A Predictive Model Analyses of Medicare and Medicaid Inpatient Stays and the Role of Recovery Audit Contracting Program (RAC)

By

Kathleen M. Kirk

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Biomedical Informatics

Department of Health Informatics

School of Health Professions

Rutgers, the State University of New Jersey

October 2017

RUTGERS
School of Health Professions

Final Dissertation Defense Approval Form

A Predictive Model Analyses of Medicare and Medicaid Inpatient Stays

and the Role of Recovery Audit Contracting Program (RAC)

BY

Kathleen M. Kirk

Dissertation Committee:

Shankar Srinivasan PhD

Barbara Gladson PT, OT, PhD

Frederick Coffman PhD

Approved by the Dissertation Committee:

_____ Date: _____

_____ Date: _____

_____ Date: _____

_____ Date: _____

_____ Date: _____

TABLE OF CONTENTS

ABSTRACT

Introduction: Advances in information acquisition techniques and the widespread use of information technologies in healthcare services has resulted in an incredible opportunity for health administrators to utilize data analytics and models to provide better health care, manage risks and improve patient outcomes. In this study two different data analytical models - a Logic Model and a Predictive Model – were formulated using datasets obtained from the Health Cost Utilization Project (HCUP) and the Recovery Audit Contractor (RAC) reports for use in Medicare and Medicaid patient hospitalization outcomes research.

Objectives: The overall goal of the study was to (1) to design an appropriate analytical model to explain the operations of the RAC process and identify the hospitalization factors that affect the efficient recovery of claims (2) to formulate a predictive model by using HCUP's Nationwide Inpatient Sample datasets to help predict those hospitalization factors above affecting the RAC claims recovery process, and (3) to determine other relevant hospital, regional and patient related variables that play a statistically significant role in both the RAC and the Hospitalization Outcomes Models.

Methods: To meet the aforementioned objectives data was extracted from both the RACTrac Website and Reports (for developing the RAC Process Model) and the HCUP Nationwide Inpatient Sample (NIS) database. Several analytical models currently in vogue in both health and finance were investigated and it was decided to adopt a Logic Model to describe the RAC claims recovery process and with its help identified the hospitalization factors related to the claims and payment issues. Secondly the Multiple Linear Regression Model was found to be the most suitable predictive model type for the

hospitalization factors identified from the RAC Logic Model. Lastly several descriptive and inferential statistics were employed to infer relationships among several patient and hospital variables with the RAC regions and their outcomes.

Results: Both Length of Stay (LOS) and Total Charges were found to be intimately related to the RAC claims recovery process and accordingly they both were employed in the development of the Multiple Linear Regression Model with several independent variables such as DRG, RAC region, Payer type (Medicare, Medicaid, Private), Number of Diagnoses and Number of Procedures resulted in a reasonably good fit (54 % to 59 %) of the model in explaining the variance of the outcome of Total Charges and not a very good fit for the LOS which was expected since LOS is not a linear variable and subject to too many constraints and hence not easily predictable. The ANOVA Tests revealed several interesting relationships between the independent variables listed above and the RAC regions with implications of import for the RAC claims recovery process.

Conclusion: This study is significant because it demonstrates the validity of the use of analytical models such as Logic Model and the Multiple Linear Regression Model in predicting Hospitalization Outcomes of interest to not only the RAC claims recovery process relevant to this study but also in other health administrative settings involving planning of budget and resource allocation. The complex process of RAC claims recovery mechanism has been duly modeled by the Logic Model technique thus making it available for future configuration modification and studies into their effect on the claims recovery process.

ACKNOWLEGEMENT

I would like to gratefully acknowledge many people who have journeyed with me in recent years as I have completed my dissertation.

I would like to express my sincere gratitude to Rutgers Biomedical Health Sciences - School of Health Professions for allowing me to fulfill my goal of pursuing higher education.

I would also like to thank the Department of Health Informatics for giving me the opportunity to complete my dissertation. I would like to express my special appreciation and thanks to my Committee Chair, Dr. Shankar Srinivasan. I am extremely grateful to my committee, Associate Dean, Dr. Barbara Gladson and Dr. Frederick Coffman, for their assistance and suggestions on my dissertation.

I would like to thank the people who have been instrumental in the successful completion of this dissertation. I extend my heartfelt thanks to the faculty in the Department of Health Informatics who offered collegial guidance and support over the years: Prof. Barbara Manger, Dr. Masayuki Shibata, Dr. Riddhi Vyas and Prof. Memory Ndanga.

Mary T. Siegert deserves a special mention for her unremitting encouragement and advice.

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

1.1 Background of Problem

In Section 306 of the Medicare Prescription Drug, Improvement, and Modernization Act 2003 (MMA), Congress directed the Department of Health and Human Services to conduct a 3-year demonstration using Recovery Audit Contractors (RAC) to detect and correct improper payments in the Medicare Fee-For-Service program[1]. The Recovery Audit demonstration was conducted from March 2005 to March 2008, in six states which had the highest Medicare expenditures. Those states were: California, Florida, New York, Arizona, Massachusetts, and South Carolina. The RAC demonstration program was to determine if Recovery Auditors could successfully be used to identify improper payments for claims paid under Medicare Part A and Part B. The RAC demonstration was a significant means in assisting preparation for and forming the RAC permanent program.

The Tax Relief and Health Care Act of 2006 (TRHCA) made the RAC program permanent, authorizing expansion of the program to all 50 states by June 10, 2010. The program's mission is to "reduce Medicare improper payments through the efficient detection and collection of overpayments, the identification of underpayments, and the implementation of actions that will prevent future improper payments."

CMS designated four RAC contractors, who each have responsibility for approximately one-quarter of the country. The RAC regions are as follows:

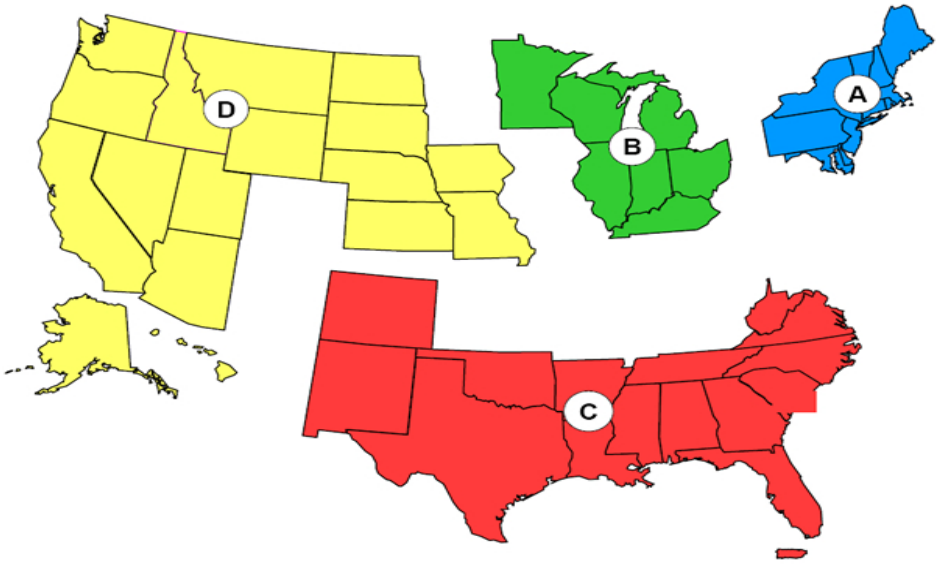


Figure 1 – Revenue Audit Contracting Program (RAC) Regions A, B, C and D (details in Table 1)

Table 1: RAC Regions and Designated States

RAC	STATES
Region A: (Region 1) Performant Recovery (formerly Diversified Collection Services)	CT, DE, DC, ME, MD, MA, NH, NJ, NY, PA, RI and VT.
Region B: (Region 2) CGI Federal, Inc.	IL, IN, KY, MI, MN, OH and WI.
Region C: (Region 3) Cotiviti, formerly Connolly, Inc.	AL, AR, CO, FL, GA, LA, MS, NM, NC, OK, SC, TN, TX, VA, WV, Puerto Rico and U.S. Virgin Islands.
Region D: (Region 4) HealthDataInsights	AK, AZ, CA, HI, ID, IA, KS, MO, MT, ND, NE, NV, OR, SD, UT, WA, WY, Guam, American Samoa and Northern Marianas.

Each of these RAC are responsible for identifying overpayments and underpayments in their specified geographical areas of the country. RAC is also responsible for reporting to CMS any common billing errors, trends, or other Medicare payment concerns.

At the end of 2015, Centers for Medicare & Medicaid Services (CMS) announced the awarding of the Region 5 RAC to Connolly, LLC. This Region is exclusively devoted to

the review of DMEPOS (durable medical equipment, prosthetic and orthotic suppliers) supplies, home health agencies and hospice providers. In 2015, Connolly, LLC was the "old" Region C RAC and was the first RAC to emerge when the permanent program started, posting its first automated issues in June 2009[2].

While the implementation of Region 5 RAC contract was expected to roll out in the Spring of 2014, delays were encountered due to pre-award protests involving Regions 1, 2, and 4. Also at that time, there was a lawsuit pending in the United States Court of Federal Claims by a RAC contractor, CGI Federal against CMS.

RAC auditors are paid on a contingency basis. RACs are paid a percentage of every dollar, ranging from 9% up to 12.5% in overpayments collected. The percentage varies among the RACs depending on their specific contract, whether the claim errors are overpayments to be recouped or underpayments to be returned to providers. This sounds reasonable; however, there is an innate pressure on RACs to bring money in. Underpayment totals have been far less than overpayments.

The RAC identify over and underpayments by reviewing the supporting medical records or through an automated analysis of certain claims. Table 2 represents the total corrections in FY 2010 by each RAC.

Table 2 – FY 2010 – Total Corrections in Millions

REGION	RAC	AMOUNT CORRECTED (Million)
Region A	Diversified Collection Services (DCS)	$ 5.9
Region B	CGI, Inc.	$ 15.5
Region C	Connolly, Inc.	$ 27.5
Region D	HealthData Insights (HDI)	$ 43.4
		TOTAL: $ 92.3

Section 6411 (b) of the Affordable Care Act (ACA) expanded the use of recovery audit contractors (RAC) to Medicare Parts C and D. CMS has initiated implementation of Part C and Part D RACs. A contract for Part D recovery auditing was awarded on January 13, 2011 to ACLR Strategic Business Solutions (ACLR's). ACLR's initial review focused on identifying improper payments for prescriptions written by excluded prescribers or filled by excluded pharmacies. Recoupment began in the first quarter of FY 2013 for those plans that did not appeal findings identified in the RAC's initial audit review. In addition to the Part D RAC procurement activity, CMS solicited comments on how best

to implement the Medicare Part C recovery auditing program through a Request for Information (RFI) that was published in the Federal Register on December 27, 2010[3].

Section 6411(a) of the Affordable Care Act amended section 1902(a)-(42) of the Social Security Act to require that States and its territories establish RAC programs.

States must contract with one or more RACs in their Medicaid programs. They are expected to administer their Medicaid RAC programs within the Federal regulatory structure established by CMS. CMS published a Notice of Proposed Rule Making (NPRM) for the establishment of Medicaid RACs on November 10, 2010. The Final rule was published on September 16, 2011 and it required states to implement their Medicaid RAC programs by January 1, 20I 2. At the conclusion of FY 2011, states made progress in implementing their Medicaid RAC programs, including several states that have Medicaid RAC contracts in place and many others that were in the procurement process.

1.2 Statement of the problem

The literature review revealed that there have not been any studies researching any associations between RAC regions and cost, length of stay, diagnoses, procedures or total charges. This study will explore the various issues that have overwhelmed the RAC program since its demonstration program began in 2005. There are several issues in particular that have delayed the RAC program from recouping reimbursement, thus impacting its recoveries to the Medicare Trust Fund. The RAC program has been monetarily beneficial in the past. For example, in FY 2013, after deducting administrative

costs and underpayments, the RAC program returned $3.03 billion to CMS for improper payments or underpayments.

1.2.1 Two Midnight Rule

However, in FY 2014, RAC dropped to 1.6 billion[4]. This decrease occurred because CMS banned its RAC contractors from reviewing hospital inpatient status in relation to the Two Midnight Rule implementation. This ban lasted for a period of 18 months. The two-midnight rule directs CMS payment contractors to presume hospital stays are appropriately billed as inpatient admissions rather than outpatient observation visits if they span two midnights[5].

On July 15, 2013, the US Department of Health and Human Services (HHS) temporarily suspended RAC audit appeals in order to allow the Recovery Auditors to decrease their backlog of approximately 357,000 claims[6].

CMS's challenges in setting expectations about the work the Part D RAC would conduct and establishing the length of time required for CMS and the RAC to reach project milestones hampered Part D RAC program implementation[7]. RAC's scope of work was unclear because CMS incorporated the terms of work set out in the performance work statement into the RAC's contract without making any changes to the performance work statement.

Medicare's recovery audit contractor program faces additional havoc as the U.S. Court of Appeals for the Federal Circuit issued a decision that stated a lower federal claims court must reconsider how CMS procures its new contracts for the RAC program.

1.2.2 RAC and American Hospital Association

There has been a growing mistrust of the RAC program across the country by hospital administrators. From the hospital's perspective, they attribute RACs' contingency-based fee structure for RACs' overzealous claims denial. RAC Auditors receive commissions on every Medicare payment the auditors deny. RACs have concentrated the bulk of their audits on inpatient hospital claims which are usually the highest reimbursement for hospitals. This incentive has led to a high volume of denied claims and reimbursement recoupment from hospitals for claims that initially should not have been denied.

Although hospital administrators understand the need for auditors to identify billing mistakes; however, responding to the increasing number of audits and challenging inappropriate denials drains hospitals' time, funding and attention that could more effectively be focused on patient care. Hospitals are facing an increase in audits of payment claims by recovery audit contractors (RACs), which subject hospitals to additional administrative burden and costly payment denials.

The American Hospital Association's (AHA) RACTrac survey collects data from hospitals on a quarterly basis to assess the impact the Medicare Recovery Audit Contractor (RAC) program on hospitals nationwide. AHA developed RACTrac in

response to the lack of data and information provided by CMS on the impact of the RAC program on providers[8]. Hospitals are reporting widespread RAC process issues via the RACTrac survey tool by issue. The data collected from the survey indicates several key issues. One issue being the two-midnight rule. The American Hospital Association (AHA) has stated that the policy is unclear and undermines the medical judgment of physicians. RAC have based their audit reviews of the physician's expectation of medical necessary care covering two or more midnights on the information available to the admitting physician at the time. In June 2016, the AHA reported that 27% of the RAC was not meeting the 60-day deadline to make a determination on a claim delaying the appeal process for hospital claims review.

There are legislators who contend that the RAC program requires an overhaul. On May, 1, 2015, the H.R. 2156, the Medicare Audit Improvement Act of 2015 was introduced by Representatives Sam Graves, a Republican from Missouri and Adam Schiff, a Democrat from California. This bipartisan legislation addressed some of the issues that made the RAC program inefficient.

On June 28, 2016, the Department of Health and Human Services issued a proposed rule that would make changes to the procedures for Administrative Law Judge (ALJ) appeals of payment and coverage determinations for items and services provided to Medicare beneficiaries, in addition to other Medicare appeals[9]. The proposed rule was issued just days before HHS's July 1, 2016 deadline to respond in court to show progress in

resolving the backlog of ALJ hearings as part of the American Hospital Association's lawsuit challenging the significant delays in ALJ hearings[10].

