Variation in Computerized Tomography Scan Utilization

Xiaojin Xie

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science

In

Industrial and Systems Engineering

Christian Wernz, Chair

C. Patrick Koelling

Anthony D. Slonim

08/27/2010

Blacksburg Virginia

Keywords: geographic variation, CT scan, utilization rate, logistic regression

Variation in Computerized Tomography Scan Utilization

Xiaojin Xie

ABSTRACT

The U.S. health care system is one of the most expensive health care systems in the world, yet it is not as efficient as it is expected. Studies have shown that the use of expensive imaging procedures, such as CT scans, was significantly increasing for the past few years. However, the increased number of CT scans may not help to improve quality of care. No studies are conducted on investigate geographic variation on CT scan usage rate. This research is the first one to examine CT scan usage rate among counties and to examine variation caused by patient and hospital characteristics. We used the 2007 HCUP-SID database provided data for the research. GIS graph was used to illustrate geographic variation on CT scan usage in New York State. Contingency tables were developed to evaluate to what extent patient and hospital characteristics contribute to the variation. A logistic regression model was built to control the variation caused by patient and hospital characteristics in order to find variation contributed by other potential factors such as availability of CT scanners and radiologists. Significant geographic variation of CT scan usage rate in the county level of New York State was found. Patient demographics, insurance status and medical conditions as well as hospital bed size and teaching status were contributing factors to the variation. After controlling these factors, significant geographic variation was still found. It indicates that other potential reasons would influence the technology use.

Acknowledgement

I would like to thank all the people who have helped and inspired me during my graduate study.

I especially want to thank my advisor, Dr. Christian Wernz, for his guidance during my research and study at Virginia Tech. His energy and enthusiasm towards research had motivated me. He is always willing to help with my research problems.

Dr. Anthony Slonim, and Dr. Patrick Koelling deserve a special thanks as my thesis committee members and advisors. In particular, I would like to thank Dr. Slonim for his professional advices and help.

I would like to thank Dr. Eric Vance from Statistic Department and Andrew Henry for their inspiration and technical help in the thesis research.

My deepest gratitude goes to my family and John Paul Dunning for their unflagging love and support throughout my life; this thesis would simply have been impossible without them.

Table of Content

List of Tables and Figures

Chapter 1

Introduction

The existence of geographic variation of the usage of medical imaging technologies, for example computed tomography (CT) scans may show inequity in health care. This is the first study that conducts geographic variation on CT scan usage. Using a number of statistical analyses, the study will identify the relationships between CT scan use and patient demographics/hospital characteristic and will examine the geographic variation in the use of the imaging technology.

An analytical model will be developed for assessing CT scan use by taking into account geographic, patient demographic and hospital factors. The specific aim is to identify the usage rate of CT scans among counties in New York State. The usage rate is measured based on the following two prospective: per-capita rate, and per-inpatient rate. It will help to assess the value of CT scan and to assist furthering health policy decision making regarding the location and use of CT scans in this state.

We are interested in understanding the contributors to the utilization of CT scans including how CT scan density, patient, hospital, and county effects contribute to variation used for hospitalized patients, with the hope that it would contribute to health policy decisions regarding the location and use of CT scans. Geographic and demographic variations will be measured by statistical methods and illustrated by GIS technique using maps. Technology usage in terms of counties, patient demographics and hospital characteristics will be described in contingency tables.

We are expected to see geographic variations in CT scan usage. However, the structure of demographics and medical conditions of inpatients as well as hospital characteristics among different areas are sources of the variation. A logistic regression model will be used as the research model to control the patient and hospital characteristics. A map will illustrate the usage rate of CT scan when all factors are controlled.

Goal

To analyze and report geographic variation of CT usage rate in New York State.

Objects

1) To examine general CT scan usage rate in terms of both per-capita and per-inpatient among counties in New York State.
2) To examine the CT scan usage rate associated with patient and characteristics and to demonstrate how these characteristics contribute to the variation found among counties.

3) To examine geographic variation among counties when control patient and characteristics to reveal the variation masked by these factors.

Chapter 2

Background

The health care in the United States (US) is one of the most advanced one in the world. Medical care technologies contribute to the high care quality of care one can receive in the US, yet there are many problems in the current system needed to be addressed as soon as possible. Not everyone in the States has the opportunity to get the high-tech medical treatment. 46.3 million Americans are without insurance coverage, which account for almost 16% of the population. [1]. Health care cost is high, comparing with every other country in the world [2]. There is the fact that the expenses of health care are climbing at rates that exceed not only those of inflation and dollar depreciation but even the Federal government itself [3]. The report “A Message to the Public” concluded that by 2019 the costs of Medicare will exceed its revenue by 20 percent. Medicare and Medicaid together accounted for 22.9 percent of the entire federal budget in 2007 and are on track to consume some 30 percent of it by 2018 [4].

One of the main reasons for the high cost is how doctors, especially specialists are getting paid. Under the fee-for-service method, doctors and hospitals get paid according to the number of the services they deliver. They send the report to insurance companies or the government to get paid for their service. The more procedures they prescribed, the more money they get paid.

In addition, health care in the United States has a high level inequity. Former studies [5-6] find that men are more likely to have access to adequate medical care than women, minorities are less likely to have access to proper medical care than non-ethnic minorities, and within all groupings, and individuals of higher socioeconomic standing are more likely to have access to adequate medical care than individuals of lower socioeconomic standing.

The expensive diagnostic imaging technology CT scan is a good window to examine the inequity in the health care system. It is a widely used standard procedure and can be used on the diagnoses of a wide range of medical problems, and most importantly it is a controversial practice, given its lack of proven benefit, cost, radiation exposure, and the risk of finding the 'incidental' abnormalities that may trigger additional investigations.

CT scan was first introduced in the 1970s. It has been used for preventive medicine or screening for disease. It is an important diagnostic technology tool to assist in diagnosing tumors, fractures, bony structures, and infections in the organs of the chest. Physicians can easily diagnose problems such as cancers, cardiovascular disease, trauma and infectious disease with the detailed pictures provided by the CT scans. CT scans improve the quality of health care by providing earlier diagnostic answers for complex medical problems. This provides for earlier interventions and perhaps improved outcomes [7].

An estimated 72 million scans were performed in the United States in 2007 [8]. The number of CT scans had been increasing dramatically during the past years [9]. The largest increase in CT scans has occurred in children and the screening of adult smokers, in virtual colonoscopy, in cardiac screening, and in whole-body CT for asymptomatic patients [10]. Meanwhile, increasing the availability of CT scans creates a supply-induced demand that is a major contributor to overuse and variability in healthcare [7, 11-13].

The cost of a CT scan ranges from $550 to $3,000. This price differential is due, in large part, to the type of CT scan and the procedure location. CT scans may not improve the diagnostic accuracy for certain conditions, such as diagnoses on appendicitis, previous studies have not found any evidence that the increasing number of CT scans has resulted in improved patients' health conditions [14-16]. As Wennberg and Gittelsohn's [17] research proved almost 40 years ago, more care is not always better care. Determining the right amount of CT scans is important. CT overuse is costly and negatively impacts healthcare quality [18-19]. Attention has been drawn to the adverse occurrences related to the overuse of CT scans including the development of cancer due to the radiation. Laws of Certificate-of-Need (CON) [20] were implemented to ensure the appropriate use of medical care, such as the usage of the expensive imaging technology. The CON is designed to promote the delivery of high-quality health care services and to ensure that facility-based health care services are aligned with community health needs. The CON process also reins in investments in excess facility capacity and unneeded medical equipment that drive up health care costs for everyone, without contributing materially to the health of our communities.

The overuse is also linked with geographic variation in which the overuse of CT scans represents inequities in the health care system [21]. Rathore and Krumholz [22] suggested that variation in health care can be considered "disparities" when the difference in health care reflects shortfalls in appropriate care that cannot be explained by other patient factors. The National Healthcare Disparities Report [23] stated that the disparities of access to health care result from many contributing factors combined. The factors may include age, gender, race and insurance status.

Geographic variation is well used in assessing the pattern of both medical diagnosis and procedures among large areas (such as countries or regions with a country) and among small areas (such as counties or hospital market areas). No study has conducted geographic variation on CT scan usage. We will investigate the usage patterns and variations associated with CT scans in specific geographic areas. The study will help to assess the value of CT scan and to assist furthering health policy decision making regarding the location and use of CT scans in this state.

Chapter 3

Literature review

This chapter provides an overview of the literature related to the study. It is divided in to the following sections: 1) geographic variation related studies; and 2) researches done on CT scan usage.

3.1 Research related to CT scan usage rate

CT scans are expensive. A physician's order of a CT scan for a patient is influenced by patient's medical condition, patient demographic, physician's experience, physician's practice habit, and availability of CT scanners. Several studies were conducted on the following topics about CT scan usage: CT scans usage pattern [24] [15], and the increasing usage of CT scans [9] [16]. The number of studies concerning CT scan usages is low. There is no study done on geographic variation of CT scan usage.

3.1.1 CT scans usage pattern

CT scan usage patterns vary by patients demographic such as age, gender and race. Mettler et al. examined the pattern of CT scan usage at the University of New Mexico Health Sciences Center [24]. The Health Sciences Center is a teaching hospital which is not only a tertiary referral but also is it one of the centers for pediatric and primary care. The authors examined if CT scans usage is associated with a patient's age and gender. They used simple descriptive statistics in the study. The research found that, in the center, 55% percent of CT scans were done on males and 45% were done on females. Patients in age group 36-50 years had the highest CT scan usage rate, and of patients having CT scans 30% had at least three scans. The study also revealed that 67% of the CT scan was the major contributor to diagnostic radiology dose. However, the authors indicated that the research result did not mean that CT scanning was being misused or overused. Note that the study used conventional age groups; however, these groups can hide variability of CT scan usage among ages. Based on the conventional age groups, we identified nine age groups to more accurately demonstrate the variability caused by age; each group has a linear trend in CT scan usage.

3.1.2 The increasing usage of CT scans

Several studies have implemented research in the emergency departments [9, 15] on the increasing usage of CT scans. All of the studies have pointed out the relationship between technology use and cost of health care. Broder and Warshauer [9] conducted a study which aimed to characterize changes in CT utilization in the adult emergency department over a 5-year period, 2000 to 2005. They analyzed CT scans ordered on adult ED patients from July 2000 to July 2005. All the CT scans can be grouped into five groups: head, cervical spine, chest, abdomen, and miscellaneous. ED patient volume and triage acuity (the severity of a patient's illness or injury and the need for immediate evaluation and treatment) were determined. The study found that in the five year period the adult emergency department patient volume increased by 13%, among which head CT increased by 51%, cervical spine CT by 463%, chest CT by 226%, abdominal CT by 72%, and miscellaneous CT by 132%, while triage acuity remained stable. The study also showed that ED CT utilization had increased at a rate far exceeding the growth in ED patient volume. The authors stated that whether this increase in utilization results in improved patient outcomes could not be determined and needed further study. Several other studies have also found that in the past few years CT scan usage increased significantly [9, 16, 19, 25]. These studies make it necessary to investigate the geographic variation of CT scan use to further evaluate the equity of health care.

3.2 Geographic Variation Related studies

3.2.1 Cornerstone of geographic variation researches

Geographic variation in medical area is a very popular topic. Researchers use geographic variation to detect problems in procedures, prescription usages and health care expenditures. Our research is not the first one to examine the geographic variation in this field. It is built upon previous works of studies on the topic of geographic variation in medical area. Wennberg and Gittelsohn [17] published the paper which was the very first paper on geographic variation.

The article examined the extent to which manpower and bed use, expenditures, and utilization vary among hospital service areas in Vermont. Both the quantity of medical services produced and the kinds of cases treated in the study were expressed as utilization rates. They believed that since the medical care in each area was delivered by local physicians, variations tended to reflect differences in the patterns of physicians or medical groups practice medicine.

In the study, the authors grouped towns into hospital service areas surrounding the hospital. The group contains the towns by which the hospital is the most frequently used. The measures of health care delivery included age-adjusted utilization rates as well as indices of manpower, facilities and expenditures. They estimated rates of use of hospital beds, hospital staffs by the proportion to the use of these facilities and staff by residents. A similar method was used to estimate the use-rate of nursing home beds, hospital expenditure, medical manpower, and non-physician.

With a comprehensive study of patterns used in three surgical procedures among the chosen areas on those factors, wide variation was found in resource input, utilization of services, and expenditures among neighboring communities. The limitation of Wennberg's research is that they did not take in to account patients' characteristics which greatly influence the procedures usage. The variation cannot be explained only by the factors they presented in the paper. However, despite the limitation, this research results can help public regulation since they took into account a community's regional ranking in regard to bed input and utilization rates. It also helped in the development of strategies to deal with inequality and uncertainty concerning the effectiveness of health care delivery. It was an important step in the development of rational public policy for health. As a corner stone of geographic variation research, many researchers have cited this research.

