease warning signs easily overlooked*. Retrieved from http://www.smgh.ca/heart-disease-warning-signs-easily-overlooked/

WHO. (n.d.). *The atlas of heart disease and stroke*. Retrieved from http://www.who.int/cardiovascular_diseases/resources/atlas/en/

Wynn, B., & Scott, M. (2007). Evaluation of severity- adjusted drg systems addendum to the interim report In B. Wynn (Ed.), *Evaluation of Severity- Adjusted DRG Systems Addendum to the*

Interim Report . Retrieved from https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Reports/downloads/wr434z1.pdf