1.2.3 Medicare Trust Fund

The American public has been hearing for years and perhaps decades that the Medicare Trust Fund may be on the brink of insolvency. Since most of the federal budget goes toward Defense, Social Security. and major Health Program (i.e, Medicare & Medicaid Programs) it would be in the best interest of Americans to protect Medicare & Medicaid's future. (see Figure 2) CMS' Medicare and Medicaid Services collectively have a total of 33% of a share of the federal budget in 2015. The RAC program following a successful pilot program yielded approximately $900 million in returns to the Medicare Trust Fund. Yet another reason to study the future viability of the RAC program and in this regard this research study aims to include the hospitalization outcomes affecting the RAC process and the ability to develop a predictive model using the two elements.

As mentioned previously, there has not been a study that attempts to associate costs, length of stay, diagnoses, procedures or total costs to the four RAC Regions as possible factors hindering the viability of the Center for Medicare and Medicaid Services RAC program.

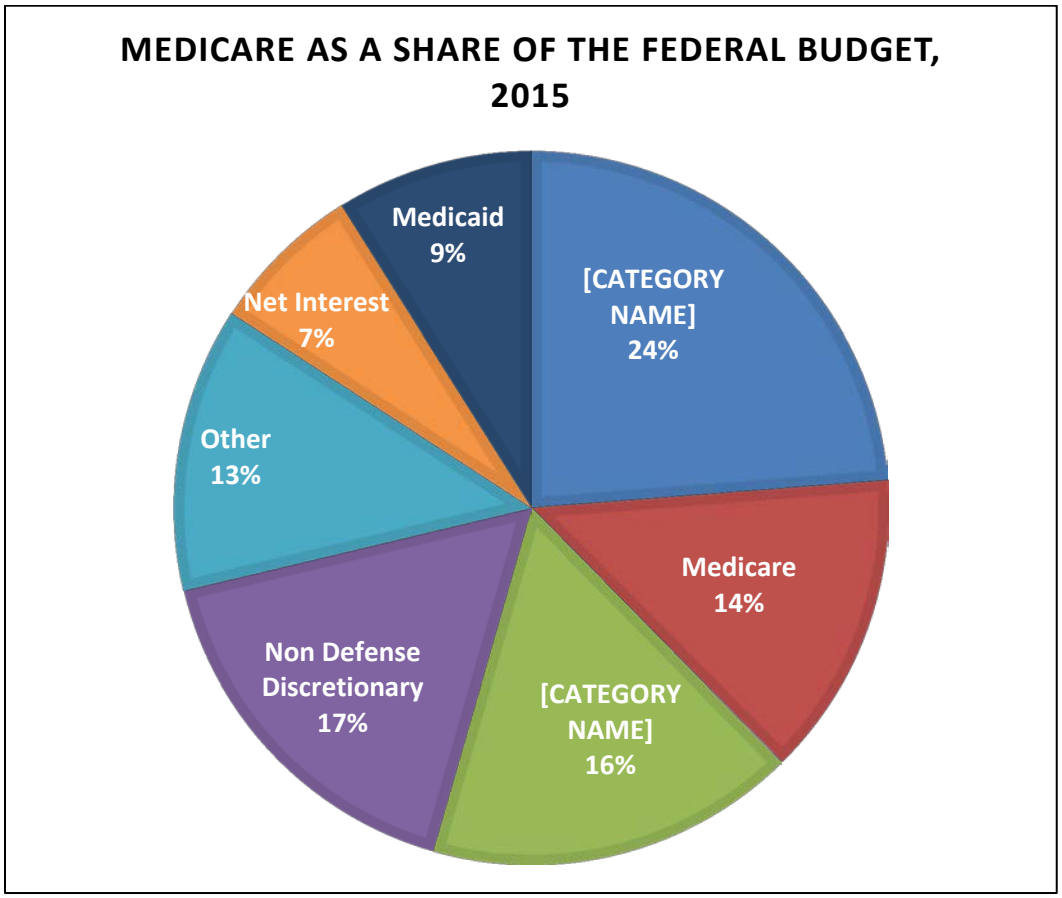


Figure 2: Medicare as a Share of the Federal Budget - 2015

There has been recently a lot of interest at the Governmental and Private Sector levels to introduce predictive analytics into healthcare. Indeed Section 4241 of the Small Business Jobs Act of 2010 authorizes the Secretary of the Department of Health and Human Services ("HHS") to use predictive modeling and other analytics technologies to identify improper claims for reimbursement and to prevent the payment of such claims under the Medicare fee-for-service program. Even though the goal of this statute was to identify fraud in billing yet the potential use of analytics lies beyond just fraud detection alone.

Indeed it can provide health administrators to utilize data analytics and predictive models to provide better health care, manage risks and improve patient outcomes. In this dissertation it was planned to use predictive and data analytics to evaluate the role and impact of certain Hospitalization Outcomes available in the Health Cost Utilization Project (HCUP) and identified in the RAC process as being significant variables affecting the RAC's performance and findings.

Accordingly some specific objectives for further research were identified and are as listed below.

1.3 Research Objectives

Specifically, the objectives were:

- to design an appropriate analytical model to explain the operations of the RAC process and identify the hospitalization factors that affect the efficient recovery of claims

- to formulate a predictive model by using HCUP's Nationwide Inpatient Sample datasets to help predict those hospitalization factors above affecting the RAC claims recovery process, and

- to determine other relevant hospital, regional and patient related variables that play a statistically significant role in both the RAC and the Hospitalization Outcomes Models.

1.4 Need and Rationale

The RAC program has been monetarily beneficial in the past but has since been overwhelmed with issues that have delayed the RAC program from recouping reimbursement, thus impacting its recoveries to the Medicare Trust Fund.

A survey of the literature reveals that there are not any studies on the RAC program that explore the possibility that there may be other factors in the four RAC regions that could impact on the future viability of the program other than those previously discussed in this chapter. Failure to explore factors such as costs associated with Medicare and Medicaid patients in terms of length of stay, DRGs, total charges, number of procedures, number of diagnoses, length of stay and race could adversely affect the financial outcome of the RAC program. These other factors could play a pivotal role in determining how to proceed in the development of new policies regarding the development of the four RAC regions and impact recoveries to the Medicare Trust Fund.

1.5 Research Hypotheses

The aim of this study is to explore factors such as costs associated with Medicare and Medicaid patients in terms of length of stay, DRGs, total charges, number of procedures, number of diagnoses, length of stay and race could adversely affect the financial outcome of the RAC program. This study will be determined by the following research hypotheses.

- **Hypothesis 1**: It is possible to derive an appropriate analytical model to explain the operations of the RAC process and identify the hospitalization factors that affect the efficient recovery of claims

- **Hypothesis 2**: It is possible to formulate a predictive model by using HCUP's Nationwide Inpatient Sample datasets to help predict those hospitalization factors above affecting the RAC claims recovery process

- **Hypothesis 3**: There are relevant hospital, regional and patient related variables that play a statistically significant role in both the RAC and the Hospitalization Outcomes Models.

CHAPTER II

LITERATURE REVIEW

2.1 Literature Search and Search Strings

The literature search consisted of a review of various articles published on the Center for Medicare and Medicaid Revenue Audit Recovery Program. Articles were searched in the PubMed database which includes Medline. Google searches were also utilized. A good portion of the articles reviewed were published on healthcare, government or law websites. The total number of articles and/or abstracts reviewed were 587. Of the 587, approximately 157 were reviewed in detail.

The following are several of the search strings used to locate articles from the databases: Search terms used were:

“Medicare” and “Hospital Length of Stay” and “Hospital Costs”

“Medicare and Medicaid” and “ICD-9-CM”

“Revenue Audit Program” and “Medicare Trust Fund”

“ CMS” and “Total Costs” and Medicare’s Trustee Report”

“RAC” and “American Hospital Association”

“Medicare & Medicaid” and “Number of Diagnoses” and “Number of Procedures”

2.2 Medicare

Medicare is the nation's health insurance program for most people age 65 and older and certain disabled individuals[11]. The majority of beneficiaries, nearly 75%, receive benefits through Medicare's fee-for-service (FFS) program, known as "original" or "traditional" Medicare. The remaining beneficiaries, approximately 25%, chose to enroll in private health care plans under Medicare Part C, the Medicare Advantage (MA) program. Approximately 73% (34.6 million beneficiaries) of Medicare beneficiaries chose to enroll in Part D, the outpatient prescription drug program[12].

2.2.1 Medicare Plans

There are various types of Medicare plans, which are divided into four distinct "parts". Medicare Parts A, B, C and D are the most common Medicare programs.

Medicare Part A and Medicare Part B are referred to as Original Medicare. Original Medicare is managed by the federal government. In general, Original Medicare provides Medicare eligible individuals with coverage for and access to physicians, hospital care (inpatient), limited home health services, skilled nursing facility care and hospice care.

The Centers for Medicare & Medicaid Services (CMS) is the federal agency that operates Medicare. CMS is part of the U.S. Department of Health and Human Services. Medicare is financed by a portion of the payroll taxes paid by workers and their employers, general revenues and beneficiary premiums.

Beneficiaries are automatically eligible for Medicare Part A at age 65. Some beneficiaries may qualify for Medicare A prior to reaching the age of 65, if they have a disability, end-stage renal disease, or amyotrophic lateral sclerosis. In order to qualify for Medicare Part A, the individuals must be either a United States citizen or a permanent resident of at least 5 continuous years. In general Medicare Part A coverage includes: hospital care (inpatient stays), limited home health services, skilled nursing facility care and hospice care. Individuals who had paid Social Security and Medicare payroll taxes for approximately 10 years any additional eligibility requirements, would not have premiums for Medicare Part A coverage. If individuals do not have sufficient payroll tax history, they may qualify for no premium Part A, based on their spouse's work history. There is an option to buy Medicare Part A with a monthly premium.

Medicare Part B is considered health insurance coverage. Most individuals pay a premium for Part B regardless of work history. Medicare Part B assists in the coverage of medically necessary services and supplies required for the diagnosis or treatment of an individual's health condition. This includes outpatient services received at a hospital, physician's office, clinic, or other health facility. Medicare Part B also assists in the coverage of various preventive services to prevent illness or detect illness at an early stage.

Medicare Part C is also referred to as Medicare Advantage. Medicare Advantage is a Medicare-approved private health insurance plan for individuals enrolled in Medicare Part A and Part B. Medicare pays private companies to cover the individual's Medicare

benefits. Medicare Advantage offers additional benefits, such as, vision, dental, hearing and several include prescription drug coverage. In addition to the Part B premium, the individual is often required to pay a monthly premium for the services based on the level of coverage selected. Each Medicare Advantage Plan has different premiums and costs for services.

2.2.2 Expansion of the RAC Program

Section 6411(b) of the Patient Protection and Affordable Care Act of 2010 (PPACA) (P.L. 111- 148), as amended by the Health Care and Education Reconciliation Act of 2010 (HCERA) (P.L 111-152), requires expansion of the RAC program to Medicare Part C. Amendments to the existing Fee-For-Service (FFS) RAC statute at §1893(h) provide CMS with general authority to enter into contracts with RACs to identify overpayments and underpayments and recoup overpayments in Medicare Part C[13]. Under the RAC contract, payment is produced to the RAC only from amounts recovered, and payment is made to the RAC on a contingent basis for collecting overpayments and in an amount determined by the Secretary for identifying underpayments.

Medicare beneficiaries are not automatically enrolled for prescription drug coverage. Medicare Part D coverage is optional. There are two options to obtain Medicare prescription drug coverage. Both of these options are offered through private insurance companies that commission with Medicare.

Option 1: Through a stand-alone Medicare Part D prescription drug plan which can be added to the Original Medicare coverage. You can enroll in any Part D prescription drug plan that serves the area where you live.

Option 2: Through a Medicare Advantage Prescription Drug plan. The beneficiary selects a Medicare Advantage (Part C) plan that includes drug coverage. This option allows the beneficiary to receive all their Medicare benefits under one plan.

The Centers for Medicare & Medicaid Services (CMS), a branch of the Department of Health and Human Services (HHS), is the federal agency that operates the Medicare Program and monitors Medicaid programs offered by each state.

In 2011, Medicare alone covered 48.7 million people. Total expenditures in 2011 were $549.1 billion. This money originates from the Medicare Trust Funds.

2.3 Medicaid

Medicaid is a joint federal and state program that together with the Children's Health Insurance Program assists with medical costs for individuals with limited income and resources which also includes children, pregnant women, parents, seniors and individuals with disabilities. Medicaid provides coverage to over 72.5 million Americans[13]. Each state has different rules about eligibility and also differs in the application process. Medicaid is the single largest source of health coverage in the United States[14].

2.4 Fraud and Abuse and Improper Payment

The Centers for Medicare & Medicaid Services' (CMS) Medicare Integrity Program (MIP) is designed to identify and address fraud, waste, and abuse which attribute to causes of improper payments. In Medicare, program integrity typically encompasses two types of activities: (1) processes directed at reducing abuse, such as payment errors or improper payments and (2) activities designed to prevent, detect, investigate, and ultimately prosecute fraud. Since 1990, the Government Accountability Office (GAO) has designated Medicare as a federal program at high risk for fraud and abuse due to its size, complexity, scope, and decentralized administrative structure[15].

In order to protect the Medicare Trust Funds from improper payments, CMS contracts with private companies to review claims to determine whether the services provided are medically reasonable and necessary. In Medicare, improper payments include both provider under- and overpayments. The majority of claims are screened and reviewed after payment has been made or post-payment.

2.5 Medicare and Medicaid Contractors

To conduct Medicare program integrity activities, CMS contracts with a several different contractors. Activities undertaken by these contractors varies depending on their Statements of Work (SOW). Several process and pay Medicare claims in addition to performing distinct program integrity functions (i.e., Medicare Administrative Contractors or MACs). Others specialize completely in program integrity activities such

as Program Safeguard Contractors (PSCs) and Zone Program Integrity Contractors (ZPICs), Medicare Drug Integrity Contractors (MEDICs), Recovery Audit Contractors (RACs), the Comprehensive Error Rate Testing (CERT) contractor, the National Supplier Clearinghouse (NSC), and the Coordination of Benefits (COB) contractor. In this study, the Medicare Administrative Contractors (MACs), the Zone Program Integrity Contractors (ZPICs), the Recovery Audit contractors (RACs) and the Medicaid Integrity Contractors (MICs) will be discussed due to type of recovery audit they conduct.

The current audit landscape includes:

- Medicare Administrative Contractors (MACs)
- Zone Program Integrity Contractors (ZPICs)
- Recovery Audit Contractors (RACs);
    - Medicare RACs
    - Medicaid RACs
- Medicaid Integrity Contractors (MICs)
    - Comprehensive Medicaid Integrity Plan (CMPI)

Each of the above mentioned Medicare contractors have various responsibilities for auditing records, claims, and payments. Each type of contractor may use different methods to conduct audits, but they must all abide by the principal Medicare guidelines for medical review, denials, appeals, and payment recovery, as set forth by the Centers for Medicare and Medicaid Services (CMS).