3.2.2 Geographic variation research focused on patient characteristics

Since 1973, an increasing number of studies were done on both large area geographic variation and small area geographic variation in various topics in the medical field. For example, in 1987, Chassin et al. studied geographic variation in the appropriateness of use of coronary angiography (CA), carotid endarterectomy (CE), and upper gastrointestinal tract (UGTE) endoscopy in the paper "Does inappropriate use explain geographic variations in the use of health care services: a study of three procedures" [26]. The research measured how appropriately these 3 procedures were performed. The researchers chose 13 sites throughout the U.S. to study the geographic variation. They selected physician claims from 1981 provided by Medicare insurance carriers from these sites. In the study, the appropriateness is divided into three distinct categories: appropriate, equivocal, and inappropriate. Chi-square tests were used to test for differences in the distribution of appropriateness ratings across sites. They found statistically significant geographic variation in appropriateness in the utilization of these 3 procedures (CA, CE, and UGTE) which meant that in terms of surgery usage it may appear inequity. Chi-square tests were used to test for differences in the distribution of appropriateness ratings across sites. They found statistically significant geographic variation in appropriateness in the utilization of these 3 procedures (CA, CE, and UGTE) which meant that in terms of surgery usage it may appear inequity. In the study, they excluded claims for patients living outside an area who came into it for services as well as services performed for residents of an area by physicians practicing outside the area. Excluding out-of-state patients can minimize the variation caused by other geographic areas. And since the sample size of in-state-patient is ideal enough, we accepted this idea when we adjusted the population of counties.

3.2.3 Geographic variation research focused on health care provider characteristics

Havranek et al. [27] presented study on how health care provider characteristics influenced geographic variation in the quality of care on heart failure in the United States among Medicare patients. Documentation of ejection fraction (EF) and angiotensin-converting enzyme (ACE) were selected as indicators of quality of care. They used the datasets from the Centers for Medicare & Medicaid Service's National heart Care Project. Cases of heart failure are identified by ICD-9 codes from the datasets. They utilized empirical Bayesian techniques produced by non-linear mixed effect models on random sample selected from the dataset to examine the variation in terms of both random variation and systematic variation. They also generated maps to illustrate geographic variation. The response variables were the two indicators EF and ACE, and the predictors are provider, hospital and patient characteristics. The study demonstrated geographic variation in the two indicators. It also indicated that the attending physician characteristic of cardiology specialty was associated with more frequently documented EF, and non teaching hospitals tend to fail to prescribe ACE inhibitors. In the North of the U.S., higher rates of EF were found, and significantly low rates of ACE inhibitor prescription were found in the South. In sum, rates of these indicators were generally highest in the northeastern states. EF documentation has greater variation than ACE inhibitor prescription. After the mixed model adjusted for the patient characteristics, both provider and hospital characteristics remained significantly associated with EF documentation but not with ACE inhibitor prescription. This means that provider and hospital characteristics have a large proportion of contribution to the geographic variation, while patient characteristics cause the large variation in the ACE inhibitors model.

The studies discussed above provide us with the insight into the potential causes of geographic variation in medical procedure utilization: hospital characteristics and patient characteristics. Most studies did not take into account patient demographics and health provider characteristics. These two are the most important frequently examined factors which influence variation of procedure use. Both factors will be measured in the study. These characteristics may influence the variation in different medical usages. In our research, we will first investigate how patient characteristics contribute to the CT scan usage variation. Then we will control patient and hospital characteristics to see if there is still variation among the geographic areas. If variation exists, it indicates that there are other factors as contributing factors to geographic variation in CT scan usage.

3.2.4 Methods in research of geographic variation

Statistics methods, such as logistic regression, chi-square test, and ANOVA, are the most common method used on testing geographic variation. We also implemented statistics method for our research. Diehr, Cain and Connell [28] conducted research to try analyze different statistical methods were good to analyze geographic variation. They developed a computer program to simulate the distribution of several descriptive statistics under the null hypothesis that all of the areas have similar underlying surgery rates.

They generated random numbers from normal distribution for each county with an underlying surgery of mean p and variance p(1-p)/n. They assigned observers to randomly select a rate for each county. The observed rates will be different depend on the observers. They calculated the min, max, max/min of the observed rates, and used 95% percentile of the EQ, chi-square, CV (coefficient of variance) and SCV (Systematic component of Variation) to test for differences in the surgery rate among counties. After evaluating the results, they showed that EQ method can well detect variation, that the chi-square method was a good way of testing for variability if the probability of readmission (a person in the numerator more than once) was zero and the expected number of cases per county is not too small. They also found out that SCV can be a measurement for comparing several surgery types, or the same surgery in two different regions. We selected Chi-square as our research method. Since we recorded whether or not a patient had CT scans instead of the total number of CT scans, there will be no problem for readmission.

3.2.5 Guidelines of geographic variation research

We need to be aware of the several guidelines when study geographic variation besides the research factors and the method. Volinn et al [29] published a paper "Why geographic variation matters". They believed that unexplained variations raise serious questions about the quality, appropriateness and cost effectiveness of health care. In the paper, they proposed seven questions in evaluating studies on geographic variation and, more specifically, small area analysis (county variation)

The seven questions are as following:

1) What events are to be analyzed
2) What geographic units are to be analyzed
3) How good are the data
4) Are differences in rates due to chance alone

5) Are high rates too high
6) How is geographic variation to be explained
7) What is the role of "presentation style" in explaining geographic variation

Based on the guideline, our research is on CT scan usage in counties in New York State. The data from AHRQ is widely used and the data was cleaned and ready to be used in the data modeling. We explained the geographic variation by patient characteristics and other potential causes such as hospital characteristics and availability of CT scanners. We used maps to illustrate geographic variation along with odds ratio table.

Chapter 4

Methods

The following sections will describe the data source, data cleaning methods and analytical methods in the study.

4.1 Data Source

4.1.1 New York State

New York is a state in the Northeastern region of the United States and is the nation's third most populous state. New York State covers 54,556 square miles (141,300 km2) and ranks as the 27th largest state by size [30]. Both well urbanized and rural counties can be found in the state. It also has large number of minorities residing in the state. These facts make the state a good place for the investigation of variation and disparity in CT scan usage.

4.1.2 State Inpatient Database

This research uses the 2007 edition of the State Inpatient Database (SID) from the Health Cost and Utilization Project (HCUP), a Federal-State-Industry partnership sponsored by the Agency for Healthcare Research and Quality (AHRQ). We used 2007 SID of New York State (NYSID2007) for our research on variation of CT scan use.

The SID is a powerful set of hospital databases from data organizations in the participating states. Twenty-six data organizations participating in the HCUP have agreed to release their SID files through the HCUP Central Distributor under the auspices of the AHRQ. The individual state databases are in the same HCUP uniform format and represent 100% of records processed by AHRQ. The SID is a well developed set of databases for research that requires complete enumeration of hospitals and discharges within hospital market areas or States. Researchers and policymakers use SID to investigate questions to one state or counties in the state; to compare data from two or more States; to conduct market area research or small area variation analyses; and to identify State-specific trends in inpatient care utilization, access, charges, and outcomes [31].

NYSID2007 contains complete discharge data from 206 hospitals in New York State in 2007. 203 of the hospitals are identified as community hospitals, while 3 are not community hospitals. American Hospital Association (AHA) defines community hospitals as “nonfederal, short-term general and special hospitals whose facilities and services are available to the public. (Specialty hospitals include obstetrics and gynecology; eye, ear, nose, and throat; rehabilitation; orthopedic; and other individually described specialty services.) Short-term general and special children's hospitals are also considered to be community hospitals” [32]. The dataset contains more than 100 clinical and nonclinical variables included in a hospital discharge abstract, such as:

patient principal and secondary diagnoses, principal and secondary procedures, admission and discharge status, patient demographics (gender, age, race and so forth), method of payment (Medicare, Medicaid, private insurance, self-pay, and other payment methods), detailed charged items for hospitalized patients as well as hospital and county identifiers that allow linkage to AHA's datasets and the Area Resource File.

4.2 Population

The accuracy of population size by county is critical for effective evaluation of CT scan usage density per capita. The U.S. Census Bureau, through their publication 'County and State Quick Facts', provides data for unadjusted population size by county for the US [33]. While the number of people residing in a given county provides an estimate of CT service demand within that county, this unadjusted population size does not effectively account for patients who travel across county lines to receive medical care. To account for cross-county travel, we adjusted the population size by taking the weighted population average in each Patient County by using the following analytic sequence.

1) P_i is the number of inpatients that reside in county i.
2) $H_{i,j}$ is the number of inpatients from county i who went to a hospital in county j.
3) C_i is the population size in county i.
4) For the new variable Adjusted County, the population ADJi= $\sum_i \frac{H_{i,j}}{P_i} \cdot C_i$.

When discussing population estimate, "per capita" will refer to the adjusted county population.

4.3 Procedure Identification

In SID, CT scans performed on each patient are recorded by several methods: ICD-9-CM codes (International Statistical Classification of Diseases and Related Health Problems), grouped ICD-9 codes which includes CCS (Clinical Classifications Software) and multi-level CCS, revenue codes, and utilization flags. ICD-9-CM for CT scans is: 87.03 for head, 87.41 for chest, 88.01 for abdomen, and 00.31, 87.71 and 88.38 for other CT scans. Single level CCS codes for CT scan are: 177 for head, 178 for chest, 179 for abdomen, and 180 for other CT scans. Multi-level CCS code for CT scan is 16.02 for all CT scans with sub-level of 16.02.01 for head, 16.02.02 for chest, 16.02.03 for abdomen, and 16.02.04 for other CT scans. When CT scans are performed, they can either be recorded by ICD-9-CM codes, revenue codes or both. A previous study [34] demonstrated that ICD-9-CM codes, CCS codes, and multi-level CCS codes underreport CT scan usage. When compared with revenue codes, ICD-9-CM procedure coding was found to be considerably underreported and variable, with only 33% of CT being reported. In the 2007 SID, CT usage was also found to be underreported by revenue codes. A Utilization Flag (UF) was created by AHRQ to track whether or not a patient had CT scans in 2007, and to track the method of how the procedures were recorded (ICD-9 code, revenue code or both). We used both ICD-9 codes and revenue codes to count the number of CT scans for each patient to assure appropriate sensitivity in capturing the CT scans performed.

4.4 Study Variables

We determined the association of CT scan usage with the following categories of independent variables obtained from SID: 1) Patient demographic characteristics, 2) Diagnosis and health status of patient, 3) Utilization, and 4) Hospital characteristics. According to National Healthcare Disparities Report [6] patient demographic and health status directly influence variation in medical utilization. Based on a common consensus in the medical field, we believed that hospital characteristics also contribute to the variation of procedure utilization rate. Patient demographics, method of payment and utilization of CT scans were drawn directly from SID. Data for population is provided by Census Bureau. We chose to measure patients' medical conditions by MDC (Major Diagnostic Categories). The MDC is formed by dividing all possible principal diagnoses (from ICD-9-CM) into 25 mutually exclusive diagnosis areas. Data of hospital characteristics (including bed size, teaching status and location) were collected through New York Department of Health. Details are listed in Table 1.

Table 1 - Variable Description

Variable		Description
Key	HCUP record identifier	A unique patient identifier assigned by HCUP. It is the key of the dataset to link with other SID datasets if needed.
Responsive variable	Number of CT scans	An integer variable created to record the number of CT scans a patient had during the year of 2007.
	CT	A binary variable to identify whether or not a patient had CT scans in 2007.
Patient related variables	Patient County	A categorical variable created by linking the FIPS code in the dataset with data provided by Census Bureau. This variable shows the name of the county each patient resided in during the year of 2007
	Age	A continuous variable which gives the age of patients at their admission in 2007
	Gender	A binary variable where 1 represents female and 0 for male.
	Race	A categorical variable with six categories of race and one missing category.
	Method of payment	A categorical variable with three categories of insured payment methods and three categories of uninsured payment methods.
	Major Diagnosis Categories (MDC)	A categorical variable with 24 major diagnosis categories and one pre-MDC category.
Hospital related variables	Hospital Bed Size	A categorical variable with three categories of measurement for number of beds in a hospital
	Hospital Teaching Status	A binary variable to identify if a hospital is a teaching hospital, 1 is for teaching hospitals, and 0 is for non-teaching hospitals
	Hospital County	A categorical variable created by linking the county code in the dataset with county data provided by SAS institute. This variable shows the name of the county each hospital resided in during the year of 2007

4.5 Analytic Method

4.5.1 Summary Statistics

To investigate the association of the independent variables on CT scan usage, descriptive statistics were calculated and were reported as the mean probability with their odds ratios and confidence intervals.

The odds ratio is the ratio of the odds of an event occurring in one group to the odds of it occurring in another group. It is a measure of effect size, describing the strength of association or non-independence between two binary data values [35]. According to the county versus CT in table 2, for example, in the county of Albany the probability of having CT scans is 26.5% while the probability of not having CT scans is 73.5%, then the odds of Albany are 26.5/73.5=0.361. While, the state level odds are 0.389, comparing with the state average level, the odds ratio of Albany is 0.361/0.389=0.928.

The summary statistics is the gateway to see the differences of utilization rate among different groups and examine the disparity and inequity in the health care system.

4.5.2 Logistic Regression

A logistic regression model was used to control the influence of combinations of variables on the outcome of interest. Logistic regression is defined by equation (1) and (2)

$$p(z) = \frac{e^z}{e^z + 1} = \frac{1}{1 + e^{-z}} \qquad (1)$$

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + ... + \beta_k x_k \qquad (2)$$

where $p(z)$ represents the probability of a patient receiving CT scans.

The values of the regression coefficients are determined as $\beta_0, \ldots, \beta_k$ for each variable $x_1, \ldots, \beta$. The x_k is a binary variable, and the value of x_k is 1 when the patient's characteristic fall in the corresponding category, and 0 otherwise. The intercept β_0 is the value of z when the value of all independent variables (predictors) is zero. Each of the regression coefficients describes the level of the contribution of that factor. A positive regression coefficient means that the explanatory variable increases the probability of the outcome, while a negative regression coefficient means that variable decreases the probability of that outcome. A large regression coefficient means that the factor strongly influences the probability of that outcome; while a near-zero regression coefficient means that the factor has little influence on the probability of that outcome [36].

The reference groups in the research are: Chemung (county), white (race), female (gender), age from 30 to 39 (age group), Eye problem (MDC), non-teaching hospital (teaching status), and medium size (hospital bed size). All the reference groups combining together create a reference patient for the research. The β_0 in the model is the point estimate for the reference patient. With setting other variables with a corresponding β to 1, we can calculate the usage rate for other types of patients. The logistic regression result shows the variation in CT scan use for the reference patient. The limitation is that we cannot take out all the influences (reference patient is needed) by patient and hospital characteristics to investigate if other potential causes exist or not. However it can discover the variation masked by the composition inside of each category factors and reveal inequity of health care delivery.

4.6 Software used in the research

Cleaning the datasets and implementing the statistical method were done by Statistical Analysis System (SAS) base/stat and Structured Query Language (SQL). SAS is a powerful integrated system of software products -provided by SAS Institute Inc. that enables programmers to perform various analytical tasks.

The usage density variables were virtualized with GIS software, ArcGIS. ArcGIS is a suite consisting of a group of geographic information system (GIS) software products produced by ESRI. It enables data presentation in a color coded county map of New York State. We chose to use a series of graduate colors to present the variation of CT scan usage rate.

Chapter 5

Results

5.1 CT scan use by Population

Figure 1 and Figure 2 demonstrate the geographic variation in inpatient CT scan utilization measured by the number of CT scans/county population and CT scans/100 adjusted county population respectively. A total of six counties had a value of 0 because these counties do not have community hospitals on file from AHRQ. These counties are: Greene, Hamilton, Lewis, Seneca, Tioga and Washington.

Significant geographic variation is found in both maps. The highest rate in Figure 2 is lower than the highest one in Figure 1. The higher usage and variation to some extent spread out to the surrounding counties. For example, in Western New York State, before adjustment of population, the usage rate in this region ranges from low to high level without significant patterns on how the usage rate distributed. It is obvious from the maps that the adjusted population soothes the variation of CT scan usage rate comparing with the rate in terms of original population. In addition, a higher usage ‘center’ can be found after adjustment of population. The same pattern is found in all other regions (Central New York, Capital District/Upper New York, and New York Metro area and Long Island) in New York State after the adjustment of population. Except for the average high usage rate in New York Metro area and Long Island, all other regions show a radical pattern. It is commonly known that medical usage pattern is radical with a center in the middle. The adjusted population is representing the reality more accurately.

Figure 2 indicates that more populated counties do not have a higher probability of inpatient CT scan use and less populated counties do not have a lower probability of inpatient CT scan use. For example, the population density in Erie County is almost three times larger than in Dutchess County, but the probabilities of receiving a CT scan from these two counties are similar. Meanwhile, the most populated Counties, New York County, Queens County and Kings County are all in the lower range of usage density.

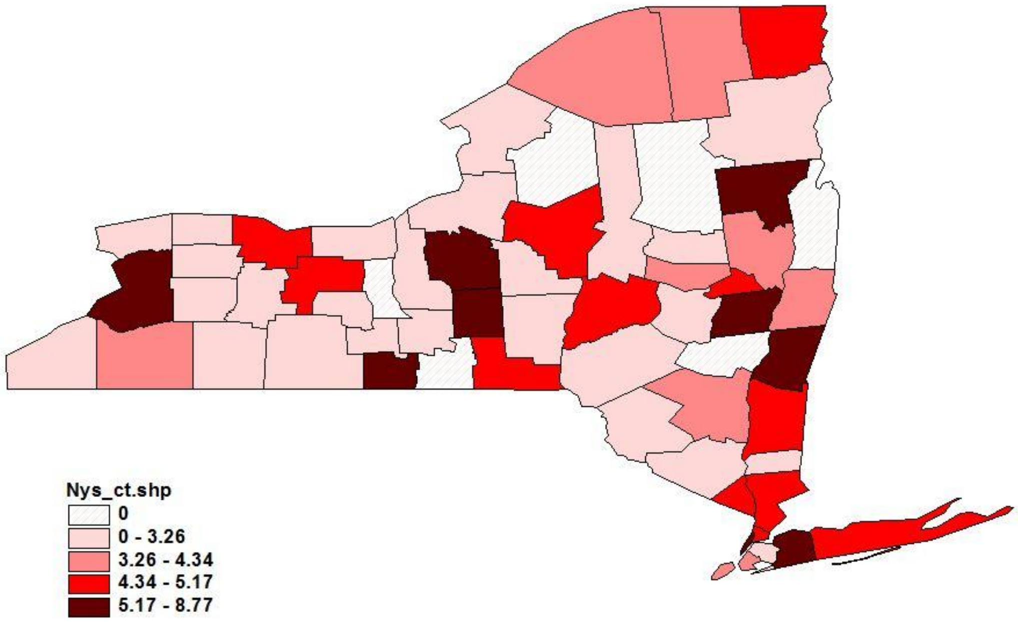

Figure 1-Computerized Tomography (CT) Scan Use per 100 Population

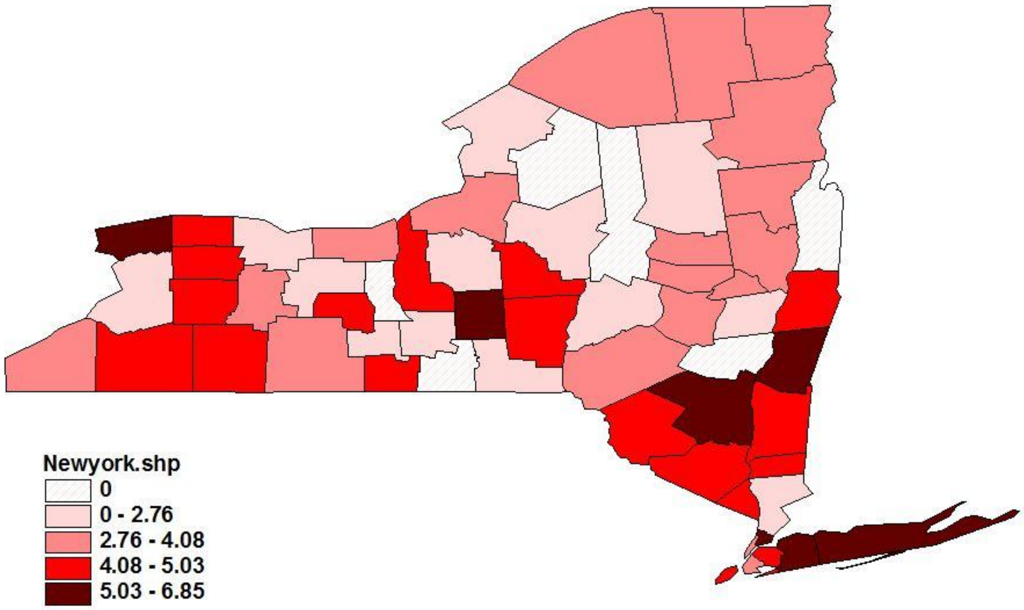

Figure 2 - Computerized Tomography (CT) Scan Use per 100 Adjusted Population

5.2 CT scan use by inpatients

The per capita variation provides a glimpse into the geographic variation in CT scan usage. It assists the analysis on how medical outreach influences the technology use in different geographic areas. Medical outreach includes availability of CT scanners, ratio of physicians to population, percentage of radiologists among physicians and so forth. Variation found in Figure 2 may be influenced by population structure. But the database only contains inpatient data. Due to the limitation of NYSID, It’s more appropriate to examine variation of CT scan usage through a per-inpatient perspective than per-capita one which contains too many variations beyond measurement, such as the medical outreach and level of people's access to health care in a certain county.

Table 2 demonstrates CT scan utilization, based on the descending odds ratios, for inpatients by county in New York State. The inpatient measurement is the percentage of inpatients that had one or more CT scans during hospitalization in that County. Cortland County had the highest rate of CT inpatient utilization with an odds ratio 1.554; Otsego County had the lowest rate with an odds ratio of 0.532.

Table 2 - Computerized Tomography (CT) Utilization, based on the descending odds ratios, for inpatients by county in New York States

County Name	Percentage of CT Scans Used	Odds Ratio	95% Confidence Interval UB	LB
Chemung, NY *	27.81%	1		
Cortland, NY	37.45%	1.554	1.615	1.493
Rockland, NY	37.23%	1.539	1.577	1.502
Suffolk, NY	34.50%	1.367	1.4	1.334
Ulster, NY	34.14%	1.345	1.39	1.301
Cayuga, NY	33.42%	1.303	1.368	1.238
Niagara, NY	33.06%	1.282	1.328	1.235
Orange, NY	32.99%	1.278	1.317	1.238
Columbia, NY	32.80%	1.267	1.327	1.207
Putnam, NY	32.16%	1.23	1.285	1.176
Rensselaer, NY	32.14%	1.229	1.274	1.185

Nassau, NY	32.02%	1.222	1.255	1.19
Dutchess, NY	31.88%	1.215	1.253	1.177
Warren, NY	31.63%	1.201	1.245	1.156
Franklin, NY	31.62%	1.2	1.262	1.139
Cattaraugus, NY	31.11%	1.172	1.226	1.118
Erie, NY	30.51%	1.14	1.173	1.106
Fulton, NY	30.21%	1.124	1.195	1.052
Broome, NY	30.20%	1.123	1.163	1.083
Schoharie, NY	30.08%	1.117	1.283	0.951
Essex, NY	30.07%	1.116	1.272	0.96
Queens, NY	29.71%	1.097	1.13	1.064
Westchester, NY	29.50%	1.086	1.12	1.053
Allegany, NY	29.08%	1.064	1.147	0.981
Onondaga, NY	29.08%	1.064	1.099	1.03
Genesee, NY	28.90%	1.055	1.119	0.991
Clinton, NY	28.71%	1.045	1.095	0.995
Herkimer, NY	28.65%	1.042	1.122	0.962
Livingston, NY	28.36%	1.027	1.106	0.949
Saratoga, NY	27.50%	0.984	1.026	0.942
Montgomery, NY	27.28%	0.974	1.035	0.913
Steuben, NY	27.24%	0.972	1.026	0.917
Sullivan, NY	27.18%	0.969	1.034	0.904
Chenango, NY	27.16%	0.968	1.064	0.871
Bronx, NY	26.89%	0.955	0.988	0.922
Orleans, NY	26.85%	0.953	1.04	0.865
Oneida, NY	26.65%	0.943	0.982	0.904
Richmond, NY	26.62%	0.942	0.978	0.906
Albany, NY	26.45%	0.934	0.97	0.898

Schenectady, NY	26.30%	0.926	0.968	0.884
Ontario, NY	26.03%	0.914	0.962	0.865
Wyoming, NY	25.83%	0.904	0.991	0.817
St. Lawrence, NY	25.80%	0.902	0.951	0.854
Madison, NY	25.66%	0.896	0.96	0.832
Wayne, NY	25.66%	0.896	0.977	0.815
Monroe, NY	25.21%	0.875	0.909	0.84
Tompkins, NY	25.09%	0.869	0.928	0.81
Kings, NY	24.80%	0.856	0.889	0.823
Chautauqua, NY	24.38%	0.837	0.885	0.789
Oswego, NY	23.88%	0.814	0.873	0.755
New York, NY	23.75%	0.809	0.841	0.776
Yates, NY	22.57%	0.757	0.887	0.626
Delaware, NY	20.66%	0.676	0.807	0.545
Jefferson, NY	20.67%	0.676	0.732	0.621
Schuyler, NY	17.28%	0.542	0.682	0.402
Otsego, NY	17.01%	0.532	0.585	0.479

*reference

Figure 3 shows the variation in CT scan utilization by 100 inpatients. There are four counties with high level usage rate: Onondaga and Cortland in Central New York, Ulster in New York Metro area, and Columbia in Upper New York. In these four counties, other than Onondaga County three of them have low population density. Onondaga County has a higher population density and a metropolitan area Syracuse.