Medicare Administrative Contractors (MAC) are private healthcare insurers who have been awarded a geographical jurisdiction to process Medicare Part A and Medicare Part B or Durable Medical Equipment (DME) claims for Medicare Fee-For-Service (FFS) beneficiaries. MACs are multi-state contractors. The MAC scope of work includes using data from other contractors to target improper payment. MACs have the ability to perform medical reviews for all claims, at their discretion, and will do so by issuing an additional documentation request (ADR) to the provider[16].

Section 911 of the Medicare Prescription Drug Improvement, and Modernization Act (MMA) of 2003 directed CMS to replace the Part A Fiscal Intermediaries (FIs) and Part B carriers with MACs. CMS procures all MAC contracts according to the Federal Acquisition Regulation.

Zone Program Integrity Contractors (ZPICs) identify and stop potential fraud and refer these cases to the Department of Health and Human Services (HHS) Office of Inspector General (OIG) Office of Investigation. The ZPICs perform functions to ensure the integrity of the Medicare Program. Generally, most MACs will interact with one ZPIC to handle fraud and abuse issues within their jurisdictions.

Recovery Audit Contractors (RACs) detect and correct improper payments in order for CMS and Carriers, Fiscal Intermediaries (FI), and MACs can implement actions that will prevent future improper payments. CMS historically contracts with Recovery Audit Contractors (RACs) to identify and correct overpayments and underpayments in Medicare Parts A, B, and D. On December 22, 2015, CMS released a Request for

Information (RFI) for CMS to enter into a contract with one or more Recovery Auditors to identify and correct overpayments and underpayments in Medicare Part C. The RFI sought to develop the Recovery Audit Program to include the identification and correction of overpayments and underpayments associated with diagnosis data submitted to CMS by Medicare Advantage Organizations for Part C payment[17].

2.6 Medicaid Integrity Program

Under the provisions of the Deficit Reduction Act (DRA) of 2005, Congress directed CMS to establish the Medicaid Integrity Program (MIP). Audit Medicaid Integrity Contractors (Audit MICs) are entities with which CMS has contracted to conduct post-payment audits of Medicaid providers. The general goal of the provider audits is to identify overpayments and to eventually decrease the payment of inappropriate Medicaid claims. At the direction of CMS, the Audit MICs audit Medicaid providers throughout the country. The audits ensure that Medicaid payments are for covered services that were actually provided and properly billed and documented. MICs perform field audits and desk audits. Any Medicaid provider will be subject to an audit by MIC, including fee-for-service providers, institutional and non- institutional, as well as managed care entities.

Umbrella contracts have been awarded to: Booz Allen Hamilton, Cognosante, IPRO, IntegriGuard, and Health Integrity, LLC. Task orders that have been issued by CMS to the following MICs are listed in Table 3.

Table 3: Medicaid Integrity Contractors listed by Region

Region	MICs
I/II	IPRO
III/IV	Health Integrity
V/VII	Health Integrity
VI/VIII	Health Integrity
IX/X	IntergiGuard

The Government Accountability Office (GAO) has identified the Medicaid program as high risk due to its size ($400 billion annually), growth, diversity of programs, and concerns about the adequacy of fiscal oversight[18].

Section 1936(d) of the Social Security Act directs the Secretary of Health and Human Services (HHS) to establish, on a recurring 5-fiscal year basis, a comprehensive plan for ensuring the integrity of the Medicaid program by combatting fraud, waste, and abuse. The Comprehensive Medicaid Integrity Plan (CMPI) establishes the strategy of the Centers for Medicare & Medicaid Services (CMS) to safeguard the integrity of the

Medicaid program during federal fiscal years 2014–2018. The implementation of the Affordable Care Act over the next five years will result in an expansion of Medicaid enrollment and an increase in the federal investment in the program. The Congressional Budget Office (CBO) projects that over federal fiscal years 2014 – 2018, Medicaid enrollment will expand by a total of 14 million beneficiaries and federal Medicaid spending will increase by a total of $119 billion over five years[19]. The Comprehensive Plan represents CMS' strategy to improve existing program integrity efforts in addition to implementing new initiatives to safeguard expanded coverage and financial investment in the Medicaid program.

2.6.1 Unified Program Integrity Contractor

To improve efficiency and coordination of federal data analysis and audit work within each region, CMS is developing a Unified Program Integrity Contractor (UPIC). According to the CMS' "Comprehensive Medicaid Integrity Plan for Fiscal Years 2014 - 2018" report:

"... CMS is developing a Unified Program Integrity Contractor (UPIC) strategy. Under this strategy, Medicare and Medicaid program integrity audit and investigation work at the federal level will be consolidated into a single contractor within a defined multi-state area, which will complement audit and investigation efforts by states. This contractor will conduct Medicare, Medicaid investigations and audits within designated geographic jurisdictions. In July 2013, CMS released a Request for Information and conducted an Industry Day targeted at gathering information from the vendor community on possible

requirements for combining Medicare and Medicaid program integrity functions. CMS expects to implement the UPIC strategy beginning with initial contract awards in FY 2015 with additional transitions to occur in subsequent fiscal years."[20]

The intent of this program is to decrease the burden placed on providers who are trying to meet all the audit requests from various auditing agencies. Implementation of the UPIC began with contract awards in FY 2015 with additional transitions to occur in subsequent fiscal years.

2.7 Types of RAC Audits and Reasons for Claims Denial

There are two types of RAC Audits: automated and complex review. RACs perform automated reviews when improper payment is obvious.

2.7.1 Automated Review

The automated review uses edits designed to detect claims for evidence of improper coding or other mistakes. For example, the RAC software checks for duplicates and improper coding such as unbundling. CMS will often have lower reimbursement for rates for groups of procedures commonly performed together such as closures incidental to surgery. Unbundling billing codes illegally increases a provider's profits by billing bundled procedures separately, which results in higher reimbursement from Medicare and Medicaid. Based on the automated reviews, RACs send the provider a letter demanding repayment. This type of review occurs when the provider receives a demand letter for

repayment from the RAC. There is no prior review of the medical record prior to receiving the demand letter. Therefore, the recoupment is automatic. The provider has 30 days to dispute the RAC overpayment determination, in what is termed a “rebuttal”. If after 30 days the provider does not successfully dispute the findings, the Fiscal Intermediary/MAC will offset the overpayment. If the claim was originally paid by an insurance carrier, the carrier will adjust the claim and the provider will receive a demand letter and a revised explanation of benefits. The provider has 41 days to repay the overpayment. There is no rebuttal period for a claim identified by a RAC when the claim was originally paid by the carrier. But the provider has the option of contesting the RAC's determination, which the RAC may review and rescind.

The demand letter articulates the amount of each overpayment identified by a RAC the method of calculating the denial; the reason the original payment was correct; regulatory and statutory basis for the denial; the provider’s option to submit a rebuttal statement; the provider’s appeal rights which are separate from the rebuttal process; and the recoupment, payment and interest options for the provider; and the related timelines.

2.7.2 Complex Review

The Complex Review is the actual physical review of the Medical Record or other documentation where by the RAC contractor individually reviews the medical record. The RAC sends a letter requesting medical records. The RAC then has 60 days to review the information and notify the provider in writing if an overpayment is discovered. Under

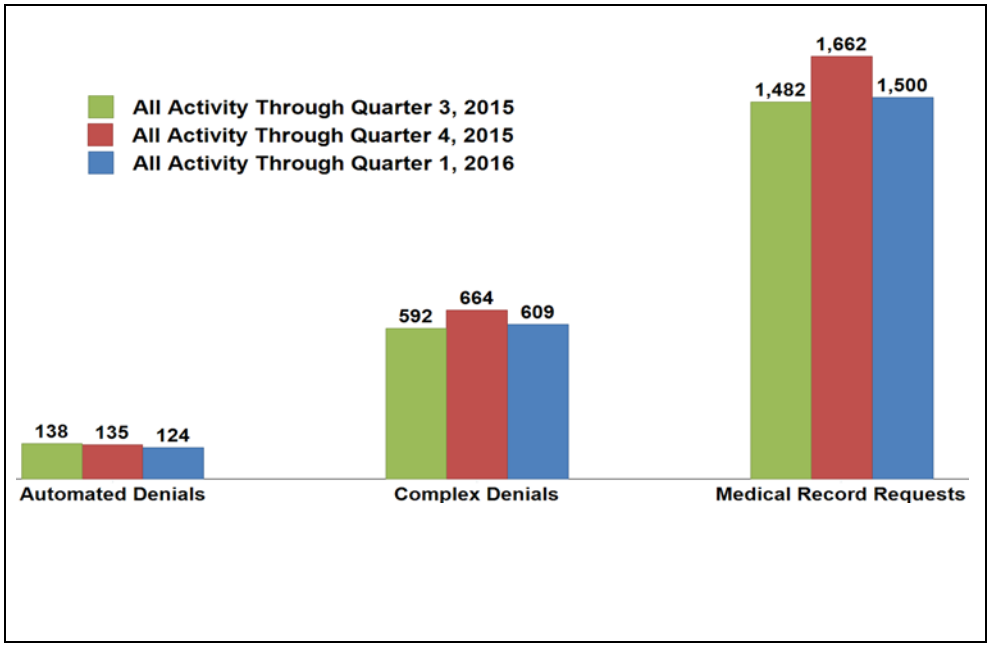
a complex review, the RAC communicates all results directly to the provider including results where no improper payment was identified.

When there has been an improper payment, the RAC must inform the provider of which coverage, coding, payment policy or article was violated. The complex review differs from the automated review because recoupment is the same regardless of whether the claim was paid by the fiscal intermediary or the carrier. The overpayment amount will be offset against each provider's future payments. If a substantial amount is outstanding, the provider may request an extended payment plan. The provider has the right to appeal the RAC's final determination.

The RAC offers the provider a discussion period for all denied claims. During the discussion period, the provider may provide additional information or documentation to the RAC for its consideration. This is not considered part of the formal Medicare Appeals process. Providers must contact the RAC within 15 calendar days of the date of the Demand Letter to begin the discussion period.

The Appeals timeframe is not put on hold for the discussion period and will run simultaneously from the date of the Demand Letter. For example, if a provider requests to stop recoupment, the provider should simultaneously file an appeal with the Carrier (MAC) at the same time the provider is discussing the audit with the RAC.

The RAC is allowed to review claims retroactively up to three years of provider claims. Claims paid prior to 10/1/07 cannot be reviewed by RAC; and going forward from 10/1/10, claims paid more than 3 years prior cannot be reviewed.



Source: AHA RACTrac Survey – June 3, 2016

Figure 3: Average Automated Denials, Complex Denials and Medical Records Requests Per Participating Hospital, through Quarter 1 – 2016

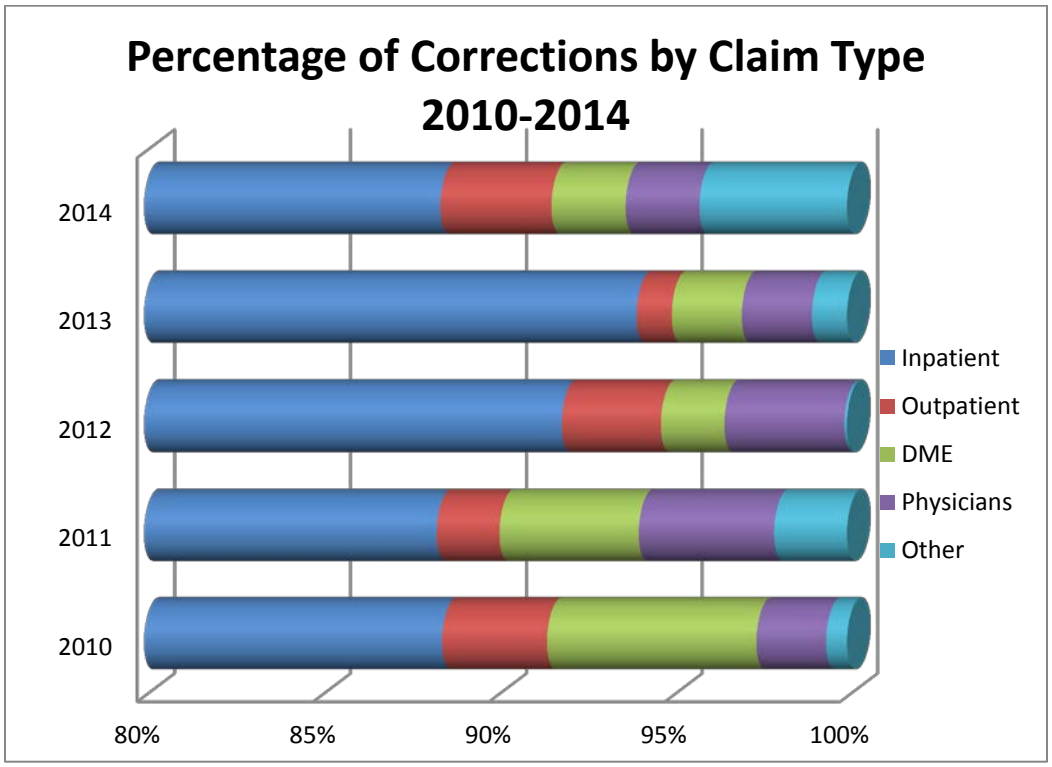


Figure 4: Percentage of Corrections by Claim Type 2010 -2014

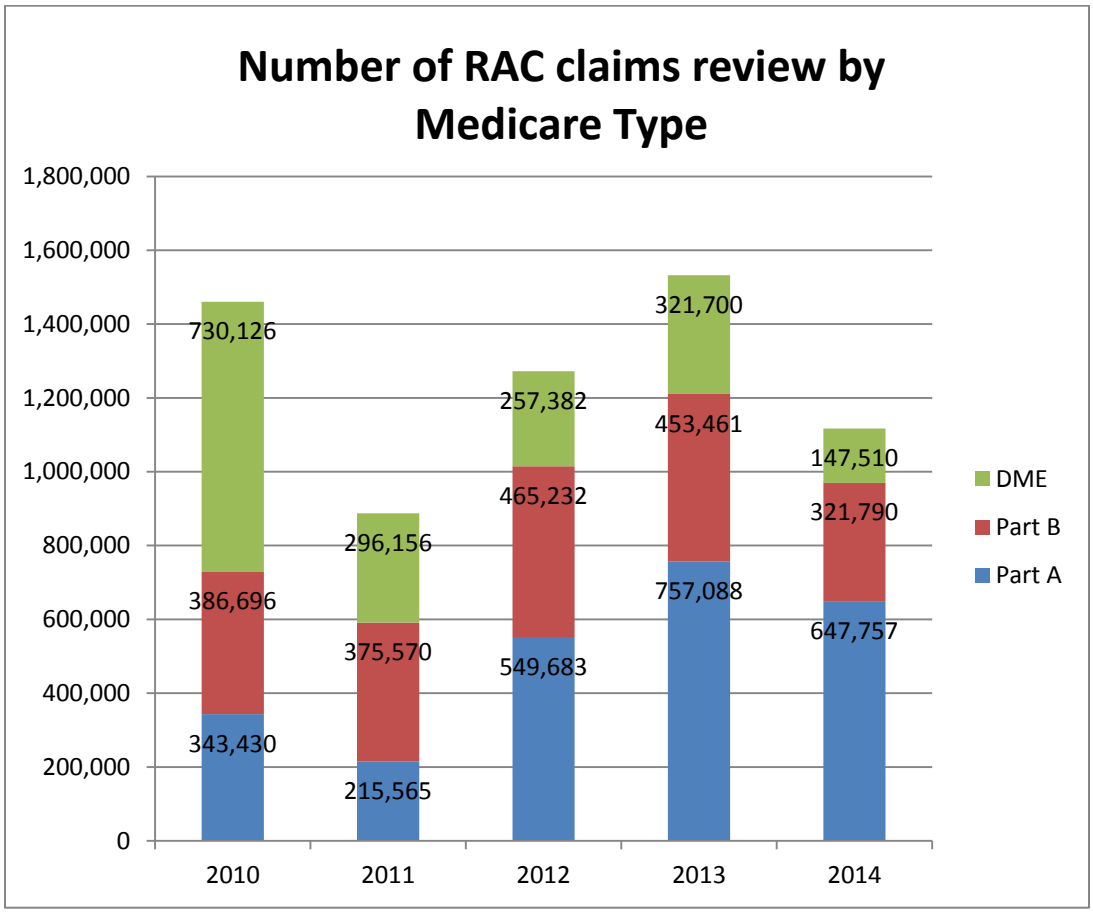


Figure 5: Number of RAC claims review by Medicare Type – 2010 – 2014

2.8 Medicare's Appeal Process

Claims denied payment following a determination made by a RAC contractor can be appealed. Medicare offers a five-level appeal process. In additional to the five-level appeal process, RAC is required to offer an opportunity for the provider to discuss the improper payment determination. This typically occurs soon after the initial RAC letter is received. Then, if the provider determines if an appeal is in order. A formal appeals process begins as shown in Table 4.

Level 1: Redetermination

Redetermination is the examination of a claim by a MAC using personnel who was not associated with the original determination. The appeal must be filed within 120 days from the date of receipt of the initial claim determination. There is no minimum monetary threshold and a decision will be made within 60 days.

Level 2: Reconsideration

Reconsideration can be filed if there is disagreement with the findings of the Level 1 - Redetermination. A Qualified Independent Contractor (QIC) will complete the review. There is no minimum monetary threshold and a decision is typically completed within 60 days of the Reconsideration request.

Level 3: Administrative Law Judge Appeal

Appealing to the ALJ, requires completion of a CMS document form and a request for a hearing. The ALJ hearing may be requested within 60 days of receipt of the Reconsideration. There are several ways these hearings may be held: by telephone, video conferencing, or face to face. CMS can elect to attend the hearing. A decision is usually rendered within 90 days.

Level 4: Appeals Council Review

The request for the Level 4 appeal must be made within 60 days of receipt of the ALJ decision and is required to list he issues contested. A decision from the Appeals Council is typically rendered within 90 days

Level 5: Judicial Review in the US District Count

The request for Level 5 would require a claim in excess of $1260.00. A request for this level of appeal must occur within 60 days of receipt of the decision from the Appeals Council. There is no statutory time limit. Since the RACs and MACs and Z-PICs have expanded their audits, the appeals have grown rapidly. Though the original demonstration RAC Audit program showed a reserved provider group only appealing 14% of the determinations, this is not currently the tone of the industry.

The Administrative Law Judges which is the third level of appeals have a backlog in excess of 357,000 determinations for the 65 Administrative Law Judges or 2.5 years before they can review another case. With the advent of the new RAC Audit group, home health and hospice agencies would be well served to understand all levels of appeal. A memo from Chief Administrative Law Judge of the Office of Medicare Hearings and Appeals (OMHA), Nancy Griswald stated the above figures as well as the fact that hearing rate requests are at an all-time high of 15,000 per week[21]. In 2013, the agency suspended acting on new requests for hearings filed by hospitals, physicians, nursing home facilities and other healthcare provides which constitute approximately 90 percent of the appeals cases. The suspension of appeals would last two to two and a half year. Processing of beneficiaries' appeals continued. Hospitals must wait an estimated two

years before their appeals are heard by an ALJ, during which time the disputed funds are recouped by CMS. According to RACTrac data, hospitals have more than $1 billion at stake in the appeals process and are now facing several years before they will receive a final determination on appealed claims. Timing is crucial in the appeals process. If in Level 1 appeals (Determination) are not filed within a specific timeframe for the appropriate level of appeal, the opportunity to appeal is lost. In all actuality, it could take 780 days for one appeal to be effected. The following table provides a brief overview of the expected time frame.

Table 4: Medicare Appeals Process

Level of Appeal	**Decision Maker**	**Timeframe a Provider is Required to File the Appeal (Days)**	**Days Until Issuance of Decision**
Redetermination by Fiscal Intermediary	MAC personnel – not associated with the initial claim determination	120	60
Reconsideration by a Qualified Independent Contractor	QIC – independent review. May include a panel of physicians or other healthcare professionals	180	60
Administrative Law Judge Hearing	ALJ makes the decision.	60	90
Medicare Appeals Council Review	Appeals Council	60	90
Judicial Review in U.S. District Court	US District Court	60	_ _ _

Figure 6 represents the funds returned to CMS as either restored underpayments made by hospitals or funds that have been overturned on appeals. Even though there has been a delay in the appeal process, as shown in Figure 6, there is a significant increase in underpayments which in 2014 total 173.1 million and an increase in overturned appeals which in 2014 total 157.4 million.

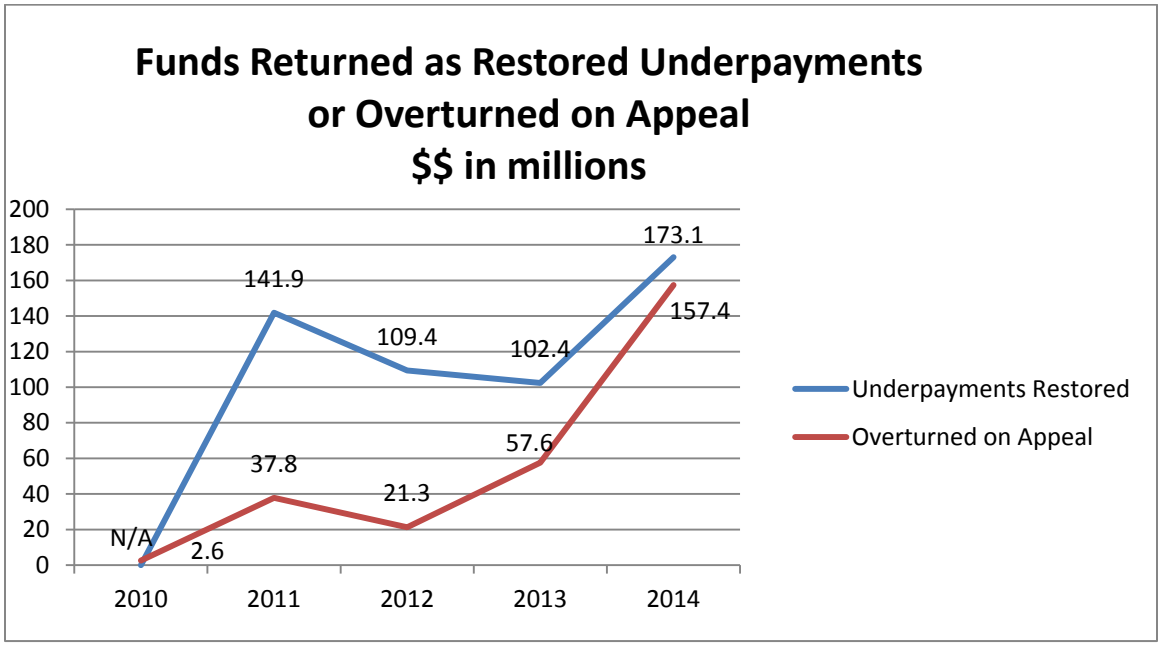


Figure 6: Funds Returned as Restored Underpayments or Overturned on Appeal in Millions

2.9 Trends in CMS Audit and Enforcement Actions

Under the Patient Protection and Affordable Care Act (ACA) legislation which was enacted in March 2010, requires CMS to expand the Recovery Audit Contractor (RAC) program to the Medicare Part C (Medicare Advantage) and Part D (Prescription Drug

Benefit) programs. The amendments to the existing Medicare Fee-for-Service (FFS) RAC statute at section 1893(h) of ACA provide CMS with general authority to enter into contracts with RACs to identify and reconcile overpayments and underpayments and recoup overpayments in Medicare Parts C and D.

In 2015. The Medicare Advantage (MA) program included 3,500 plan options. MA enrolled approximately 16.7 million beneficiaries which equates to 30% of all beneficiaries. CMS paid approximately $170 billion to cover Part A and Part B services. MA plans increased by about 6% to 16.7 million enrollees between 2014 and 2015. While private fee-for-service plans decrease, MA plans continue to increase.

On December 22, 2015, the Centers for Medicare & Medicaid Services (CMS) released a request for information (RFI) and a proposed statement of work (SOW) seeking industry comments on the expansion of the recovery audit contractor (RAC) program to Medicare Part C through the proposed incorporation of RACs into CMS' Risk Adjustment Data Validation (RADV) audit process. CMS contracts with RACs to identify and correct overpayments and underpayments in Medicare Parts A and B and Section 6411(b) of the Patient Protection and Affordable Care Act of 2010 (PPACA) required expansion of the RAC program to Medicare Part C. The RFI sought comment on expanding the Recovery Audit Program to include the identification and correction of overpayments and underpayments associated with diagnosis data submitted to CMS by Medicare Advantage Organizations (MAOs) for Part C payment.

There are two ways in which RACs will be incorporated into CMS' Risk Adjustment Data Validation (RADV) audit process. RACs would perform a comprehensive audit and condition-specific audits. Condition-specific RADV focus on specific conditions audits and or diagnostic codes that are inclines to have increased rates of error in payment. RADV audits are not new to CMS. CMS has been conducting this type of audit on approximately 5% of its Medicare Advantage (MAO) contracts. CMS' future goal is to conduct condition-specific RADV audits on all MAO contracts.

In 2015, approximately 39 million Medicare beneficiaries were enrolled in Medicare Part D. The 39 million beneficiaries equate to 70% of all beneficiaries. Of the 39 million beneficiaries enroll 61% participated in stand-alone prescription drug plans and the remaining 39% participated in Medicare Advantage Prescription Drug plans. In 2014, Medicare spent over $78 billion on Part D.

For purposes of insuring the integrity of the Part D Recovery Audit Contractor (RAC) program, CMS implemented a Data Validation Contractor (DVC) program. The DVC will perform an independent quality check for CMS that will confirm the RAC's improper payment findings and will measure the RAC's accuracy rate. The DVC approves or disapproves RACs improper payment referrals. The DVC is required to validate the RAC's improper payment results before the RAC is permitted to pursue overpayments from Part D sponsors. CMS has contracted Livanta, LLC to perform the Part D RAC data validation functions under vigilant guidance established by CMS.

The Center for Medicare & Medicaid Services (CMS) imposed financial penalties on Medicare Part C and Part D health plan sponsors. CMS imposed more than $9.1 million in Civil Monetary Penalties (CMP) which were linked to its 2015 program audits[22]. The CMP's were considerably higher than in 2014 which totaled $3.75 million. Numerous health insurers that provide Medicare Advantage and Part D plans have had fines alieved against them for various violations[23]. In 2015, one of the largest fines, totaling $3.1 million, was levied against Humana. Humana has well over 3 million Medicare Advantage enrollees which makes Humana a leading health insurer in the industry. Fines were also levied against Envision Pharmaceutical Services for a total of $2.6 million. Envision is a pharmacy benefit manager which is defined as a third-party administrator of prescription drug programs. Envision was acquired by the drug store chain Rite Aid.

At the start of 2016, the Centers for Medicare & Medicaid (CMS) had levied sanctions against Cigna Corporation. CMS communicated to Cigna Corporation, a private health insurance organization that it will be preventing Medicare enrollees from joining its medical services and prescription drug programs. This was due to an audit of the Cigna Corporation's compliance with Medicare rules found significant problems. CMS had previously warned Cigna Corporation of their compliance issues. The sanctions followed an October audit of the insurance organization, which at the time was seeking an antitrust regulators' approval of a merger with Anthem Incorporated.

In summary the goal of the RAC program is to reduce Medicare & Medicaid improper payments through the detection and collection of overpayments and underpayments and

further to prevent the improper payments from occurring in the future. The analytical models developed in this dissertation are purported to help identify and predict the specific hospitalization factors influencing the claims and with future mitigation of the payment recovery and effort by the RAC[24]. Details on the methods and the results that ensued are provided in the following chapters.

CHAPTER III

RESEARCH METHODOLOGY

3.1 Healthcare Cost and Utilization Project (HCUP)

The HCUP is a system of databases that was first developed in 1994 and is sponsored by the Agency for Healthcare Research and Quality (AHRQ)[25]. First collected in 1988, these categories of interrelated databases contain inpatient, outpatient and emergency department patient information that is updated annually. The HCUP is a compilation of patient data collected by state data organizations, hospital associations, private data organizations and the federal government. These combined efforts have made the HCUP into the largest collection of multiyear hospital care data in the United States.

3.2 Nationwide Inpatient Sample Data

The sample data consist of inpatient hospital stay file from the HCUP Nationwide Inpatient Sample (NIS)[26]. The NIS is nationwide database of community hospital inpatient stays. Research and policymakers use NIS data to identify, track and analyze trends in health care utilization, access, charges, quality and outcome. The NIS is nationally representative of all community hospitals (i.e. short-term, non-federal, non-rehabilitation hospitals). The NIS is a sample of hospitals and includes all patients from each hospital, regardless of payer including uninsured. It is drawn from a sampling frame

that contains hospitals comprising about 95 percent of all discharges in the United States. The time frame to be used in this analysis is 2007 inclusive and 2011 inclusive.

3.3 Research Design and Methods

In this study we plan to utilize the datasets obtained from the Nationwide Inpatient Sample (NIS) database towards our analyses of Medicare and Medicaid. The NIS is the largest all-payer inpatient care database in the United States containing data from 1998 to 2013. It contains data from approximately 8 million hospital stays each year accruing from all discharge data from 1,051 hospitals located in 45 States, approximating a 20-percent stratified sample of U.S. community hospitals. The sampling frame for the 2011 NIS is a sample of hospitals that comprises approximately 95 percent of all hospital discharges in the United States. The NIS includes more than 100 clinical and nonclinical data elements for each hospital stay. These include:

- Primary and secondary diagnoses
- Primary and secondary procedures
- Admission and discharge status
- Discharge status

- Patient demographics (e.g., gender, age, race, median income for ZIP Code)
- Expected payment source
- Total charges
- Length of stay
- Hospital characteristics (e.g., ownership, size, teaching status).

The NIS is the only national hospital database containing charge information on all patients, regardless of payer, including persons covered by Medicare, Medicaid, private insurance, and the uninsured.

We have acquired all NIS data for 2007 to 2011 and the statistical analysis software SAS 9.2, SPSS and Microsoft Excel will be employed to process the datasets and perform the analyses. The figure below illustrates the conceptual model employed in this research project.