New York Metro area has a relevant overall higher CT scan usage rate than other regions in the state. However, surprisingly New York City area (Manhattan, Kings, Queens, Bronx and

Staten Island) has only mid-level usage rate. Variation in Western New York is lower while it is higher in Central and Upper New York.

The inpatient CT scan usage is unrelated to the number of inpatients. For example, in the southeastern area of New York State, Manhattan (New York County) and Brooklyn (Kings County) for instance, have lower odds of having CT scans (0.809 and 0.856, respectively). The correlation coefficient between the number of inpatients and number of CT scan is 0.038 indicating that higher inpatient volume does not lead to the increase in CT scan utilization for inpatients.

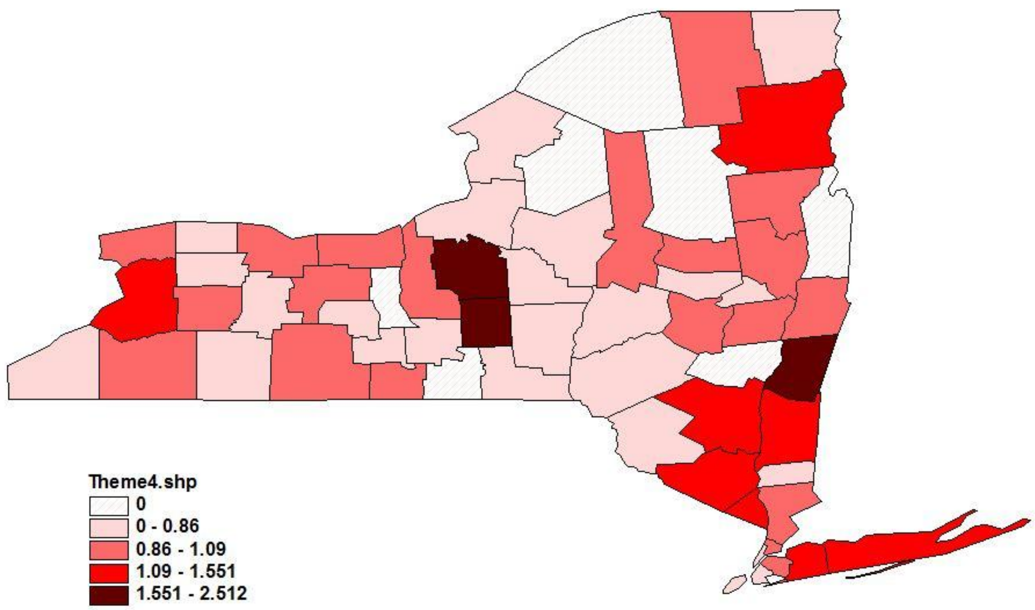

Figure 3 - Percentage of Inpatients with Computerized Tomography (CT) Use

5.3 Patient Characteristics associated with Inpatient CT Scan Use

5.3.1 Patient age and CT usage rate

To follow the research convention in the medical field, ages were put in age groups instead of being used as a continuous variable. To appropriately develop the boundary for age groups, the relationship between age and CT utilization among hospitalized patients are plotted in Figure 3. The graph shows significant variation in the use of CT scans across ages. The number of CT scans per patient increased with age for patients < 18 years and between the ages of 18 and 30 CT scan utilization per patient decreased. After age 30, the number of CT scans per patient again increases with age. Within each of these intervals, age and CT usage rates have a linear relationship. For future research convenience we developed more detailed age groups. These age groups follow the conventional pediatric age groups for ages under 18, young and mid-aged adults above 18, and also break ages over 65 in one group due to Medicare.

Based on figure 4 and conventional age groups we developed the nine age groups listed in Table 3 for further analysis. The nine age groups are birth to year of 1 not including 1, 1 to 4, 5 to 13, 14 to 17, 18 to 29, 30 to 39, 40 to 64, 65 to 84 and 85 and above. These groups minimized the variation of CT scans in each of them to the utmost so that the variation among age groups would not be masked by the variation within the groups.

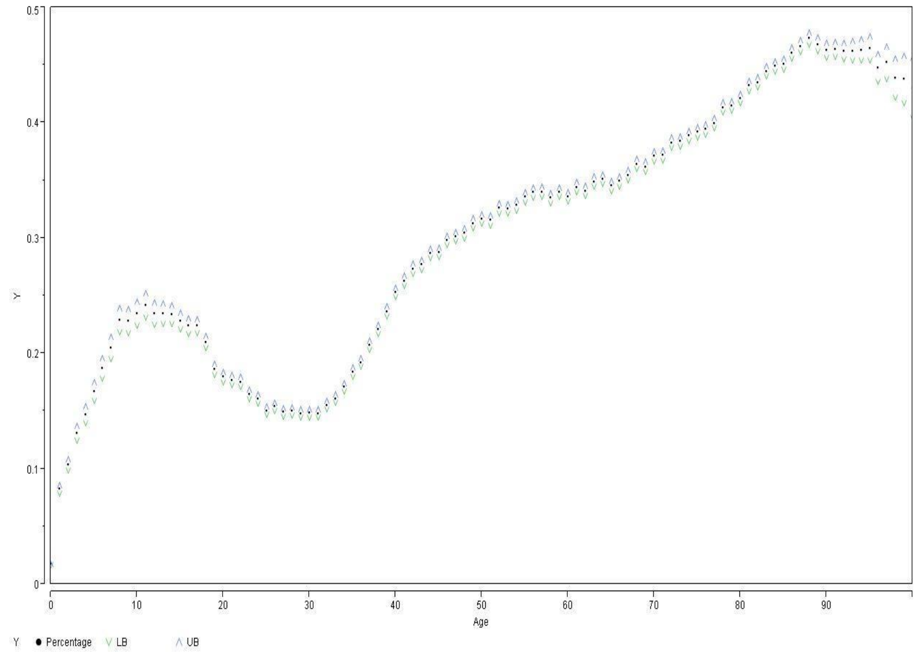

Figure 4 - Age and Computerized Tomography (CT) Use (Point Estimates and 95% Confidence Intervals)

5.3.2 Patient demographics and CT usage rate

There were 2,608,615 inpatients in New York State in 2007, of which 730,197 (27.99%) of received CT scans. 43.37% of the inpatients were male and 56.62% were female. The race variable had six categories with 59.25% White, 18.48% Black, 1.41% Native Americans, 3.63% Asian, 0.8% Pacific Islander, 13.35% with two and more races and 3.07% not reported. The nine age groups were 10.59% for ages 0 to 1 years not including 1, 1.51% for ages 1 to 4 years, 1.49% for ages 5 to 13 years, 1.72% for ages 14 to 17 years, 10.71% for age 18 to 29 years, 10.49% for age 30 to 39 years, 29.84% for age 40-64 years, 25.59% for age 65 to 84 years, and 8.07% for age 85 years and above. There were 35.51% of the inpatients with Medicare, 24.39 with

Medicaid, 32.22% were covered by private insurance and 7.88% were uninsured (5.60% self-pay, 0.16% no charge and 2.13% with other methods of payment).

Table 3 provides the estimates of inpatient CT utilization for various patient characteristics. Males were more likely to have CT scans than females. Whites were more likely to have CT scans than inpatients in other racial and ethnic group. Medicare beneficiaries had the highest odds of receiving a CT scan. Inpatients older than age 65 had considerably higher chance of receiving a CT scan than any other age groups; whereas, infants have the lowest odds of receiving a CT scan.

Table 3 - Inpatient Demographic and Computerized Tomography (CT) Utilization

Variable	Percentage of CT Scan Used	Odds Ratio	95% Confidence Interval	
			UB	LB
Female*	26.69%	1		
Male	29.73%	1.162	1.167	1.157
White*	29.78%	1		
Black	27.17%	0.88	0.887	0.873
Native American	24.36%	0.759	0.782	0.736
Asian	22.36%	0.679	0.694	0.664
Pacific Islander	27.06%	0.875	0.904	0.846
Other	24.24%	0.755	0.763	0.747
Missing	24.41%	0.762	0.778	0.746
30 to 39*	18.23%	1		

Birth to 1 not include 1	1.75%	0.08	0.11	0.05
1 to 4	10.72%	0.538	0.572	0.505
5 to 12	21.43%	1.224	1.25	1.198
13 to 17	22.78%	1.323	1.347	1.299
18 to 29	16.44%	0.882	0.896	0.868
40 to 64	31.64%	2.076	2.087	2.066
65 to 84	39.50%	2.929	2.94	2.918
85 and older	46.13%	3.84	3.853	3.827
Private Insurance*	22.66%	1		
Medicare	39.72%	2.249	2.253	2.245
Medicaid	18.83%	0.792	0.799	0.785
Self-pay	24.59%	1.113	1.124	1.102
No charge	33.20%	1.696	1.757	1.635
Other	26.42%	1.225	1.244	1.206

*reference

5.3.3 Patient medical condition and CT usage rate

Table 4 presents the association of MDC categories with CT scan utilization. Trauma was the diagnostic category with the highest CT scan utilization 45.08%, while Human Immunodeficiency Virus Infection had the lowest (9.23%).

Table 4 - MDC and Computerized Tomography (CT) Utilization

MDC	Description	Percentage of CT Scan Used	Odds Ratio	95% Confidence Interval	
				ub	lb
1*	Nervous System	0.303	1		
2	Eye	0.222	0.658	0.725	0.591
3	Ear, Nose, Mouth And Throat	0.228	0.681	0.707	0.655
4	Respiratory System	0.281	0.898	0.91	0.887
5	Circulatory System	0.315	1.06	1.07	1.051
6	Digestive System	0.299	0.981	0.992	0.97
7	Hepatobiliary System And Pancreas	0.274	0.871	0.889	0.852
8	Musculoskeletal System And Connective Tissue	0.33	1.136	1.148	1.124
9	Skin, Subcutaneous Tissue And Breast	0.251	0.771	0.79	0.753
10	Endocrine, Nutritional And Metabolic System	0.273	0.864	0.88	0.848
11	Kidney And Urinary Tract	0.278	0.886	0.902	0.871
12	Male Reproductive System	0.283	0.907	0.944	0.871
13	Female Reproductive System	0.302	0.998	1.019	0.978
14	Pregnancy, Childbirth And Puerperium	0.254	0.785	0.796	0.774
15	Newborn And Other Neonates (Perinatal Period)	0.261	0.813	0.825	0.802
16	Blood and Blood Forming Organs and Immunological Disorders	0.237	0.716	0.741	0.69
17	Myeloproliferative DDs (Poorly Differentiated Neoplasms)	0.231	0.69	0.72	0.661
18	Infectious and Parasitic DDs	0.291	0.946	0.964	0.928
19	Mental Diseases and Disorders	0.246	0.753	0.768	0.738
20	Alcohol/Drug Use or Induced Mental Disorders	0.194	0.555	0.573	0.538
21	Injuries, Poison And Toxic Effect of Drugs	0.28	0.896	0.92	0.872
22	Burns	0.251	0.771	0.853	0.689
23	Factors Influencing Health Status	0.271	0.855	0.876	0.834

24	Multiple Significant Trauma	0.451	1.892	1.959	1.824
25	Human Immunodeficiency Virus Infection	0.092	0.234	0.295	0.174

*reference

5.4 Hospital Characteristics Associated with Inpatient CT Scan Use

In general, hospital bed size, teaching status and location are the three primary indicators to assist with measuring quality of care. Bed size of a hospital determines the level of inpatients. Figure 5 indicates that bed size was not a major contributor to the variation of CT usage rate. In addition, it is well believed in the medical field that teaching status influences the number of CT scan prescribed. A teaching hospital tends to order more CT scans than a non-teaching hospital due to the lack of experiences of interns. However, based on table 5, teaching status of a hospital is not a contributor to the variation either.