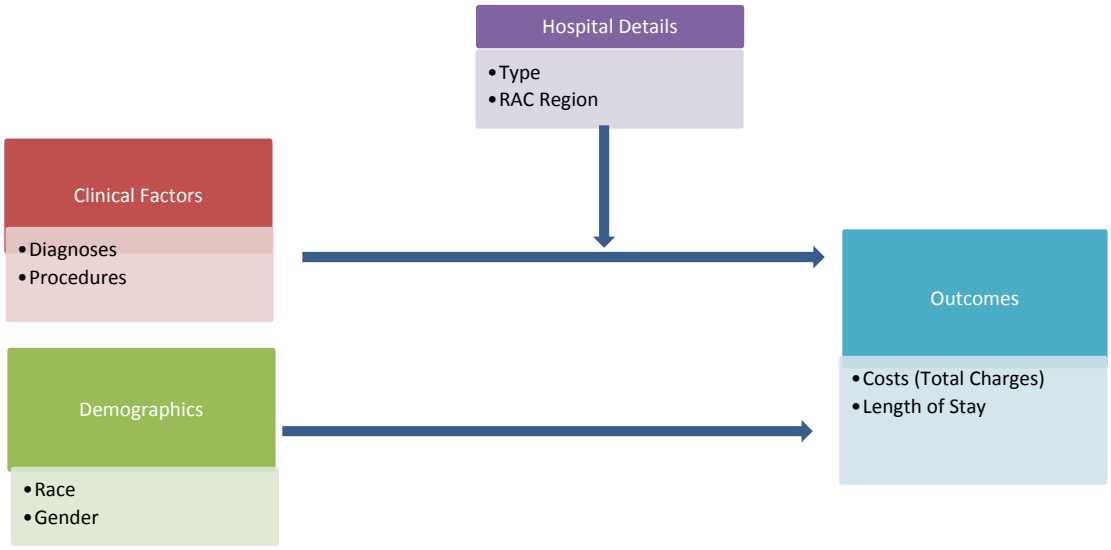


Figure 7: Conceptual model for identifying the factors determining Length of Stay and Costs**.**

Essentially it shows the factors that are hypothesized to affect the research outcomes such as length of stay, payer types and also the hospitalization costs for inpatients data acquired from the NIS database. These factors are categorized as being: clinical such as the type of diagnoses, the number and the number of procedures; demographics group delineates the race of the patient as also the type of insurance (Private, Medicare and Medicaid) while the Hospital Details category consists of its location in the US, and whether it is in the rural, metro or urban areas and this research project specifically speaks to the RAC Regions across the US. The outcomes of interest as identified in this proposal are the length of stay and the costs involved. Using the datasets obtained from

the NIS database appropriate descriptive and inferential statistics (such as multiple linear regression and ANOVAs) will be effected. To relate the factors associated with the research outcome, the length of stay and the costs a predictive model using multiple regression technique will be setup and validated.

Length of Stay (LOS) is an important measure of utilization volume for service in any given hospital. Length of stay is the number of days patient occupied a hospital bed. The total length of stay indicates the number of days care was provided to patients were either discharged or died.

3.4 Statistical Methodology and Analyses

The following methods will be used to analyze the data as appropriate.

Parametric methods such as:

- One-way ANOVA
- Multiple Linear Regression

Data will be categorized as appropriate to investigate the research questions. All computations will be performed with SAS software running on a Windows operating system. All invalid data will be reported and a reason given for why the data is considered invalid (example –missing value). Where outlying data are observed, analyses will be performed with and without the outlying data. Sound statistical evidence that the data are outlying (i.e. outlying data is more than 4 standard deviations beyond the mean of comparable data) will be documented.

Outlying data can be removed from an analysis if it can be shown to improve the power of the statistical tests or if not removing it would skew the result[27].

Descriptive and distribution analyses will be performed for all appropriate variables. Continuous variables will be assessed for normality. If the data is normally distributed, parametric methods will be used to analyze data otherwise non-parametric methods will be used. Non-parametric methods will be used to analyze score data. Categorical analyses with the appropriate methods will be used to compare categorical variables.

SAS procedures will be used to perform the analyses: the GLM Procedure for Total Charges for 2007 inclusive through 2011 inclusive. In SAS, a one-way multivariate regression analysis will be used to determine whether there are any differences between the number of procedures, the number of diagnoses, total charges and length of stay in relationship to the RAC Regions.

Analysis of variance was developed to analyze difference among group means and their associated procedures. This method was developed by Ronald Fisher. In ANOVA setting, the observed variance in a particular variable is partitioned into components attributed to different sources of variation. ANOVA provides a statistical test of whether or not the means of several groups are equal, and therefore generalizes the t-*test* to more than two groups. ANOVA is useful in comparing three or more means for statistical significance. In this thesis, ANOVA is used to study and assess the difference among means of the data set for 2007-2011 by comparing the value of *F ratio* to the *F crit.* The *F ratio* represents the ratio of the variance between groups to the variance within groups. The *F*

crit represents the threshold value when the test is to be rejected[28]. ANOVA will be necessary to evaluate the difference among RAC regions within the years 2007-2011. The results obtained from ANOVA analysis will give a better understanding to the underlying reasons for differences in outcomes among the four RAC regions.

Simple Linear Regression and Multiple Linear Regression are related statistical methods for modeling the relationship between two or more random variables using a linear equation. Simple linear regression refers to a regression on two variables while multiple regression refers to a regression on more than two variables[29]. A simple linear regression equation is represented below.

$$\textbf{Simple Linear: } y = \alpha + \beta_1 X_1$$

A multiple regression equation is represented below.

$$\textbf{Multiple Regression: } y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \varepsilon$$

In both the simple linear and multiple regression models α is the constant. β represents the coefficient for the independent variable(s) 1 through n. Where *n* is the number for the subscript for the last independent variable. In the case of the simple linear model there is only one independent variable. ε is the error value in the multiple regression model.

A related method to determine if a model is a "good" fit is called the Akaike information criterion (AIC). The AIC was developed by Hirotsugu Akaike in 1971 and is a measure

of the goodness of fit for an estimated statistical model. The AIC trades the complexity of an estimated model against how well the model fits the data.

Simple variable selection algorithms are ad hoc or method based. A common method is the greedy hill climbing approach. This approach evaluates a variable subset and then modifies that subset to determine if an improved subset exists. Thus, this greedy algorithm adds or deletes the respective best or worst variable. The stepwise regression method is a popular choice, which demonstrates a greedy algorithm.

Stepwise regression is an automatic procedure for statistical model selection where there are a large number of variables added or dropped. New variables are added at each stage in the process and variables are checked to see if some can be deleted without increasing the Residual Sum of Squares (RSS). The stepwise regression process stops when a selection is maximized or when it can no longer be improved. Backward regression and forward regression are variations[30].

A confidence of 95% will be used unless there is a need to increase or decrease the percentage. Thus, a p value < 0.05 for a 95% confidence would be significant. In many cases all the data records will be used in the analysis unless otherwise stated.

In summary, the regression methods will determine what data elements or explanatory variables significantly influence or 'predict' the outcome. Multiple Linear regression using the Stepwise Regression technique as available in SAS using maximum R^2 option will be employed.

CHAPTER IV
RESULTS AND DISCUSSION

The research goal of this dissertation as indicated in the previous chapters were two-fold: one to formulate a model that will serve to describe the RAC process to help identify the critical variables affecting its performance and second to utilize those identified variables as outcomes to be predicted using the hospital discharge datasets available from HCUP.

Towards meeting the first goal several process modeling techniques were studied with a view to identifying the most appropriate for the RAC process and the best modeling approach with the most potential for utilization in not only this dissertation study but also for potential future studies was the Logic Model[31,32].

4.1 Logic Model of the Recovery Audit Contract Process

A Logic Model is a systematic way of presenting the visualization of a process with the relationships amongst its various elements, the activities involving them and the results achieved through their combination[33]. Specifically they serve as actionable plans or visualizations with clear outcomes and explicit steps for solving program related issues or problems. Thus they are useful to program administrators who are trying to implement changes in their program or indicate such changes for future process modifications[34,35].

The Recovery Audit Contractor (RAC) program was developed as an integrity plan to protect the Medicare Trust Fund from improper payments to hospitals on patient stays. Although the program has been successful by recovering over $9 billion in improper Medicare payments since its inception, it is not without scrutiny from stakeholders. Protecting the Trust Fund is a high priority for the Centers of Medicare and Medicare. As described herein, the Medicare RAC program is moving forward despite significant concerns from healthcare providers and hospital administrators. This situation implores the question as to whether the RAC program will be able to sustain its mission and goal in the future. The Logic Model in Figure 8 lays out a graphic overview of the RAC process along with an underlying set of assumptions for the purposes of identifying if there is a future for the RAC program.

The assumptions reflect the identified various elements that are in place and are necessary to proceed with the future mission and goal of the RAC program.

The tangible resources (inputs) invested in the RAC program are software billing edits, ICD-9 coding materials, staffing and a considerable amount of time. Another important input displayed in the logic model is the congressional support. Section 1893(h)(8) requires the Department of Health and Human Service Secretary to annually submit to Congress a report on the use of recovery audit contractors. The report is comprehensive including information on the performance of the RAC regions and the under and over payments identified by the RAC contractors.

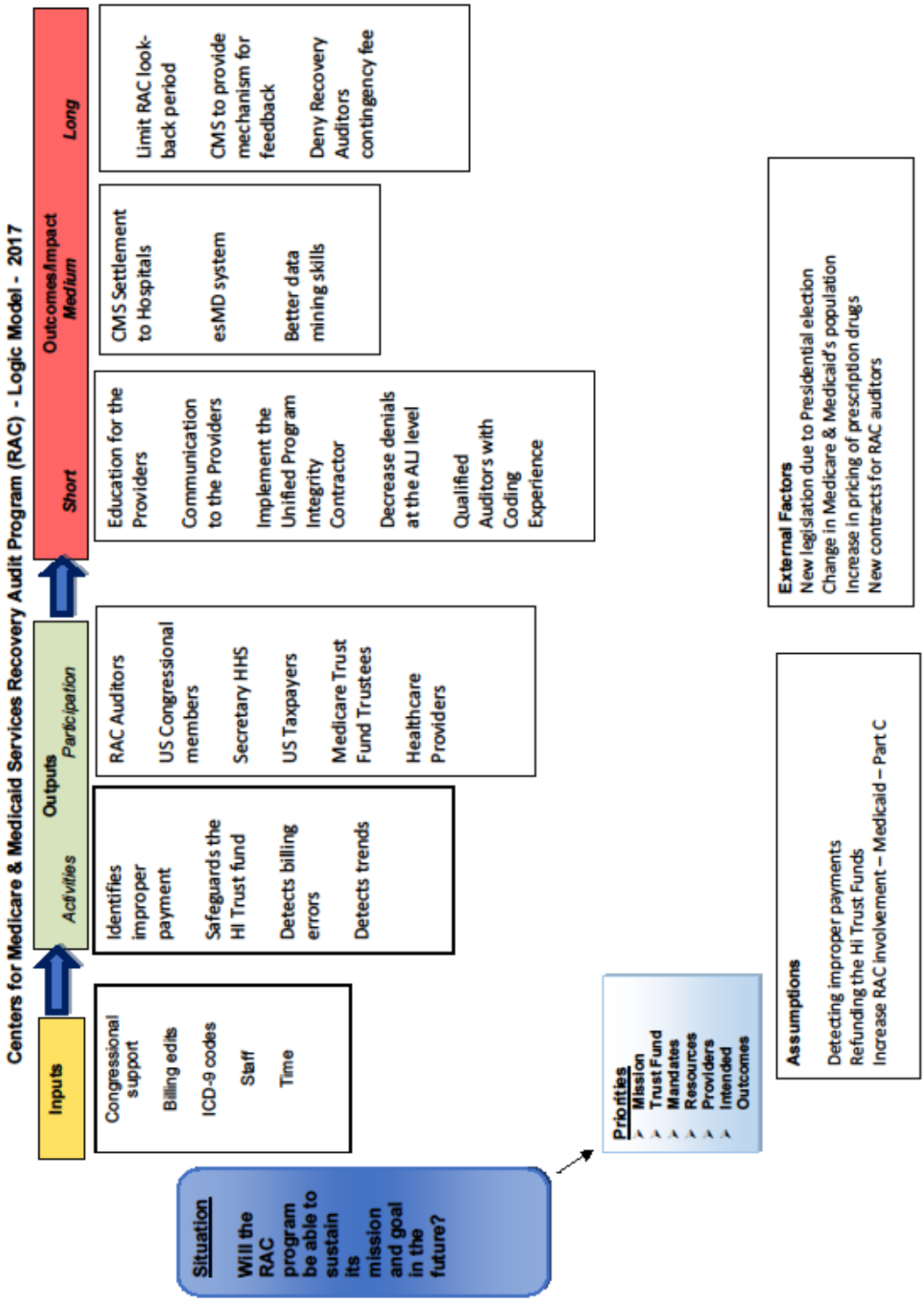


Figure 8 Schematic Diagram of the RAC Process Logic Model

In terms of the outputs displayed in the logic model, the activities include the mission of the RAC program which is to identify improper payment, detect billing errors, detect trends in billing, and ultimately these activities safe guards the Medicare Trust fund.

The Logic Model captures not only the participants involved in the RAC program but also those who are affected by the RAC program. Identified are the RAC auditors, the Secretary of Health and Human Services, healthcare provides, US Congressional members, Medicare Trust Fund Trustees and the U.S. taxpayer. Each of the identified participants in the Logic Model can induce change in the RAC program.

The outcomes section of the Logic Model is divided into three sections: short-term outcomes, intermediate outcomes and long-term outcomes. The short-term impact of the RAC program aligns with the output activities and can be immediately implemented for positive change.

In the short term, better communication to the providers can alleviate the frustration felt by the providers and hospitals who complain that they receive little or no guidance on prevention of billing errors. CMS needs to provide a mechanism to assist providers on how to prevent errors before they occur.

Other issues that can be immediately addressed would be RAC providing training for the providers of any updates in process change and the hiring of qualified auditors with medical coding experience to better understand the national coding guidelines. There is a lengthy delay at the Administrative Law Judges appeals level. The appeals backlog is

significant to the providers as Medicare will offset the provider's Medicare funds to recoup the disputed overpayment. Steps are required to encourage the RACs to make every effort to make accurate claims denial. As a short-term impact to the RAC program, CMS planned to establish a Unified Program Integrity contractor (UPIC) to reorganize and consolidate the work of the Medicaid Integrity Contractors (MICs) and the Medicare Zone Program Integrity Contractors (ZPICs). CMS expects contracts with ZIPCs and MICs to end once UPICs are implemented to specific regions. The introduction of UPICs will bring further consolidation and increased claims data transparency and availability to integrity contractors. There will be increased potential for government scrutiny of claims payment by federal healthcare programs.

The three main areas that would have an intermediate impact on the RAC program are: CMS reinstituting settlement payments to eligible hospitals and providers, uniformity in utilizing the esMD system and utilizing data analytics and mining.

In August 2014, CMS made available an administrative settlement process to eligible hospitals to alleviate administrative burdens for all parties involved. Hospitals who were willing to withdraw their pending appeals received a timely payment at 68 percent of the net allowable for the claims associated with those appeals. CMS paid over $1.5 billion to the eligible hospital providers[36].

CMS needs to make this settlement process available again to alleviate the financial burdens to the hospital providers as a good faith act. CMS can best accomplish this process as an intermediate outcome. Another intermediate outcome for the RAC

program's viability would be if every RAC contractor used the esMD system. This system allows contractors to electronically receive responses to the Additional Documentation Request (ADR) letter sent by providers during the claims review process. Not all contractors participate in the enhanced functionality esMD offers. By not participating in this paperless transmission of electronic medical records, it creates a paper burden for the hospitals and providers who are forced to manually submit the requested record either by sending a fax or letter.

RACs receive data from CMS containing information for claims within their specified region. RACs have access to a massive database. RACs review the data using a proprietary data mining methodology. Although the methodology can identify correlations and aberrations in billing and coding errors, the data mining software is not very sophisticated. To improve the data mining process for RACS and providers, other data mining software needs to be reviewed.

In the long term, there are several outcomes that could improve RACs program. They are:

- limit the RACs look-back period from 3 years to 6 months;
- provide a mechanism for providers to provide feedback to CMS on the Recovery Auditors performance; and
- deny Recovery Auditors their contingency fee until after the second level of appeal is finished.

Currently CMS has limited the look-back period for Recovery Auditor reviews to a maximum of 3 years with a maximum look-back date of October 1, 2007. In 2015, CMS limited the look-back to 6 months in instances where the hospital submits claims within 3 months of the date of service. CMS needs to continue this process in future years. In the past, providers and hospitals did not have a mechanism to voice their concerns to CMS about a RACs performance. CMS will need to provide such a mechanism, such as, a provider survey or evaluation which allows providers a channel to voice concerns to CMS. Providers have previously been concerned that Recovery Auditors were paid immediately upon denial and recoupment of claims. If the Recovery Auditors did not receive their contingency fees until after the second level appeal is finished, it would help reassure providers that the claim decision made by the Recovery Auditor was correctly based on regulations and Medicare statutes.

The external factors listed in the logic model are elements that affect the RAC program over which there is little or no control. It is expected that the newly elected U.S. President will introduce either new legislation to CMS or introduce reforms to the existing model. Consequences to these actions are, at this juncture in time, unpredictable. Future Medicare expenditures will depend on a number of factors such as the change in CMS's population, its growth, death rates, wage increases and composition of the population eligible for Medicare benefits. CMS Office of the Actuary (OACT) projects a comparatively higher per capita growth rate in the coming years for Medicare Part D than for the other parts of the program due to higher costs associated with expensive specialty

drugs. Per capita spending growth is projected to be 5.8% for Part D, compared to 3.2% for Part A and 4.6% for Part B[37].

On October 31, 2016, CMS awarded Medicare Fee-for-Service Recovery Contractor (RAC) contracts to:

- Region 1 – Performant Recovery, Inc.;
- Region 2 – Cotiviti, LLC;
- Region 3 – Cotiviti, LLC;
- Region 4 – HMS Federal Solutions; and
- Region 5 – Performant Recovery, Inc.

There is a need to transition work from the outgoing RAC to a different incoming RAC. The transition is not always successful. Timeframes may overlap involving outstanding claims and appeals, which do not transition to the incoming RAC.

It is very likely that the RAC program will be able to sustain its mission and goal in the future as CMS is committed to working with RACS, providers and various stakeholders to continue to enhance the RAC program.

4.2 Analyses of Hospitalization Variables across RAC Regions

In view of developing a predictive model for the outcomes of Length of Stay and Total Charges across the RAC regions it was imperative to identify the nature and variations of these two and any other significant hospitalization variables that were gleaned from

literature to determine those two outcomes. Accordingly, several Analyses of Variance (ANOVA) were done using the NIS data for the years 2007 to 2011 for the four RAC regions of interest.

4.2.1 Analysis of Variance (ANOVA) for the RAC Regions for 2007 to 2011

An analysis of variance (ANOVA) was used in order to analyze the differences between group means and their associated procedures. In this case the groups were the RAC regions and the means tested belonged to four different hospitalization variables of interest, namely, Length of Stay, Total Charges, Number of Diagnoses and Number of Procedures. ANOVA was performed for all the years 2007 to 2011 –however for the sake of brevity only the results of 2007, 2010 and 2011 are shown in the sections below since they all revealed similar patterns of variations across the four RAC regions over the years.

Furthermore since the ANOVA for the four RAC regions showed statistically significant differences hence a Post Hoc Tukey HSD test was performed for each of the years to indicate what the variation was for each pair of regions for the four hospitalization variables of interest namely Length of Stay, Number of Diagnoses, Number of Procedures and Total Charges.

The results of the ANOVA and the Post Hoc Tukey HSD tests for the years 2007, 2010 and 2011 are as presented below.

ANOVA 2007

		N	Mean	Std. Deviation
Number of procedures on this record	1	1633351	1.81	2.159
	2	1582525	1.49	1.940
	3	2971295	1.55	1.988
	4	1856244	1.61	2.020
	Total	8043415	1.61	2.025
Number of diagnoses on this record	1	1633351	6.66	4.096
	2	1582525	6.95	4.228
	3	2971295	6.78	4.389
	4	1856244	6.55	4.615
	Total	8043415	6.73	4.356
Total charges (cleaned)	1	1632519	25823.61	45782.989
	2	1568624	21127.52	35796.287
	3	2936434	25381.38	43332.481
	4	1743371	32469.16	60960.093
	Total	7880948	26194.21	47170.178
Length of stay (cleaned)	1	1633327	4.97	7.520
	2	1582480	4.30	5.756
	3	2971246	4.65	6.805
	4	1855981	4.31	7.146
	Total	8043034	4.57	6.852

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
Number of procedures on this record	2184.119	3	8043411	.000
Number of diagnoses on this record	2505.668	3	8043411	.000
Total charges (cleaned)	26339.539	3	7880944	.000
Length of stay (cleaned)	4243.144	3	8043030	.000

		Sum of Square	df	F
Number of procedures on this record	Between Groups	94680.219	3	7719.317
	Within Groups	32885116.81	8043411	
	Total	32979797.03	8043414	
Number of diagnoses on this record	Between Groups	155819.786	3	2740.478
	Within Groups	152445752.8	8043411	
	Total	152601572.6	8043414	
Total charges (cleaned)	Between Groups	1.111E+14	3	16746.807
	Within Groups	1.742E+16	7880944	
	Total	1.754E+16	7880947	
Length of stay (cleaned)	Between Groups	514293.608	3	3655.884
	Within Groups	377152618.9	8043030	
	Total	377666912.5	8043033	

		Sig.
Length of stay (cleaned)	Between Groups	.000
	Within Groups	
	Total	
Number of diagnoses on this record	Between Groups	.000
	Within Groups	
	Total	
Number of procedures on this record	Between Groups	.000
	Within Groups	
	Total	
Total charges (cleaned)	Between Groups	.000
	Within Groups	
	Total	

As can be seen in the table above, the following results were found.

There was a significant effect of Number of Procedures at the $p<.05$ level for the four RAC regions [$F(3, 804341) = 7719.317$, $p = 0.00$].

There was a significant effect of Number of Diagnoses at the $p<.05$ level for the four RAC regions [$F(3, 804341) = 2740.478$, $p = 0.00$].

There was a significant effect of Total Charges at the $p<.05$ level for the four RAC regions [$F(3, 788094) = 16746.80$, $p = 0.00$].

There was a significant effect of Length of Stay at the $p<.05$ level for the four RAC regions [$F(3, 804303) = 3655.884$, $p = 0.00$].

Since all the four variables yielded significant effects a Post-Hoc Tukey Test was accordingly performed to determine which of the four RAC regions differed and by how much from each other. The Post Hoc Test results are as presented below. The numbers associated with the Mean Difference with an asterisk indicates a significant difference for a particular pair of RAC regions.

For example, RAC regions 1 and 2 differed by a mean difference of 0.314 procedures with RAC region 1 having a higher number of procedures than RAC region 2 and similarly for the other pairs indicated in the table below.

Post Hoc Tests(2007)

Dependent Variable		(I) RACRegion	(J) RACRegion	Mean Difference (I-J)
Number of procedures on this record	Tukey HSD	1	2	.314*
			3	.255*
			4	.201*
		2	1	-.314*
			3	-.059*
			4	-.113*
		3	1	-.255*
			2	.059*
			4	-.053*
		4	1	-.201*
			2	.113*
			3	.053*
Number of diagnoses on HSD this record	Tukey	1	2	-.296*
			3	-.118*
			4	.111*
		2	1	.296*
			3	.178*
			4	.407*
Dependent Variable		(I) RACRegion	(J) RACRegion	Mean Difference (I-J)
		3	1	.118*
			2	-.178*
			4	.229*
		4	1	-.111*
			2	-.407*
			3	-.229*

Total charges (cleaned)	Tukey HSD	1	2	4696.090*
			3	442.225*
			4	-6645.548*
		2	1	-4696.090*
			3	-4253.865*
			4	-11341.638*
		3	1	-442.225*
			2	4253.865*
			4	-7087.773*
		4	1	6645.548*
			2	11341.638*
			3	7087.773*
Length of stay (cleaned)	Tukey HSD	1	2	.664*
			3	.315*
			4	.656*
		2	1	-.664*
			3	-.348*
			4	-.008
		3	1	-.315*
			2	.348*
			4	.341*
		4	1	-.656*
			2	.008
			3	-.341*

Post Hoc comparisons using the Tukey HSD test indicated that the mean value for the Number of Procedures for RAC Region 1 (M = 1.81, SD = 2.159) was significantly higher than the mean value for RAC Region 2 (M = 1.49, SD = 1.940), significantly higher than the mean value for RAC Region 3 (M = 1.55, SD = 1.988) and significantly higher than the mean value for RAC Region 4 (M = 1.61, SD = 2.020).

Similarly Post Hoc comparisons using the Tukey HSD test indicated that the mean value for the Number of Diagnoses for RAC Region 1 was significantly lesser than the mean value for RAC Region 2 (Mean Difference = -0.296), significantly lesser than the mean value for RAC Region 3 (Mean Difference = -0.118) and significantly higher than the mean value for RAC Region 4 (Mean Difference = 0.111).

Post Hoc comparisons using the Tukey HSD test indicated that the mean value for the Total Charges for RAC Region 1 was significantly higher than the mean value for RAC Region 2 (Mean Difference = 4696.09 dollars), significantly higher than the mean value for RAC Region 3 (Mean Difference = 442.225 dollars) and significantly lesser than the mean value for RAC Region 4 (Mean Difference = -6645.548 dollars).

Lastly Post Hoc comparisons using the Tukey HSD test indicated that the mean value for the Length of Stay for RAC Region 1 was significantly higher than the mean value for RAC Region 2 (Mean Difference = 0.664 days), significantly higher than the mean value for RAC Region 3 (Mean Difference = 0.315 days) and significantly higher than the mean value for RAC Region 4 (Mean Difference = 0.656 days).

It was interesting to note that RAC Region 4 incurred a greater cost with a shorter hospital stay regardless of there being slightly lesser number of diagnoses on average per patient and slightly above average number of procedures conducted on such patients compared to the other 3 regions.

This could perhaps be explained due to RAC region 4 having more large sized hospitals and many more of teaching type hospitals than non-teaching type with possibility of innovative procedures and availability of specialists in such hospitals requiring shorter stay but effective diagnosis, treatment and discharge.

ANOVA 2010

		N	Mean	Std. Deviation
Length of stay (cleaned)	1	1646229	4.98	7.383
	2	1527659	4.40	6.270
	3	2792898	4.77	6.915
	4	1830752	4.34	7.154
	Total	7797538	4.64	6.958
Number of diagnoses on this record	1	1646252	7.71	4.590
	2	1527811	8.85	5.654
	3	2795396	8.13	5.329
	4	1830982	7.79	5.657
	Total	7800441	8.10	5.345
Number of procedures on this record	1	1646252	1.79	2.180
	2	1527811	1.59	2.146
	3	2795396	1.59	2.072
	4	1830982	1.57	2.080
	Total	7800441	1.63	2.114
Total charges (cleaned)	1	1638785	32099.20	55451.119
	2	1526585	28213.29	45569.777
	3	2769754	32500.75	54590.063
	4	1719259	38669.13	67411.965
	Total	7654383	32945.18	56423.225

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
Length of stay (cleaned)	3098.595	3	7797534	.000
Number of diagnoses on this record	32414.291	3	7800437	.000
Number of procedures on this record	2302.535	3	7800437	.000
Total charges (cleaned)	14098.248	3	7654379	.000

		Sum of Square	df	F
Length of stay (cleaned)	Between Groups	489151.052	3	3372.058
	Within Groups	377037048.7	7797534	
	Total	377526199.8	7797537	
Number of diagnoses on this record	Between Groups	1303474.299	3	15300.583
	Within Groups	221509408.9	7800437	
	Total	222812883.2	7800440	
Number of procedures on this record	Between Groups	59383.437	3	4438.785
	Within Groups	34785550.96	7800437	
	Total	34844934.40	7800440	
Total charges (cleaned)	Between Groups	9.223E+13	3	9693.578
	Within Groups	2.428E+16	7654379	
	Total	2.437E+16	7654382	

		Sig.
Length of stay (cleaned)	Between Groups	.000
	Within Groups	
	Total	
Number of diagnoses on this record	Between Groups	.000
	Within Groups	
	Total	
Number of procedures on this record	Between Groups	.000
	Within Groups	
	Total	
Total charges (cleaned)	Between Groups	.000
	Within Groups	
	Total	

As can be seen in the table above, the following results were found.

There was a significant effect of Number of Procedures at the $p<.05$ level for the four RAC regions [$F(3, 780043) = 4438.785$, $p = 0.00$].

There was a significant effect of Number of Diagnoses at the $p<.05$ level for the four RAC regions [$F(3, 780043) = 15300.58$, $p = 0.00$].

There was a significant effect of Total Charges at the $p<.05$ level for the four RAC regions [$F(3, 765437) = 9693.578$, $p = 0.00$].

There was a significant effect of Length of Stay at the $p<.05$ level for the four RAC regions [$F(3, 779753) = 3372.058$, $p = 0.00$].

Since all the four variables yielded significant effects a Post-Hoc Tukey Test was accordingly performed to determine which of the four RAC regions differed and by how much from each other. The Post Hoc Test results are as presented below. The numbers associated with the Mean Difference with an asterisk indicates a significant difference for a particular pair of RAC regions.

Post Hoc Tests (2010)

Dependent Variable		(I) RACRegion	(J) RACRegion	Mean Difference (I-J)
Length of stay (cleaned)	Tukey HSD	1	2	.582*
			3	.209*
			4	.637*
		2	1	-.582*
			3	-.373*
			4	.055*
		3	1	-.209*
			2	.373*
			4	.428*
		4	1	-.637*
			2	-.055*
			3	-.428*
Number of diagnoses on HSD this record	Tukey	1	2	-1.145*
			3	-.418*
			4	-.077*
		2	1	1.145*
			3	.727*
			4	1.068*
Dependent Variable		(I) RACRegion	(J) RACRegion	Mean Difference (I-J)
		3	1	.418*
			2	-.727*
			4	.341*
		4	1	.077*
			2	-1.068*
			3	-.341*

Number of procedures on HSD this record	Tukey	1	2	.203*
			3	.209*
			4	.227*
		2	1	-.203*
			3	.005
			4	.024*
		3	1	-.209*
			2	-.005
			4	.019*
		4	1	-.227*
			2	-.024*
			3	-.019*
Total charges (cleaned)	Tukey HSD	1	2	3885.904*
			3	-401.555*
			4	-6569.935*
		2	1	-3885.904*
			3	-4287.459*
			4	-10455.839*
		3	1	401.555*
			2	4287.459*
			4	-6168.380*
		4	1	6569.935*
			2	10455.839*
			3	6168.380*

Post hoc comparisons using the Tukey HSD test indicated that the RAC Region 4 once again incurred a greater cost regardless of there being lesser length of stay on average per patient compared to the other 3 regions.