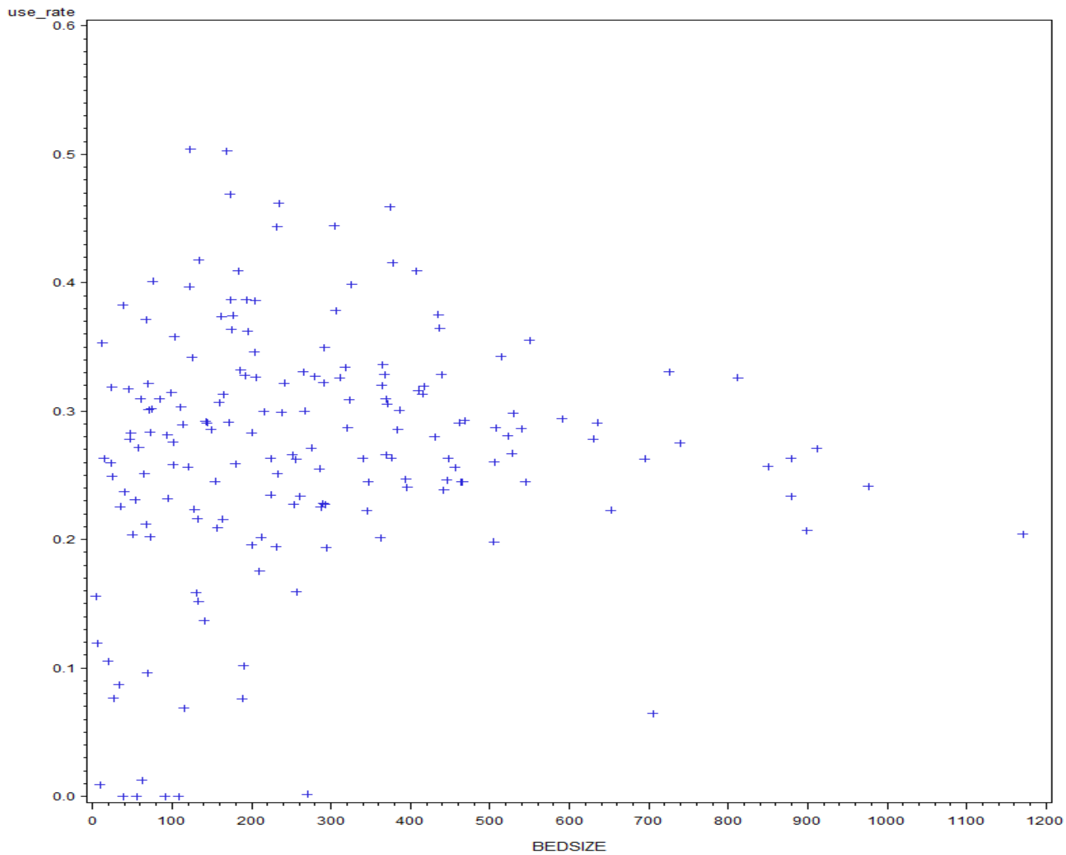

Figure 5 - Computerized Tomography (CT) Utilization and Bed Size

Usage rates of CT scans between two teaching status and among the three bed sizes are listed below in table 5.

Table 5 - Computerized Tomography (CT) Utilization associated with Hospital Teaching Status and Bed Size

Variable	Percentage	Odds Ratio	Confidence Interval	
			UB	LB
Non-teaching*	30.66%	1.000		
Teaching	27.09%	0.961	0.968	0.955
Medium*	30.02%	1.000		
Large	27.79%	0.893	0.899	0.886
Small	25.70%	0.806	0.815	0.798

*reference

5.5 Logistic Regression Results

The fixed effects, logistic regression results for CT scan utilization are presented in Table 6.

In this research the reference is chosen as a white female patient from Chemung County with private insurance who is between age 30 and 39 without any MDC, and goes to a medium-sized non-teaching hospital. There are 103 predictors in the logistic regression model.

The model controls patient and hospital characteristic. The result of logistic regression does not have the influences from the composition of patient and hospital, yet geographic variation is still found. The variation is illustrated in figure 6. In addition after controlling patient characteristics, it appears that teaching hospitals have higher CT scan usage rate than non-teaching hospitals.

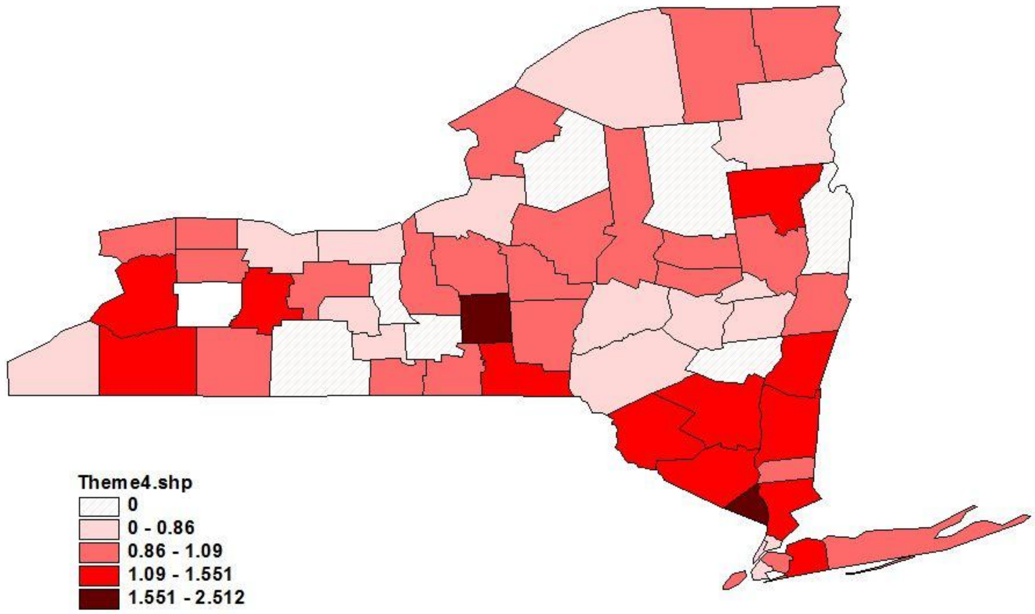

Figure 6 - Geographic Variation from Logistic Regression

Two counties are found to have high usage rate, they are: Cortland and Rockland. Neither of these two counties have major hospitals or cities residing in them. New York Metro area still have overall higher usage rate than other regions. The variations of usage rate in areas other than New York Metro are lowered compared with Figure-3 which means that geographic variation is reduced when taking out the variation contributed by composition of patient and hospital characteristics. Yet the variation is still significant. In addition, no regular patterns are found in the variation.

In addition, if one needs to know the real probability of getting CT scans in a certain county when a specific kind of patients goes to the hospital. We'll use the equation of the model to get the probability. Given a patient who is a 65 year old white with Medicare from New York

County with head trauma from a large size teaching hospital, the probability of him getting CT scans is calculated as below:

Set the predictor of male (gender), age group (8), Medicare, New York County, MDC 24, teaching, and large size as 1(no need to set the race since it's the same as the reference).

The value of logistic regression model is z= 1.0159+(-0.0246)*1+0.2326*1+(-0.2839)*1+0.0528*1+1.2027*1+(-0.0435)*1+(-0.0955)*1=2.0565

Convert the z value into probability: pr=1/(1+e^(-z))=88.66%

So the probability of this type of patients getting CT scans is 88.66%.

Table 6 - Logistic Regression Results

		Odds Ratio Estimates (compared with baseline)		
Effect	Estimates	Odds Ratio	95% Wald Confidence Limits	
Intercept	1.0159			
Albany, NY	-0.209	0.811	0.777	0.847
Allegany, NY	-0.0836	0.92	0.832	1.017
Bronx, NY	-0.1636	0.849	0.816	0.884
Broome, NY	0.2214	1.248	1.189	1.31
Cattaraugus, NY	0.1203	1.128	1.058	1.203
Cayuga, NY	-0.0397	0.961	0.891	1.037

Chautauqua, NY	-0.2361	0.79	0.747	0.835
Chenango, NY	-0.1231	0.884	0.789	0.99
Clinton, NY	0.0627	1.065	1.004	1.129
Columbia, NY	0.1633	1.177	1.097	1.263
Cortland, NY	0.4562	1.578	1.464	1.701
Delaware, NY	-0.6146	0.541	0.469	0.623
Dutchess, NY	0.2768	1.319	1.26	1.38
Erie, NY	0.0944	1.099	1.056	1.144
Essex, NY	-0.2647	0.767	0.647	0.911
Franklin, NY	-0.032	0.969	0.9	1.042
Fulton, NY	0.0252	1.025	0.942	1.117
Genesee, NY	0.0056	1.006	0.932	1.085
Herkimer, NY	-0.1038	0.901	0.822	0.988
Jefferson, NY	-0.1263	0.881	0.825	0.941
Kings, NY	-0.2012	0.818	0.786	0.851
Livingston, NY	0.1236	1.132	1.033	1.24
Madison, NY	0.0588	1.061	0.984	1.143
Monroe, NY	-0.1776	0.837	0.803	0.873
Montgomery, NY	-0.0161	0.984	0.916	1.057
Nassau, NY	0.1677	1.183	1.137	1.23
New York, NY	-0.2839	0.753	0.724	0.783

Niagara, NY	-0.0294	0.971	0.919	1.026
Oneida, NY	-0.1422	0.867	0.829	0.908
Onondaga, NY	-0.0024	0.998	0.956	1.041
Ontario, NY	-0.1146	0.892	0.842	0.944
Orange, NY	0.1297	1.139	1.086	1.193
Orleans, NY	0.0482	1.049	0.949	1.16
Oswego, NY	-0.232	0.793	0.741	0.849
Otsego, NY	-0.8483	0.428	0.403	0.455
Putnam, NY	0.0778	1.081	1.014	1.153
Queens, NY	0.0827	1.086	1.043	1.131
Rensselaer, NY	0.0829	1.086	1.031	1.145
Richmond, NY	-0.0246	0.976	0.934	1.019
Rockland, NY	0.6978	2.009	1.92	2.103
Saratoga, NY	-0.0486	0.953	0.907	1.001
Schenectady, NY	-0.3016	0.74	0.704	0.777
Schoharie, NY	-0.3283	0.72	0.6	0.865
Schuyler, NY	-0.6811	0.506	0.433	0.591
St. Lawrence, NY	-0.2193	0.803	0.759	0.85
Steuben, NY	-0.0995	0.905	0.85	0.965
Suffolk, NY	0.3208	1.378	1.325	1.434

Sullivan, NY	0.0112	1.011	0.937	1.091
Tompkins, NY	0.0774	1.081	1.007	1.159
Ulster, NY	0.2268	1.255	1.189	1.323
Warren, NY	0.1688	1.184	1.124	1.247
Wayne, NY	-0.175	0.839	0.763	0.923
Westchester, NY	0.1529	1.165	1.119	1.213
Wyoming, NY	6.5318	686.652	<0.001	>999.999
Yates, NY	-5.3082	0.005	<0.001	>999.999
Birth to 1 not include 1	-2.2436	0.106	0.101	0.111
1 to 4	-1.829	0.161	0.155	0.167
5 to 12	-0.8829	0.414	0.401	0.426
13 to 17	-0.3426	0.71	0.69	0.731
18 to 29	0.0401	1.041	1.023	1.059
40 to 64	0.0365	1.037	1.024	1.051
65 to 84	0.2326	1.262	1.243	1.281
85 and older	0.4725	1.604	1.576	1.633
Black	0.1282	1.137	1.126	1.148
Native American	0.0271	1.027	0.999	1.057
Asian	0.1243	1.132	1.11	1.155
Pacific Islander	0.1758	1.192	1.148	1.239
Other	0.1129	1.12	1.107	1.132
Missing	-0.0562	0.945	0.927	0.964
Medicare	0.0528	1.054	1.043	1.065
Medicaid	0.052	1.053	1.042	1.064
Self-pay	0.2133	1.238	1.218	1.258

No charge	0.4652	1.592	1.48	1.713
Other	-0.0192	0.981	0.959	1.003
MALE	-0.0246	0.976	0.969	0.982
MDC 2	-0.498	0.608	0.571	0.647
MDC 3	-1.0082	0.365	0.355	0.375
MDC 4	-1.4793	0.228	0.224	0.231
MDC 5	-2.1625	0.115	0.113	0.117
MDC 6	-0.9	0.407	0.4	0.413
MDC 7	-0.9376	0.392	0.384	0.4
MDC 8	-2.1637	0.115	0.113	0.117
MDC 9	-2.2851	0.102	0.1	0.104
MDC 10	-1.9577	0.141	0.138	0.144
MDC 11	-1.2558	0.285	0.28	0.29
MDC 12	-2.8867	0.056	0.053	0.059
MDC 13	-2.8762	0.056	0.055	0.058
MDC 14	-4.9858	0.007	0.007	0.007
MDC 15	-3.3097	0.037	0.034	0.039
MDC 16	-2.0411	0.13	0.126	0.134
MDC 17	-1.751	0.174	0.168	0.179
MDC 18	-1.103	0.332	0.325	0.339
MDC 19	-2.9645	0.052	0.05	0.053
MDC 20	-3.0951	0.045	0.044	0.046
MDC 21	-1.6595	0.19	0.185	0.195
MDC 22	-2.947	0.052	0.046	0.06
MDC 23	-2.3226	0.098	0.096	0.101
MDC 24	1.2027	3.329	2.964	3.74
MDC 25	-0.8929	0.409	0.394	0.426
Non-teaching	-0.0435	0.957	0.948	0.967

Bed size L	-0.0955	0.909	0.901	0.917
Bed size S	-0.2445	0.783	0.774	0.792

Chapter 6

Conclusion

6.1 General variation in CT scan use

The study found significant geographic variation in CT scan use in New York State among hospitalized patients. The four counties with highest CT usage rate are Onondaga, Cortland Ulster, and Columbia York. The New York Metro area has a higher CT scan usage rate than other regions in the state. Variation of usage in Western New York is lower while it is higher in Central and Upper New York.

Disparity was found in CT scan usage. Males are more likely to have CT scans than females. Whites are more likely to have CT scans than inpatients in other racial groups. Medicare beneficiaries have the highest odds of receiving a CT scan. The number of CT scans per patient increases with age for patients < 18 years and decreases between the ages of 18 and 30. After age 30, the number of CT scans per patient increases again. Inpatients older than age 65 have considerably higher odds of receiving a CT scan than any other age groups, and infants have the lowest odds of receiving a CT scan.

Teaching hospitals ordered less CT scans than non-teaching hospitals. Medium-size hospitals performed more CT scans than small-size and large-size hospitals.

6.2 Geographic variation after controlling factors

We used reference groups to control the influences caused by composition of patient and hospital factors. After controlling for factors that contribute to variation in CT scan usage, geographic variation of CT scan usage remains. Cortland and Rockland have the relevantly higher usage rate than any other counties. Neither of these two counties have major hospitals or cities residing in them. The New York Metro area has the highest usage rate among the regions. The variations of CT usage rate in areas other than New York Metro are lowered compared with Figure 2 which means that it is reduced when taking out the usage variation contributed by composition of patient and hospital characteristics. Yet the variation is still significant.

Chapter 7

Discussion

Studies have been showing that CT scan usage is rising in the past years. None of them has been done on the geographic variation of the CT scans. Our research conducted generally geographic variation analysis on CT scan usage and the examined how demographic and medical condition of patients influence in the variation in New York State. We can see the variation of CT scan usage among counties with Cortland as the highest and Otsego as the lowest CT usage county.

We used geographic variation which Wennberg and Gittelsohn [17] first developed in 1973 for an imaging procedure. They used geographic variation on examine the manpower, bed input and utilization rates in a hospital associated with three surgical procedures. Later on, several studies were done from this prospective. However, those factors are greatly influenced by structure of population demographic and medical condition. Without controlling the factors, a geographic variation analysis on medical procedures could be biased. Our analysis found geographic variation after controlling variation caused by those factors mentioned above.

In Wennberg's study, he grouped towns into hospital service areas surrounding the hospital used most frequently by the town. We improved this idea by adjusting the population with the weighted average. Volinn [29] suggested that the number of residents should be used in geographic variation instead of adjusted population which may cause big inflation in some areas that lead to a "centered" pattern of medical usage according to them. But our study showed that the adjusted population actually smoothed the variation among counties.

Our results on the relationship between inpatients' demographics and CT usage frequency agrees with the broad consensus that health care disparity caused by patients' demographics and medical characteristics (age, gender, race, method of payment, and MDC) contribute to the variation of health care received by patients. Our study found that male, white and old adult inpatients are more likely to have CT scans than other groups. However, when other factors were controlled, male patients have less CT scans than female patients

In addition, inpatients' insurance status is one of the several contributors to CT scan usage variation. Our research analysis shows that CT usage rates vary among inpatients with different insurance status. Medicare beneficiaries have a much higher chance to have CT scans than any other payment groups. Uninsured inpatients have slightly higher probability to have CT scan than patients with insurance that is not Medicare.

Undoubtedly, a CT scan is ordered for a patient by a doctor based on the patient's medical conditions. Consequently, patients' health condition also directly influences CT scan usage variation. Our study finds that trauma is the top cause of CT scan usage. When a certain area has a high rate of trauma, if other conditions hold average, the area has a higher CT usage rate which causes variation.

Meanwhile, hospital's characteristics are contributing factors to the variation of CT scan usage rate. We found great variability in CT scan usage among different bed sizes. And hospital teaching status is believed in the medical field to influence usage rate.

We took each demographic and medical factor of inpatients and hospital characteristics mentioned above into consideration in the logistic regression model, and controlled the structure of them, yet the geographic variation still exists. The variation may indicate the overuse of CT scans in some counties. Causes of such variation can be the level of CT scanners availability [37], number of radiologists and other physicians in the region, practice habit of physicians [38] and so forth. Determining the relationship between CT usage and these potential causes is beyond the data availability of SID. The aforementioned potential causes are hypotheses that can inform future research.

Several study limitations must be considered. The age groups we developed are different from the original ones provided by Census Bureau. For further usage to investigate the relationship of CT and population, this may cause limitations in comparability since the population data of new age groups are not directly available. The SID only includes the inpatient data from community hospitals in New York State; it does not cover the CT scans performed on outpatients and patients from other hospitals. Some of the limitations can be eliminated by third party, such as AHRQ, by improving the quality of the data.

Despite these limitations, the study has significant strength since it demonstrates and analyzes the area variation of an expensive medical image technology in terms of patient characteristic. In addition, the relationship between CT scan usage and age is an important finding of the study and can serve as a guideline for physicians' practice. Further, the results can guide to enhance the policy of Certificate-of-Need (CON) on regulating CT scan usage, which constrains investments in unneeded medical equipment that drive up health care costs for everyone, without contributing materially to the health of the community. No other study has illustrated the relationship between age and CT utilization. Our study also provides suggestions for future research on the causes of geographic variation.

Reference

1. Hall, M. and J. Fritze, *Health care: Five faces of the uninsured*, in *USA Today*. 9/25/2009.
2. Dykman, J., *Truths About Health Care in America*, in *TIME*. 12/1/2008.
3. Sennholz, H.F., *Why is Medical Care so Expensive*, in *Mises Daily*. 8/22/2006.
4. *The State of American Hospitals-Taking the Pulse*. 2009, American Hospital Association.
5. Merzel, C., *Gender differences in health care access indicators in an urban, low-income community*. American Journal of Public Health, 2000. **90**(6): p. 909-916.
6. *National Healthcare Disparities Report*. 2007, Agency for Healthcare Research and Quality (AHRQ).
7. Brown, J. *MRI And CT Scans For Children: Improving Health Care Delivery*. 2006; Available from: http://www.medicalnewstoday.com/articles/43288.php.
8. Gonz ález, A.B.d., et al., *Projected cancer risks from computed tomographic scans performed in the United States in 2007. Arch. Intern. Med.*, 2009. **169**(22): p. 2071-2077
9. Broder, J. and M. Warshauer, *Increasing utilization of computed tomography in the adult emergency department, 2000-2005.* Emergency Radiology, 2006. **13**: p. 25-30.
10. Wikipedia. Available from: http://en.wikipedia.org/wiki/X-ray_computed_tomography.
11. Gazelle, G.S., et al., *Cost-effectiveness Analysis in the Assessment of Diagnostic Imaging Technologies.* Radiology, 2005. **235**: p. 361-370.
12. Goodman, C.S., *Healthcare Technology Assessment: Methods, Framework, and Role in Policy Making.* Am J Manag Care, 1998. **4**: p. 200-214.
13. Skinner, J.S., D.O. Staiger, and E.S. Fisher, *Is Technological Change in Medicin Always Worth it? The Case of Acute Myocardial Infarction.* Health Affairs, 2006. **25**(2).
14. Lee, S.L., A.J. Walsh, and H.S. Ho, *Computed Tomography and Ultrasonography Do Not Improve and May Delay the Diagnosis and Treatment of Acute Appendicitis.* Arch Surg, 2001. **136**(5): p. 556-562.
15. Klassen, T.P., et al., *Variation in Utilization of Computed Tomography Scanning for the Investigation of Minor Head Trauma in Children: A Canadian Experience.* Academic Emergency Medicine, 2000. **7**(7): p. 739-744.
16. Patrick, D.A., et al., *Increased CT Scan Utilization Does Not Improve the Diagnostic Accuracy of Appendicitis in Children.* Journal of Pediatric Surgery, 2003. **38**(5): p. 659-662.
17. Wennberg, J. and A. Gittelsohn, *Small Area Variations in Health Care Delivery: A Population-based Health Information System Can Guide Planning and Regulatory Decision Making.* . Science, 1973. **182**: p. 1102-1108.
18. Brenner, D.J., et al., *Estimated Risks of Radiation-Induced Fatal Cancer from Pediatric CT.* American Journal of Roentgenology, 2001. **176**: p. 289-296.
19. Nelson, R., *Thousands of New Cancers Predicted Due to Increased Use of CT.* Arch Intern Med, 2009. **169**: p. 2049-2050, 2071-2077, 2078-2086.
20. Health, N.Y.D.o. *Certificate of Need*. Available from: http://www.health.state.ny.us/facilities/cons/.
21. Rao, V., et al., *Use Trends and Geographic Variation in Neuroimaging: Nationwide Medicare Data for 1993 and 1998.* American Journal of Neuroradiology, 2001. **22**: p. 1643-1649.

22. Rathore, S. and M. Krumholz, *Differences, Disparities, and Biases: Clarifying Racial Variations in Health Care Use.* Ann Intern Med, 2004. **141**: p. 635-638.
23. in *National Healthcare Disparities Report (NHDR)*. 2009, U.S. Department of Health & Human Services, Agency for Healthcare Research and Quality.
24. Mettler, F.A., et al., *CT scanning: patterns of use and dose.* Journal of Radiological Protection, 2000. **20**(4).
25. *Growth Of CT Scan Use May Lead To Significant Public Health Problem*. 11/29/2007 Available from: http://www.medicalnewstoday.com/articles/90114.php.
26. MR, C., et al., *Does inappropriate use explain geographic variations in the use of health care services? A study of three procedures.* JAMA, 1987. **258**(18): p. 2533-2537.
27. Havranek, E.P., et al., *Provider and Hospital Characteristics Associated With Geographic Variation in the Evaluation and Management of Elderly Patients With Heart Failure.* Arch Intern Med, 2004. **164**: p. 1186-1191.
28. Diehr, P., et al., *What is too much variation? The null hypothesis in small-area analysis.* Health Services Research, 1990. **24**(6): p. 741-771.
29. Volinn, E., et al., *Why Does Geographic Variation in Health Care Practice Matter?* SPINE, 1994. **19**(185): p. 2092-2100.
30. *Land and Water Area of States*. 2000; Available from: http://www.infoplease.com/ipa/A0108355.html.
31. *Introduction to the HCUP State Inpatient Database (SID)*. 2010, AHRQ.
32. *Community Hospital Definition*, State of Rhode Island Executive Office of Health and Human Services.
33. *State & County QuickFacts*, U. S. Census Bereau.
34. Dismuke, C.E., *Underreporting of computed tomography and magnetic resonance imaging procedures in inpatient claims data.* Medical Care, 2005. **43**(7): p. 713-717.
35. Mosteller, F., *Association and Estimation in Contingency Tables.* Journal of the American Statistical Association, 1968. **63**(321): p. 1-28.
36. Hosmer, D. W, and S. Lemeshow, *Applied Logistic Regression*. 2nd ed. 2000: Wiley.
37. *MRI Services in Western New York 2003*. 2004, The Niagara Health Quality Coalition.
38. Gawande, A., *The Cost Conundrum*, in *The New Yorker*. June 1st, 2009.