ANOVA 2011

		N	Mean	Std. Deviation
Length of stay (cleaned)	1	1758518	5.00	7.621
	2	1532273	4.39	5.833
	3	2934194	4.67	6.645
	4	1798254	4.23	7.065
	Total	8023239	4.59	6.829
Number of diagnoses on this record	1	1758552	8.45	5.318
	2	1532282	9.42	6.194
	3	2934261	8.73	5.680
	4	1798495	8.61	6.149
	Total	8023590	8.77	5.823
Number of procedures on this record	1	1758552	1.74	2.120
	2	1532282	1.57	2.231
	3	2934261	1.57	1.999
	4	1798495	1.61	2.140
	Total	8023590	1.62	2.104
Total charges (cleaned)	1	1757761	36536.67	74290.690
	2	1531049	28287.01	46006.057
	3	2923701	35575.04	61300.648
	4	1629042	40865.94	76449.775
	Total	7841553	35466.78	65437.363

		Sum of Square	df	F
Length of stay (cleaned)	Between Groups	606680.163	3	4343.147
	Within Groups	373579905.5	8023235	
	Total	374186585.7	8023238	
Number of diagnoses on this record	Between Groups	882573.035	3	8703.555
	Within Groups	271207213.4	8023586	
	Total	272089786.4	8023589	
Number of procedures on this record	Between Groups	39159.191	3	2951.399
	Within Groups	35485674.00	8023586	
	Total	35524833.19	8023589	
Total charges (cleaned)	Between Groups	1.285E+14	3	10038.174
	Within Groups	3.345E+16	7841549	
	Total	3.358E+16	7841552	

		Sig.
Length of stay (cleaned)	Between Groups	.000
	Within Groups	
	Total	
Number of diagnoses on this record	Between Groups	.000
	Within Groups	
	Total	
Number of procedures on this record	Between Groups	.000
	Within Groups	
	Total	
Total charges (cleaned)	Between Groups	.000
	Within Groups	
	Total	

As can be seen in the table above, the following results were found.

There was a significant effect of Number of Procedures at the $p<.05$ level for the four RAC regions [$F(3, 802323) = 2951.399_{2951.399}$, $p = 0.00$].

There was a significant effect of Number of Diagnoses at the p<.05 level for the four RAC regions [F(3, 802323) = 8703.555, p = 0.00].

There was a significant effect of Total Charges at the p<.05 level for the four RAC regions [F(3, 784154) = 10038.174, p = 0.00].

There was a significant effect of Length of Stay at the p<.05 level for the four RAC regions [F(3, 802323) = 4343.147, p = 0.00].

Since all the four variables yielded significant effects a Post-Hoc Tukey Test was accordingly performed to determine which of the four RAC regions differed and by how much from each other. The Post Hoc Test results are as presented below.

Post Hoc Tests(2011)

Dependent Variable		(I) RACRegion	(J) RACRegion	Mean Difference (I-J)
Length of stay (cleaned)	Tukey HSD	1	2	.609*
			3	.321*
			4	.766*
		2	1	-.609*
			3	-.287*
			4	.157*
		3	1	-.321*
			2	.287*
			4	.445*
		4	1	-.766*
			2	-.157*
			3	-.445*

Number of diagnoses on HSD this record	Tukey	1	2	-.972*
			3	-.280*
			4	-.160*
		2	1	.972*
			3	.692*
			4	.812*

Dependent Variable		(I) RACRegion	(J) RACRegion	Mean Difference (I-
		3	1	.280*
			2	-.692*
			4	.120*
		4	1	.160*
			2	-.812*
			3	-.120*

Number of procedures on HSD this record	Tukey	1	2	.172*
			3	.178*
			4	.136*
		2	1	-.172*
			3	.006*
			4	-.036*
		3	1	-.178*
			2	-.006*
			4	-.042*
		4	1	-.136*
			2	.036*
			3	.042*

Total charges (cleaned)	Tukey HSD	1	2	8249.661*
			3	961.634*
			4	-4329.267*
		2	1	-8249.661*
			3	-7288.027*
			4	-12578.928*
		3	1	-961.634*
			2	7288.027*
			4	-5290.901*
		4	1	4329.267*
			2	12578.928*
			3	5290.901*

Post hoc comparisons using the Tukey HSD test indicated that the RAC Region 4 once again incurred a greater cost regardless of there being lesser length of stay on average per patient compared to the other 3 regions.

4.3 Predictive Model using Multiple Linear Regression

Towards meeting the second research goal in this dissertation of formulating a predictive model for the identified variables of Length of Stay and Total Charges a Multiple Regression Modeling technique was employed for the data contained in the years 2007 to 2011. One of the best measures of effectiveness of the predictive model is the use of test measure R^2. It is a measure in statistics of how close the data are to the fitted regression line. It is also known as the coefficient of determination, or the coefficient of multiple determinations for multiple regression. It is the percentage of the response variable variation that is explained by a linear model.

$$\mathbf{R^2 = Explained\ variation\ /\ Total\ variation}$$

R^2 is always between 0 and 100%. 0% means the model explains none of the variability of the response data around its mean. 100% indicates that the model explains all the variability of the response data around its mean. Generally, the higher the R-squared, the better the model fits the data[38]. Statistical significance was defined as a p-value < 0.05.

There were 18 independent variables found to be significant (after the performing of the stepwise regression modelling technique) and used in the Multiple Linear Regression.

The variables are as follows:

- Payer Type 1 – Medicare
- Payer Type 1 – Medicaid
- Race 1 - White
- Race 2 - Black
- Race 3 - Hispanic
- Race 4 - Asian or Pacific Islander
- Race 5 - Native American
- Race 6 - Other
- Gender – Male – Female
- RAC Region 1 (CT, DE, DC, ME, MD, MA, NH, NJ, NY, PA, RI and VT)
- RAC Region 2 (IL, IN, KY, MI, MN, OH and WI)
- RAC Region 3 (AL, AR, CO, FL, GA, LA, MS, NM, NC, OK, SC, TN, TX, VA, WV, Puerto Rico and U.S. Virgin Islands)
- RAC Region 4 (AK, AZ, CA, HI, ID, IA, KS, MO, MT, ND, NE, NV, OR, SD, UT, WA, WY, Guam, American Samoa and Northern Marianas)

- Age – 65 and older
- LOS – length of stay
- Number of Procedures
- Number of Diagnoses
- DRG - Diagnostic Related Groups

All the years 2007 to 2011 yielded very good fitting models to explain the Total Charges each year based on the explanatory variables above. Just the R^2 values for the 5 years is shown below indicating the power of this model in predicting Total Charges for each of the RAC Regions 1 to 4 and including the explanatory variables alongside or even on their own to make appropriate comparisons perhaps between the Total Charges expected from a Medicaid patient belonging to a specific RAC region or of a specific gender and so on. To date no such model is available in being able to predict the Total Charges based on the four RAC Regions making it one of the novel contributions of the dissertation.

Following the listing of the R^2 values details on the Regression Model itself for the years 2007, 2010 and 2011 are also presented below.

Table 5: R^2 of Models for Predicting Total Charges (TOTCHG) 2007 to 2011

2007

R- Squared	Coeff Var	Root MSE	TOTCHG Mean
0.578956	116.3967	32774.61	28157.67

2008

R- Squared	Coeff Var	Root MSE	TOTCHG Mean
0.616915	108.3469	36948.40	34101.95

2009

R- Squared	Coeff Var	Root MSE	TOTCHG Mean
0.597506	116.6897	42871.04	36739.34

2010

R- Squared	Coeff Var	Root MSE	TOTCHG Mean
0.590900	115.2319	43618.78	37853.03

2011

R- Squared	Coeff Var	Root MSE	TOTCHG Mean
0.542953	133.9324	55232.76	41239.27

4.3.1. Multiple Linear Regression Model – 2007

Dependent Variable: TOTCHG Total charges (cleaned)					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	18	7.4755967E15	4.1531093E14	386632	<.0001
Error	5.06E6	5.4366036E15	1074175177.6		
Corrected Total	5.06E6	1.29122E16			

R-Square	**Coeff Var**	**Root MSE**	**TOTCHG Mean**
0.578956	116.3967	32774.61	28157.67

The Multiple Linear Regression Model for the Year 2007 for Total Charges as the Outcome to be predicted by the various Hospitalization variables especially the four RAC regions yielded a very reasonably good fit model with a R2 value of 0.57896.

The various coefficients associated with the explanatory variables (i.e. the hospitalization variables) in the model are as shown below.

Parameter		Estimate		Std Error	t value	Pr>|t|
Intercept		1571.075	B	103.1072	15.24	<.0001
MEDICAID		-4130.19	B	41.8989	-98.58	<.0001
MEDICARE		-3336.23	B	43.21097	-77.21	<.0001
RACE	1	647.2929	B	83.13986	7.79	<.0001
RACE	2	627.9819	B	89.46166	7.02	<.0001
RACE	3	2182.937	B	90.34392	24.16	<.0001
RACE	4	-550.831	B	121.5424	-4.53	<.0001
RACE	5	-932.415	B	184.3123	-5.06	<.0001
RACE	6	0	B	.	.	.
FEMALE	0	2592.324	B	29.90491	86.69	<.0001
FEMALE	1	0	B	.	.	.
RACRegion	1	-15087.9	B	41.83146	-360.68	<.0001
RACRegion	2	-17117.8	B	60.56956	-282.61	<.0001
RACRegion	3	-9914.83	B	38.48729	-257.61	<.0001
RACRegion	4	0	B	.	.	.
AGE		62.86813		0.850522	73.92	<.0001
LOS		3987.622		2.361792	1688.39	<.0001
NPR		7833.382		7.634388	1026.07	<.0001
NDX		558.4523		4.251403	131.36	<.0001
DRG		-1.89697		0.077746	-24.4	<.0001
DISPUNIFORM		480.5029		3.883253	123.74	<.0001
ASOURCE		-864.102		8.314413	-103.93	<.0001

4.3.2. Multiple Linear Regression Model – 2010

Dependent Variable: TOTCHG Total charges (cleaned)					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	18	3.6353094E15	2.0196163E14	106150	<.0001
Error	1.32E6	2.5168437E15	1902597680.9		
Corrected Total	1.32E6	6.1521531E15			

R-Square	Coeff Var	Root MSE	TOTCHG Mean
0.590900	115.2319	43618.78	37853.03

The Multiple Linear Regression Model for the Year 2011 for Total Charges as the Outcome to be predicted by the various Hospitalization variables especially the four RAC regions yielded a very reasonably good fit model with a R^2 value of 0.590900.

The various coefficients associated with the explanatory variables (i.e. the hospitalization variables) in the model are as shown below.

Parameter		Estimate		Std Error	t value	Pr>|t|
Intercept		24816.47	B	360.959158	70.58	<.0001
MEDICAID		-4158.45	B	41.8989	-42.19	<.0001
MEDICARE		-2665.8	B	43.21097	-26.24	<.0001
RACE	**1**	-3297.3	B	282.719206	2.05	<.0001
RACE	**2**	-6189.68	B	303.843927	-3.93	<.0001
RACE	**3**	-5848.72	B	302.094575	-5.89	<.0001
RACE	**4**	948.7444	B	355.593261	-0.94	<.0004
RACE	**5**	-8997.27	B	1102.500758	-4.08	<.0001
RACE	**6**	0	B	.	.	.
FEMALE	**0**	2061.019	B	99.407245	31.76	<.0001
FEMALE	**1**	0	B	.	.	.
RACRegion	**1**	-32632.1	B	121.747189	-282.1	<.0001
RACRegion	**2**	-24672.7	B	212.987642	-127.28	<.0001
RACRegion	**3**	-22146.6	B	231.288061	-90.22	<.0001
RACRegion	**4**	0	B	.	.	.
AGE		-14.5172		2.860885	-0.34	<.0001
LOS		4724.663		6.666881	729.08	<.0001
NPR		10228.5		24.804388	539.65	<.0001
NDX		1108.453		10.932351	75.03	<.0001
DRG		-24.023		0.537852	-93.37	<.0001
DISPUNIFORM		361.2778		11.688359	47.53	<.0001
ASOURCE		-946.323		24.835887	39.58	<.0001

4.3.3. Multiple Linear Regression Model – 2011

Dependent Variable: TOTCHG Total charges (cleaned)					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	19	4.7022022E15	2.4748433E14	81124.9	<.0001
Error	1.3E6	3.9582227E15	3050658022.1		
Corrected Total	1.3E6	8.6604249E15			

R-Square	Coeff Var	Root MSE	TOTCHG Mean
0.542953	133.9324	55232.76	41239.27

The Multiple Linear Regression Model for the Year 2011 for Total Charges as the Outcome to be predicted by the various Hospitalization variables especially the four RAC regions yielded a very reasonably good fit model with a R^2 value of 0.54295.

The various coefficients associated with the explanatory variables (i.e. the hospitalization variables) in the model are as shown below.

Parameter		Estimate		Std Error	t value	Pr>|t|
Intercept		25477.71739	B	360.959158	70.58	<.0001
MEDICAID		-5788.72	B	41.8989	-42.19	<.0001
MEDICARE		-3753.4	B	43.21097	-26.24	<.0001
RACE	1	580.5354	B	282.719206	2.05	<.0001
RACE	2	-1195.04	B	303.843927	-3.93	<.0001
RACE	3	-1778.34	B	302.094575	-5.89	<.0001
RACE	4	-334.374	B	355.593261	-0.94	<.0001
RACE	5	-4498.23	B	1102.500758	-4.08	<.0001
RACE	6	0	B	.	.	.
FEMALE	0	3156.75	B	99.407245	31.76	<.0001
FEMALE	1	0	B	.	.	.
RACRegion	1	-34344.4	B	121.747189	-282.1	<.0001
RACRegion	2	-27109.1	B	212.987642	-127.28	<.0001
RACRegion	3	-20866.4	B	231.288061	-90.22	<.0001
RACRegion	4	0	B	.	.	.
AGE		-0.96843		2.860885	-0.34	<.0001
LOS		4860.666		6.666881	729.08	<.0001
NPR		13385.74		24.804388	539.65	<.0001
NDX		820.2627		10.932351	75.03	<.0001
DRG		-50.2194		0.537852	-93.37	<.0001
DISPUNIFORM		555.5423		11.688359	47.53	<.0001
ASOURCE		982.914		24.835887	39.58	<.0001

4.3.4. Multiple Linear Regression Model for Length of Stay (LOS)

The Multiple Linear Regression modeling technique which resulted in a very good fit for explaining the Total Charges based on the RAC Regions and the other significant hospitalization variables did not yield satisfactory results for the Length of Stay as can been observed from the results below for one of the years of analysis.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	15	35773924.6	2384928.3	60110.3	<.0001
Error	3.9E6	154830598.3	39.7		
Corrected Total	3.9E6	190604522.9			

R-Square	Coeff Var	Root MSE	LOS Mean
0.187687	132.1702	6.298876	4.765729

The R^2 value of 0.1877 indicates that only around 18 percent of the total variation in the Length of Stay could be explained in terms of the independents hospitalization variables and the RAC Regions. This is to be expected because the LOS is not a linear variable with a zero as a possible value for the inpatient data that was considered in this study since by definition, an inpatient is one who is admitted for a minimum of one day. Since similar R^2 scores were obtained for all the five years of analysis it was therefore abandoned as a source of information regarding prediction based on the four RAC Regions. However, since the claims recovery process is more intimately connected with the Total Charges incurred in the patient discharges the predictive models developed and

reported before are seen to be a significant contribution of this dissertation warranting further studies into the use of predictive analytics in the RAC claims recovery process.