Appendix

1) Map of New York State

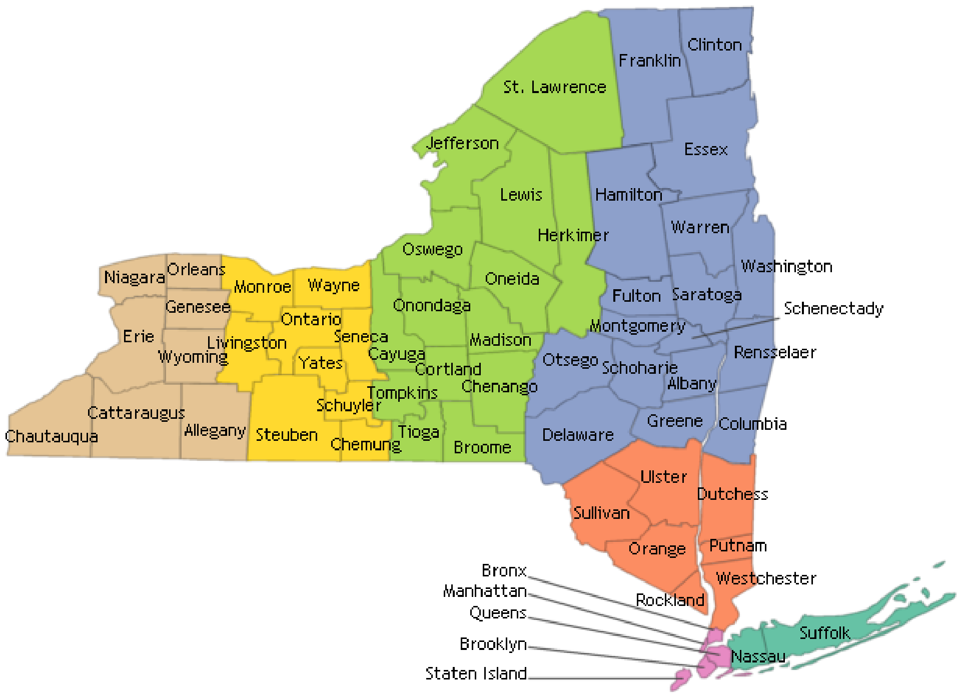

Figure 7 - Map of New York State

2) Categories of variables from SID

Table 7 - Categories of Variables from SID

Utilization Flag: Computed Tomography Scan

Value	Description
0	Service is not reported on the record
1	Service is indicated only by a UB-92 revenue code
2	Service is indicated only by a ICD-9-CM procedure code or CCS procedure category
3	Service is indicated both by a UB-92 revenue code and ICD-9-CM procedure code or CCS procedure category.

Method of Payment

Value	Description
1	Medicare
2	Medicaid
3	Private Insurance
4	Self-pay
5	No charge
6	Other
.	Missing
.A	Invalid

Race

Value	Description
1	White
2	African American (Black)
4	Asian
5	Native Hawaiian or Other Pacific Islander
3	Native American (American Indian, Eskimo, Aleut)
88	Other
99, Blank	Missing

3) Rates of CT scan use

Table 8 - Data for Maps

	CT/Adj 100 capita	CT/100 capita	CT/100 patient	CT usage after logistic
Albany County	4.57	7.89	38.14	0.811
Allegany County	4.19	1.92	32.72	0.92
Bronx County	5.08	4.98	33.81	0.849
Broome County	3.69	4.65	32.49	1.248
Cattaraugus County	5.06	3.79	35.24	1.128
Cayuga County	4.25	2.29	37.76	0.961
Chautauqua County	3.26	3.13	28.00	0.79
Chemung County	5.56	7.23	36.23	1
Chenango County	4.18	1.45	34.13	0.884
Clinton County	3.86	4.54	32.83	1.065
Columbia County	6.86	5.74	58.80	1.177
Cortland County	5.92	5.61	51.37	1.578
Delaware County	3.16	0.74	24.57	0.541
Dutchess County	4.84	4.85	40.05	1.319
Erie County	5.17	5.93	40.25	1.099
Essex County	3.77	0.80	42.24	0.767
Franklin County	4.23	4.23	37.55	0.969
Fulton County	4.66	2.62	36.26	1.025
Genesee County	4.32	3.05	31.89	1.006
Greene County	0.00	0.00	0.00	0
Hamilton County	0.00	0.00	0.00	0
Herkimer County	4.69	1.89	36.72	0.901
Jefferson County	2.59	2.17	23.70	0.881
Kings County	3.82	3.33	30.31	0.818

Lewis County	0.00	0.00	0.00	0
Livingston County	3.67	1.72	31.82	1.132
Madison County	3.67	2.78	31.51	1.061
Monroe County	3.98	5.10	35.57	0.837
Montgomery County	4.05	3.65	27.30	0.984
Nassau County	5.41	8.23	42.47	1.183
New York County	4.02	8.48	29.63	0.753
Niagara County	5.60	2.36	37.23	0.971
Oneida County	4.53	4.91	32.07	0.867
Onondaga County	5.80	8.77	49.37	0.998
Ontario County	3.81	4.50	34.21	0.892
Orange County	5.66	2.97	42.30	1.139
Orleans County	3.98	1.97	30.64	1.049
Oswego County	3.62	1.92	28.78	0.793
Otsego County	2.56	5.07	21.47	0.428
Putnam County	3.59	2.63	32.17	1.081
Queens County	4.54	3.01	38.24	1.086
Rensselaer County	4.34	3.85	36.65	1.086
Richmond County	3.99	3.71	30.29	0.976
Rockland County	4.94	4.70	40.64	2.009
Saratoga County	4.09	3.78	35.22	0.953
Schenectady County	3.98	4.95	31.65	0.74
Schoharie County	3.39	0.65	33.12	0.72
Schuyler County	2.17	1.33	18.06	0.506
St. Lawrence County	3.92	3.79	29.75	0.803
Seneca County	0.00	0.00	0.00	0

Steuben County	4.12	2.94	31.06	0.905
Suffolk County	5.40	4.64	43.34	1.378
Sullivan County	4.55	2.30	31.50	1.011
Tioga County	0.00	0.00	0.00	0
Tompkins County	2.09	1.94	25.09	1.081
Ulster County	6.02	3.90	46.84	1.255
Warren County	4.13	8.76	35.03	1.184
Washington County	0.00	0.00	0.00	0
Wayne County	4.02	1.32	35.71	0.839
Westchester County	4.26	4.59	34.89	1.165
Wyoming County	4.31	2.52	36.51	0.932
Yates County	3.68	1.70	31.98	0.005

Table 9 - Detailed Rates of CT Utilization among Different Patient Demographic Factors

Detailed Rates of CT Usage among Male and Female

County Name	Gender	
	Male	Female
Albany, NY	28.00%	25.18%
Allegany, NY	33.69%	26.23%
Bronx, NY	28.01%	26.04%
Broome, NY	34.04%	27.26%
Cattaraugus, NY	32.29%	30.27%
Cayuga, NY	34.35%	32.82%
Chautauqua, NY	26.05%	23.19%
Chemung, NY	29.35%	26.63%

Chenango, NY	32.87%	23.83%
Clinton, NY	30.74%	27.17%
Columbia, NY	35.58%	30.88%
Cortland, NY	41.95%	34.60%
Delaware, NY	19.60%	21.49%
Dutchess, NY	33.28%	30.80%
Erie, NY	32.28%	29.21%
Essex, NY	29.26%	30.56%
Franklin, NY	34.72%	29.36%
Fulton, NY	34.04%	27.76%
Genesee, NY	30.14%	28.08%
Herkimer, NY	29.13%	28.33%
Jefferson, NY	25.53%	17.59%
Kings, NY	27.10%	23.17%
Livingston, NY	32.01%	25.98%
Madison, NY	30.47%	22.72%
Monroe, NY	27.42%	23.49%
Montgomery, NY	28.23%	26.59%
Nassau, NY	33.27%	31.06%
New York, NY	25.19%	22.53%
Niagara, NY	35.02%	31.79%
Oneida, NY	28.04%	25.59%
Onondaga, NY	31.09%	27.59%
Ontario, NY	27.84%	24.73%
Orange, NY	34.06%	32.16%
Orleans, NY	27.08%	26.69%

Oswego, NY	25.28%	22.93%
Otsego, NY	18.95%	15.43%
Putnam, NY	33.77%	30.89%
Queens, NY	32.28%	27.76%
Rensselaer, NY	32.04%	32.22%
Richmond, NY	26.45%	26.78%
Rockland, NY	39.47%	35.60%
Saratoga, NY	31.13%	25.08%
Schenectady, NY	28.24%	24.84%
Schoharie, NY	31.48%	29.01%
Schuyler, NY	20.97%	15.50%
St0 Lawrence, NY	26.55%	25.27%
Steuben, NY	28.96%	26.13%
Suffolk, NY	37.32%	32.48%
Sullivan, NY	29.61%	24.90%
Tompkins, NY	27.80%	23.36%
Ulster, NY	33.47%	34.69%
Warren, NY	34.70%	29.28%
Wayne, NY	25.56%	25.74%
Westchester, NY	31.10%	28.27%
Wyoming, NY	30.03%	22.96%
Yates, NY	21.09%	23.99%

Detailed Rates of CT Scan Usage among Age Groups

County Name	AGE GROUPS								
	age 0-1	age 1-4	age 5-13	age 14-17	age 18-29	age 30-39	age 40-64	age 65-84	age 85+

Albany, NY	2.07%	13.55%	19.81%	27.26%	18.36%	19.29%	29.87%	34.93%	42.33%
Allegany, NY	0.00%	8.11%	24.29%	23.81%	12.62%	22.17%	39.53%	43.60%	45.79%
Bronx, NY	1.27%	6.64%	16.07%	18.78%	16.35%	21.37%	31.78%	40.27%	44.93%
Broome, NY	1.10%	12.61%	33.03%	40.49%	14.71%	20.46%	36.51%	38.14%	43.92%
Cattaraugus, NY	0.65%	8.00%	24.53%	21.54%	13.73%	18.22%	35.10%	43.76%	44.84%
Cayuga, NY	0.66%	5.26%	30.77%	29.31%	13.30%	19.78%	37.63%	40.33%	42.50%
Chautauqua, NY	0.68%	11.63%	20.33%	13.64%	10.10%	15.28%	28.45%	33.30%	36.58%
Chemung, NY	2.14%	8.07%	31.53%	29.44%	11.90%	19.00%	29.25%	39.56%	46.52%
Chenango, NY	0.34%	0.00%	30.00%	33.33%	8.16%	24.80%	35.92%	38.31%	34.96%
Clinton, NY	0.18%	16.95%	13.27%	14.33%	14.05%	19.63%	34.70%	38.20%	43.21%
Columbia, NY	0.00%	0.00%	46.43%	36.90%	11.16%	25.06%	37.38%	41.56%	44.46%
Cortland, NY	0.16%	10.64%	29.63%	28.77%	17.55%	27.96%	44.93%	52.71%	55.07%
Delaware, NY	0.00%	0.00%	0.00%	0.00%	6.00%	14.04%	16.53%	22.04%	28.83%
Dutchess, NY	0.56%	8.28%	29.45%	25.21%	21.90%	19.91%	35.69%	44.13%	51.02%
Erie, NY	3.21%	16.69%	28.60%	28.32%	16.99%	19.32%	33.13%	41.48%	47.57%
Essex, NY					37.50%	43.75%	37.16%	26.09%	30.54%
Franklin, NY	0.78%	12.50%	24.49%	38.37%	15.71%	24.55%	35.27%	39.88%	45.88%
Fulton, NY	0.00%	7.81%	30.56%	22.22%	10.13%	29.32%	39.86%	41.09%	43.98%
Genesee, NY	0.17%	0.00%	30.00%	25.00%	8.97%	13.82%	32.09%	40.44%	44.99%
Herkimer, NY	0.43%	10.53%	33.33%	40.63%	12.05%	20.77%	31.38%	34.46%	37.15%
Jefferson, NY	0.49%	4.90%	29.25%	23.47%	6.93%	12.96%	31.25%	36.55%	40.18%
Kings, NY	1.29%	8.82%	20.08%	21.53%	13.51%	17.11%	30.89%	37.94%	40.79%
Livingston, NY	0.00%	0.00%	35.71%	25.00%	10.50%	27.62%	34.50%	35.98%	41.32%
Madison, NY	0.23%	11.29%	38.03%	31.08%	12.45%	23.98%	29.55%	35.06%	42.86%
Monroe, NY	1.44%	13.98%	21.44%	19.69%	14.98%	17.62%	28.93%	33.44%	41.21%
Montgomery, NY	0.19%	4.17%	26.47%	13.79%	9.50%	17.64%	27.85%	36.62%	42.52%