CHAPTER V

SUMMARY AND CONCLUSIONS

The overall goal of the study was to (1) to design an appropriate analytical model to explain the operations of the RAC process and identify the hospitalization factors that affect the efficient recovery of claims (2) to formulate a predictive model by using HCUP's Nationwide Inpatient Sample datasets to help predict those hospitalization factors above affecting the RAC claims recovery process, and (3) to determine other relevant hospital, regional and patient related variables that play a statistically significant role in both the RAC and the Hospitalization Outcomes Models.

To meet the objectives above data was extracted from both the RACTrac Website and Reports (for developing the RAC Process Model) and the HCUP Nationwide Inpatient Sample (NIS) database. Several analytical models currently in vogue in both health and finance were investigated and it was decided to adopt a Logic Model to describe the RAC claims recovery process and with its help identified the hospitalization factors related to the claims and payment issues. Secondly the Multiple Linear Regression Model was found to be the most suitable predictive model type for the hospitalization factors identified from the RAC Logic Model. Lastly several descriptive and inferential statistics (ANOVA, Chi-Square) were employed to infer relationships among several patient and hospital variables with the RAC regions and their outcomes.

Both Length of Stay (LOS) and Total Charges were found to be intimately related to the RAC claims recovery process and accordingly they both were employed in the

development of the Multiple Linear Regression Model with several independent variables such as DRG, RAC region, Payer type (Medicare, Medicaid, Private), Number of Diagnoses and Number of Procedures resulted in a reasonably good fit (54 % to 59 %) of the model in explaining the variance of the outcome of Total Charges and not a very good fit for the LOS which was expected since LOS is not a linear variable and subject to too many constraints and hence not easily predictable. The ANOVA Tests revealed several interesting relationships between the independent variables listed above and the RAC regions with implications of import for the RAC claims recovery process. RAC Region 4 incurred a greater cost with a shorter hospital stay regardless of there being slightly lesser number of diagnoses on average per patient and slightly above average number of procedures conducted on such patients compared to the other 3 regions. This could perhaps be explained due to RAC region 4 having more large sized hospitals and many more of teaching type hospitals than non-teaching type with possibility of innovative procedures and availability of specialists in such hospitals requiring shorter stay but effective diagnosis, treatment and discharge. These findings were further corroborated by the model that ensued from the Multiple Linear Regression technique which also revealed that RAC Region 4 would have significantly more Total Charges compared to other three regions. This has far reaching implications in what can be predicted for future claims recovery and also studies into characterization of the hospitalization costing process, reimbursement, fraud detection and healthcare resource allocation across the RAC regions.

Nearly one-third of all health care spending in the United States is attributed to inpatient hospital services[39]. Between 1997 and 2010, aggregate inflation-adjusted hospital costs

grew by 3.8 percent annually[40]. Inpatient hospital costs vary substantially by condition[41]. Hospital costs also vary by other factors, such as patient age and primary payer[42].

The HCUP State Inpatient Databases (SID) from 2003 to 2011 include about 305 million inpatient discharges from 47 States. Despite the major benefits that come from the use of the HCUP database, there are also studies that stress the need for careful consideration that must be used in order to prevent bias or underreporting of certain variables

Limitations of the HCUP database include the following:

• Limited availability of clinical data

• Possible bias from ICD-9-CM coding inaccuracies

• Deficiency of representation of all hospital types

• Deficiency of information on revenue or cost

• Varied data elements between different states

Although HCUP data provides a current view of healthcare across the four RAC regions in terms of ICD-9-CM codes. The use of ICD9 codes has a few problems. Validity of information on procedures in administrative discharge data appears to be related to type of procedures[43]. Major procedures that are usually performed in operating rooms are reasonably well-coded[44]. However minor procedures that are routinely performed on wards or in radiology departments are generally undercoded[45].

The NIS data are not patient level data and there are no identifiers to connect multiple admissions, therefore individuals who are readmitted to the hospital will be present

multiple times in the analyses. Currently there is not a method to distinguish these as the same patient with the variables collected in NIS.

The sample used in this study was limited to only including data from 44 states so there could be sampling bias when using it to derive national estimates.

Mandated under the Affordable Care Act to meaningfully share its massive volume of data with healthcare stakeholders, the Centers for Medicare and Medicaid Services continues to ramp up its data collection, analysis and dissemination[46, 47]. For instance, in one study it shows the per capita cost down to the county level for individuals with 6 or more chronic conditions. Medicare now offers provider utilization and payment, hospital charge data, and prescribing data on more than one million Medicare providers. Data analytics however is on the rise with several businesses adopting it for improving performance, cost reduction and efficiency[48, 49, 50].

Predictive models could be developed for each RAC region such as probability of a length of stay for specific diseases. The predictive models will assist the RAC contractors in targeting specific states in their pursuit of over and under payments. Further research could compare the ICD- 9-CM codes to the newly introduced ICD-10-CM codes to billable charges for each RAC region.

The Centers for Medicare and Medicaid have only recently established a data analytics collaborative which provides a forum for states to consider how data analytics can be used to evaluate Medicaid outcomes[51]. Data provided for the states could provide valuable information to CMS on Medicaid program operations. In July 2016, CMS collected $42 billion in Medicaid and Medicare fraud using predictive analytics[52]. In this

context, the present work reported in here is indicative of the further use of predictive analytics in predetermination of claims and their possible variations across the RAC regions and thus enable a faster and possibly more accurate identification of erroneous billing and/or fraud limiting the processing time, resources and money expended in the effort.

Lastly this study is significant because it demonstrates the validity of the use of analytical models such as Logic Model and the Multiple Linear Regression Model in predicting Hospitalization Outcomes of interest to not only the RAC claims recovery process relevant to this study but also in other health administrative settings involving planning of budget and resource allocation. The complex process of RAC claims recovery mechanism has been duly modeled by the Logic Model technique thus making it available for future configuration modification and studies into their effect on the claims recovery process. To date there is no study in the literature on both the Logic Model approach to characterizing the RAC Claims Recovery Process and the Multiple Linear Regression Modeling to determine the Total Charges using the RAC Regions along with other explanatory hospitalization variables making both the models as novel contributions of this dissertation research study.

REFERENCES

1. Centers for Medicare & Medicaid Services (CMS), Public Law No: 108-173 (12/08/2003) Medicare Prescription Drug, Improvement, and Modernization Act of 2003 - Title III: Medicare Prescription Drug Benefit (Sec. 306)

2. Centers for Medicare & Medicaid Services (CMS), RCF-CMS-2007-0022, Office of Acquisition and Grants Management, Statement of Work for the Recovery Audit Contractor Program, J-I RAC SOW – Amendment 1.

3. Locke C, Sheehy AM, et al. “Changes to inpatient versus outpatient hospitalization: Medicare's 2-midnight rule”, Journal of Hospital Medicine, 2015, 10(3):194-201.

4. Harrison JP, Barksdale RMJ, “The impact of RAC audits on US hospitals”, Health Care Finance. 2013, 39(4):1-14.

5. Ollapally VM, “Recovery audit contractors: an update”, Bulletin of American College of Surgeons, 2010, 95(8):8-10

6. Harris SL, Kelly J, “Building Clarity on the Two-Midnight Rule”, Journal of AHIMA, 2015, 86(10):68-70

7. Ollapally VM, “CMS eases restrictions imposed under the Two-Midnight Rule”, Bulletin of American College of Surgeons, 2016, 101(3):43-44.

8. Goedert J, “Sorting through Medicare RAC audits”, Health Data Manager, 2011, 19(3):34-6, 38-9

9. Feehan PR, Brod BA, “Preparing for an RAC audit under the Affordable Care Act”, Cutis, 2013, 91(3):121-2

10. Lubell J, “AHA tracking RAC impact”, Modern Healthcare, 2009, 39(18):17-18

11. Zigmond J , “RAC appeals data sought. 85% overturned in favor of providers, AHA says”, Modern Healthcare, 2011, 41(18):8-9

12. Kaplan L, “A primer on Medicare and changes resulting from the Patient Protection and Affordable Care Act”, Nurse Practice, 2013, 38(2):8-9

13. Satiani B, “A Medicare primer”, Journal of Vascular Surgery, 2009, 50(2):453-60

14. Davidson BN, “Designing health insurance information for the Medicare beneficiary: a policy synthesis”, Health Services Research. 1988, 23(5):685-720

15. Daly R, “Medicaid RAC overkill. Reviews of medical treatment worry hospitals”, Modern Healthcare, 2011, 41(38):8-9

16. Kaplan L, “A primer on Medicaid and its expansion in 2014”, Nurse Practice, 2012, 37(12):7-8

17. Omachonu V.K, Suthummanon S., Akcin M., Asfour S., Predicting length of stay for Medicare patients at a teaching hospital Health Services Management Research, 17 (1) (2004), pp. 1–12

18. Taheri, D.A. Butz, Greenfield, L.J., Length of stay has minimal impact on the cost of hospital admission Journal of the American College of Surgeons, 191 (2) (2000), pp. 123–130

19. Lagoe R., Pernisi L., Luziani M., Littau S., Addressing hospital length of stay outlier patients: a community wide approach, Advances in Bioscience and Biotechnology (2014) 2014.

20. https://www.cms.gov/Regulations-and-Guidance/Legislation/DeficitReductionAct/Downloads/cmip2014.pdf Accessed November 2, 2016

21. Showalter JS. Healthcare Finance Management. 2014 Sep;68(9):46-50

22. Tomes JP. Journal of Health Care Finance. 2013 Summer;39(4):28-35

23. Stanowski AC, Simpson K, White A. Journal of Healthcare Management 2015 Jul-Aug;60(4):268-85

24. Moore KD, Eyestone K, Coddington DC. How business intelligence can improve value. Healthcare Financial Management. 2012; 66(10): 112-4

25. Agency for healthcare research and quality (AHRQ). Web site. www.ahrq.gov Updated 2013. Accessed October 11, 2016

26. Healthcare cost and utilization project (HCUP). Healthcare Cost and Utilization Project (HCUP) Databases Web site. http://www.hcup-us.ahrq.gov/nisoverview.jsp, Updated 2013. Accessed October 11, 2016

27. Loether HJ, McTavish DG., *Descriptive and Inferential Statistics: An Introduction*. Boston, MA: Allyn and Bacon; 1993

28. Heeringa, S.G., West, B.T., Berglund, P.A. , Applied Survey Data Analysis. 2010, Taylor and Francis Group, LLC.: Boca Raton, FL

29. Pallant, J., SPSS Survival Manual. 2005, New York, NY: Open University Press

30. Tabachnick, B.G., Fidell, L.S., Using Multivariate Statistics. 2007, Boston, MA: Pearson Education, Inc

31. Millar, A., Simeone, R., Carnevale, J. (2001). Logic models: A systems tool for performance management. Evaluation and Program Planning, 24, 73-81

32. Julian, D. (1997). The utilization of the logic model as a system level planning and evaluation device. Evaluation and Program Planning, 20 (3), 251-257

33. Jordan, G., Mortensen, J. (1997). Measuring the performance of research and technology programs: A balanced scorecard approach. Journal of Technology Transfer, 22 (2), 13-20

34. Moyer, A., Verhovsek H., et al. (1997), "Facilitating the shift to population-based public health programs: innovation through the use of framework and logic model tools." Canadian Journal of Public Health 88(2): 95-98.

35. Rugg, D., Buehler B., et al. (1999). "Evaluating HIV prevention: a framework for national, state, and local levels." American Journal of Evaluation 20(1): 35-36.

36. Centers for Medicare & Medicaid Services (CMS), https://www.cms.gov/Research-Statistics-Data-and-Systems/Monitoring-Programs/Medicare-FFS-Compliance-Programs/Medical-Review/Downloads/Hospital_Appeals_SettlementFAQs_10312014_508.pdf, Accessed: July 15, 2016.

37. http://kff.org/medicare/issue-brief/the-facts-on-medicare-spending-and-financing/, Accessed July 20, 2016.

38. Schindler, C.a., *Business Research Methods*. 2005, McGraw-Hill: Boston, MA

39. Kashihara D, Carper K. National health care expenses in the U.S. civilian noninstitutionalized population, 2009. MEPS Statistical Brief #355. January 2012. Agency for Healthcare Research and Quality, Rockville, MD.

40. Pfuntner A, Wier LM, Steiner C. Costs for hospital stays in the United States. HCUP Statistical Brief\#146. January 2013. Agency for Healthcare Research and Quality, Rockville, MD. http://www.hcupus.ahrq.gov/reports/statbriefs/sb146.pdf. Accessed November 25, 2016

41. Dieleman JL et al., "US Spending on Personal Health Care and Public Health 1996-2013", Journal of American Medical Association, 2016; 316(24):2627-2646.

42. Centers for Medicare & Medicaid Services (CMS), http://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/nationalhealthexpenddata/nhe-fact-sheet.html, Accessed November 25, 2016

43. Quan H, Parsons GA, Ghali WA. "Validity of procedure codes in International Classification of Diseases, 9th revision, clinical modification administrative data", Medical Care , 2004, 42(8):801-9.

44. Maclean D., Younes HB, Forrest M, Towers HK., "The accuracy of real-time procedure coding by theatre nurses: a comparison with the central national system", Health Informatics Journal, 2012, 18(1):3-11.

45. Cacace, M and Schmid, A, "The role of diagnosis related groups (DRGs) in healthcare system convergence", BMC Health Services Research, 2009, 11-18.

46. http://www.healthdatamanagement.com/news/cms-data-chief-urges-use-of-medicare-data-analytics-tools Published: July 14, 2015. Accessed Dec. 1, 2016

47. Anvita Health. Clinical Analytics: Can Organizations Maximize Clinical Data? HIMSS Analytics Website. http://www.himssanalytics.org/docs/clinical_analytics.pdf, , Accessed December 22, 2016

48. Ferranti JM, Langman MK, Tanaka D, McCall J, Ahmad A. Bridging the gap: leveraging business intelligence tools in support of patient safety and financial effectiveness. Journal of American Medical Informatics Association, 2010; 17(2): 136–143.

49. Fireman J., The Vital Role of Business Intelligence in Improving Healthcare Delivery. PICIS Website., http://www.picis.com/~/media/resource%20library/whitepapers/white_paper_business_intelligence.ashx, Accessed December 3, 2016.

50. Using Analytics to Navigate Health Care Reform. Five strategic imperatives for providers facing dramatic transformations in the US health care system – and why analytics is the key to executing them. SAS Website., http://www.healthcareitnews.com/sites/default/files/resource-media/pdf/105715_0412.pdf, Accessed October 2, 2016.

51. Brennan N, Oelschlaeger A, Cox C, Tavenner M, "Leveraging the big-data revolution: CMS is expanding capabilities to spur health system transformation", Health Affairs, 2014 33(7):1195-202

52. "CMS snags $42 billion in Medicare and Medicaid fraud with predictive analytics", http://www.healthcareitnews.com/news/cms-snags-42-billion-medicare-and-medicaid-fraud-predictive-analytics, Accessed July 25, 2016.