Nassau, NY	1.62%	11.18%	19.49%	23.92%	20.26%	18.44%	34.76%	42.94%	50.97%
New York, NY	1.25%	11.33%	15.47%	16.19%	16.29%	15.12%	26.74%	33.94%	44.58%
Niagara, NY	0.11%	9.09%	45.00%	44.55%	17.39%	23.59%	36.39%	39.48%	41.43%
Oneida, NY	0.39%	8.30%	33.79%	31.00%	11.42%	16.89%	26.62%	35.80%	41.74%
Onondaga, NY	1.64%	13.62%	22.38%	21.83%	16.66%	20.70%	33.74%	40.20%	49.32%
Ontario, NY	0.07%	7.89%	41.38%	22.99%	10.54%	15.94%	30.19%	34.62%	37.34%
Orange, NY	0.19%	8.38%	34.11%	25.17%	15.41%	22.25%	36.19%	45.24%	48.60%
Orleans, NY	0.00%	14.29%	30.77%	26.92%	10.92%	9.68%	28.97%	34.11%	41.90%
Oswego, NY	0.13%	3.05%	23.53%	6.22%	9.82%	17.92%	28.10%	35.37%	37.34%
Otsego, NY	0.00%	0.99%	16.79%	11.81%	8.72%	11.58%	19.70%	22.59%	24.20%
Putnam, NY	0.00%	0.00%	52.00%	50.79%	16.83%	17.49%	31.58%	43.12%	53.00%
Queens, NY	0.88%	7.64%	21.61%	29.91%	17.39%	20.07%	34.99%	44.33%	48.80%
Rensselaer, NY	0.08%	6.48%	27.08%	31.96%	17.92%	23.32%	32.81%	43.19%	45.82%
Richmond, NY	1.42%	14.39%	31.97%	33.97%	17.65%	16.47%	28.20%	39.08%	42.87%
Rockland, NY	25.76%	50.20%	66.82%	54.25%	28.39%	27.85%	40.16%	41.85%	50.67%
Saratoga, NY	0.04%	5.45%	35.53%	29.91%	11.81%	17.38%	32.63%	39.33%	43.88%
Schenectady, NY	0.15%	5.77%	13.60%	8.89%	12.52%	17.02%	29.25%	34.38%	43.45%
Schoharie, NY				66.67%	17.65%	33.33%	25.83%	33.92%	27.56%
Schuyler, NY	0.00%	7.69%	42.86%	60.00%	11.41%	19.47%	33.94%	17.35%	15.17%
St0 Lawrence, NY	0.08%	12.35%	34.53%	32.97%	14.51%	18.79%	29.78%	34.89%	36.08%
Steuben, NY	0.25%	9.26%	19.83%	7.71%	10.49%	20.35%	31.68%	39.23%	40.57%
Suffolk, NY	1.50%	13.55%	30.47%	33.64%	21.50%	20.37%	39.62%	48.03%	54.34%
Sullivan, NY	0.38%	2.91%	18.92%	17.76%	16.25%	15.28%	30.55%	41.18%	45.50%
Tompkins, NY	0.20%	15.38%	29.63%	10.83%	13.09%	12.94%	31.94%	37.52%	40.93%
Ulster, NY	0.12%	5.77%	42.22%	36.00%	13.43%	20.55%	35.83%	44.66%	47.76%
Warren, NY	0.26%	7.41%	33.79%	23.48%	12.53%	21.79%	36.52%	41.94%	46.29%

Wayne, NY	0.00%	0.00%	27.78%	26.32%	11.30%	15.71%	28.10%	40.14%	43.97%
Westchester, NY	1.78%	11.41%	23.54%	24.44%	19.08%	16.56%	31.64%	42.96%	49.00%
Wyoming, NY	0.30%	0.00%	72.73%	26.32%	8.11%	11.61%	27.82%	41.13%	41.10%
Yates, NY		0.00%	33.33%	100.00%	7.19%	14.89%	21.75%	26.46%	30.49%

Detailed Rate of CT Scan Usage among Race Groups

County Name	Race						
	White	Black	Native American	Asian	Pacific Islander	Other	Missing
Albany, NY	27.87%	26.78%	33.33%	17.89%	12.40%	20.67%	23.12%
Allegany, NY	29.13%	23.81%				16.67%	33.33%
Bronx, NY	32.18%	27.01%	24.68%	24.62%	32.65%	24.10%	30.63%
Broome, NY	31.61%	27.66%	25.00%	17.65%		21.98%	23.53%
Cattaraugus, NY	31.74%	16.60%	30.32%	25.00%		12.00%	66.67%
Cayuga, NY	33.83%	32.20%	0.00%	20.00%		0.00%	0.00%
Chautauqua, NY	24.96%	19.94%	24.66%	13.71%		18.87%	15.29%
Chemung, NY	28.11%	24.50%	33.33%	20.59%		15.71%	27.91%
Chenango, NY	26.12%	57.14%	0.00%	0.00%		100.00%	32.73%
Clinton, NY	28.96%	27.56%	25.00%	30.00%		22.12%	15.71%
Columbia, NY	33.94%	28.61%	33.33%	10.53%		17.27%	9.52%
Cortland, NY	37.74%	31.43%	0.00%	50.00%		18.18%	29.63%
Delaware, NY	20.66%	0.00%	0.00%			45.45%	0.00%
Dutchess, NY	34.22%	29.22%	39.47%	21.83%		24.25%	9.95%
Erie, NY	31.33%	28.84%	27.65%	16.95%	0.00%	24.75%	31.89%
Essex, NY	29.79%	50.00%				50.00%	50.00%
Franklin, NY	31.79%	27.87%	36.90%	0.00%		26.79%	27.94%

Fulton, NY	30.40%	21.43%		50.00%		0.00%	
Genesee, NY	29.33%	21.93%	26.79%	0.00%		20.34%	10.00%
Herkimer, NY	28.75%	15.00%	0.00%	25.00%		0.00%	37.50%
Jefferson, NY	21.70%	11.43%	33.33%	10.53%		6.80%	13.33%
Kings, NY	24.87%	26.43%	16.55%	21.47%	10.59%	24.14%	8.64%
Livingston, NY	28.90%	28.00%		16.67%		9.09%	17.54%
Madison, NY	25.82%	20.00%	23.81%	7.14%	33.33%	0.00%	20.00%
Monroe, NY	25.37%	23.93%	25.77%	14.74%	28.08%	19.95%	17.99%
Montgomery, NY	27.82%	13.19%	40.00%	0.00%		10.81%	14.91%
Nassau, NY	32.88%	32.73%	38.15%	22.76%	35.71%	29.16%	18.62%
New York, NY	23.97%	25.24%	22.37%	20.57%	17.79%	22.83%	23.70%
Niagara, NY	34.07%	24.87%	21.84%	13.79%	0.00%	17.02%	24.86%
Oneida, NY	27.39%	22.87%	31.25%	16.72%		12.33%	9.88%
Onondaga, NY	30.34%	27.70%	33.22%	21.62%	33.33%	16.22%	11.15%
Ontario, NY	26.44%	20.16%	30.00%	9.52%		17.07%	4.00%
Orange, NY	34.78%	30.53%	26.09%	18.75%		22.76%	15.38%
Orleans, NY	26.94%	31.39%	15.38%	50.00%		15.09%	0.00%
Oswego, NY	23.86%	23.40%	0.00%	9.09%		23.92%	32.50%
Otsego, NY	17.21%	15.82%	12.90%	10.71%		5.73%	15.79%
Putnam, NY	32.44%	32.48%	0.00%	39.29%		29.67%	14.75%
Queens, NY	35.13%	27.73%	26.09%	25.17%	22.90%	26.14%	18.98%
Rensselaer, NY	33.67%	23.36%	26.67%	17.65%		17.98%	23.01%
Richmond, NY	30.31%	20.43%	28.57%	30.21%		19.62%	3.04%
Rockland, NY	35.60%	40.17%	46.15%	42.18%		45.52%	3.15%
Saratoga, NY	28.68%	30.54%	23.08%	14.56%		17.88%	19.44%
Schenectady, NY	27.17%	24.33%	10.00%	13.64%		17.38%	22.95%

Schoharie, NY	30.29%	20.00%		50.00%		0.00%	
Schuyler, NY	16.82%	100.00%	0.00%			0.00%	19.53%
St0 Lawrence, NY	26.03%	27.21%	20.00%	25.00%		16.41%	23.81%
Steuben, NY	27.86%	21.69%	11.11%	14.29%		7.48%	8.89%
Suffolk, NY	35.68%	32.77%	27.41%	23.43%	0.00%	23.46%	38.14%
Sullivan, NY	29.74%	24.97%	0.00%	36.00%		16.22%	11.11%
Tompkins, NY	26.35%	27.15%	0.00%	9.17%		11.22%	6.88%
Ulster, NY	35.49%	29.32%	38.89%	21.35%		20.08%	14.29%
Warren, NY	31.66%	31.68%	40.00%	22.22%		25.00%	37.50%
Wayne, NY	27.46%	18.05%	0.00%	30.77%		16.78%	14.84%
Westchester, NY	31.97%	27.63%	24.21%	22.99%	36.90%	24.39%	16.12%
Wyoming, NY	25.78%	36.92%	0.00%	33.33%		13.79%	0.00%
Yates, NY	23.13%	8.00%				0.00%	0.00%

4) Sample codes

a. Combine the data sets

Combine the core dataset with hospital county info, vt_sidc_2007_ahal dataset has county zip code stored in it.

```
data temp;
merge vt_sidc_2007_core
        vt_sidc_2007_ahal;
by HOSPID;
```

```
run;
```

Assign county name by zip code. The data can be obtained from Census Bureau. We saved the data in a SAS data file called "county", and county name is stored in variable "county_name".

```
data NewYorktemp;
merge temp
        county;
by hospzip;
run;
```

b. Select "utilization flag" for CT (U_CTSCAN) from vt_sidc_2007_dx_pr_grps dataset.

```
data NewYork;
merge vt_sidc_2007_dx_pr_grps (keep=key U_CTSCAN)
         NewYorktemp(keep=key age female race_x pay1 hospid MDC  county_name);
by key;
if age<8 and age>=0 then agegroup=1;
  else if age>=8 and age<=18 then agegroup=2;
  else if age>18 and age<30 then agegroup=3;
  else if age>=30 and age<65 then agegroup=4;
  else if age>=65 then agegroup=5;
```

```
  run;
```

c. Labeling category

```
libname MDC 'C:\Documents and Settings\...\My Documents\My SAS Files\9.2';
proc format lib=MDC;
   value mdc
        0='pre-MDC'
        1='Nervous System'
        2='Eye'
        3='Ear, Nose, Mouth and Throat'
        4='Respiratory System'
        5='Circulatory System'
        6='Degestive System'
        7='Hepatobiliary System and Pancreas'
        8='Musculoskeletal System and Connective Tissue'
        9='Skin, Subcutaneous Tissue and Breast'
        10='Endocrine, Nutritional and Metabolic System'
        11='Kidney and Urinary Tract'
```

```
12='Male Reproductive System'
13='Female Reproductive System'
14='Pregnancy, Childbirth and Puerperium'
15='Newborn and Other Neonates (Perinatal Period)'
16='Blood and Blood Forming Organs and Immunological Disorders'
17='Myeloproliferative DDs (Poorly Differentiated Neoplasms)'
18='Infectious and Parasitic DDs'
19='Mental Diseases and Disorders'
20='Alcohol/Drug Use or Induced Mental Disorders'
21='Injuries, Poison and Toxic Effect of Drugs'
22='Burns'
23='Factors Influencing Health Status'
24='Multiple Significant Trauma'
25='Human Immunodificiency Virus Infection';
run;
```

Other category variables are labeled the same way.

Now we have the basic New York State data for further analysis.

d. Create New York CT and New York Inpatient data

Create a binary variable to record if the patient had CT scan(s) or not, we use variable "tag" to identify it.

```
proc sql;
 alter table NewYork
 add tag;


update NewYork
set tag=1 where U_CTSCAN in (1, 2,3);


update NewYork
set tag=0 where U_CTSCAN=0;
```

e. Create summary statistics, P includes age group, gender, MDC, method of payment

```
%let variable=P
ods html body='E:\ 'County Level";
proc tabulate data=NewYorkct;
 class county_name &P;
```

```
 var key;
 table county_name,
       key*&P* (n)/ Box='county name';
run;
ods html close;
```

f. Logistic regression

```
ods html body='C:\Documents and Settings\...\logistic';
proc logistic data=NewYork descending;
 class county_name (ref='Chemung, NY') agegroup (ref='4')
       race_x (ref='01') pay1 (ref='3') female (ref='1') MDC (ref='1')/param=ref;
 model CT=county_name agegroup race_x pay1 female MDC;
run;
ods html close;
```