DATA ANALYTICS OF CODIFIED PATIENT DATA: IDENTIFYING FACTORS INFLUENCING CODING TRENDS, PRODUCTIVITY, AND QUALITY

by

Zahraa M. Alakrawi

Bachelor of Science, University of Dammam, 2007

Master of Science, University of Pittsburgh, 2012

Submitted to the Graduate Faculty of

School of Health and Rehabilitation Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2017

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCES

This dissertation was presented

by

Zahraa M. Alakrawi

It was defended on

March 31, 2017

and approved by

Mervat Abdelhak, Ph.D., RHIA, FAHIMA
Chair and Associate Professor, Health Information Management

Melanie M. Hughes, Ph.D.
Associate Professor, Sociology

Patricia Anania Firouzan, MSIS, RHIA
Assistant Professor, Health Information Management

Patty Sheridan, MBA, RHIA, FAHIMA,
Ciox Health, Senior Vice President, HIM Business Development

Dissertation Advisor: Valerie Watzlaf, Ph.D., RHIA, FAHIMA,
Associate Professor, Health Information Management

Copyright © by Zahraa M. Alakrawi

2017

DATA ANALYTICS OF CODIFIED PATIENT DATA:

IDENTIFYING FACTORS INFLUENCING CODING TRENDS, PRODUCTIVITY, AND QUALITY

Zahraa M. Alakrawi, Ph.D.

University of Pittsburgh, 2017

Cost containment and quality of care have always been major challenges to the health care delivery system in the United States. Health care organizations utilize coded clinical data for health care monitoring, and reporting that includes a wide range of diseases and clinical conditions along with adverse events that could occur to patients during hospitalization. Furthermore, coded clinical data is utilized for patient safety and quality of care assessment in addition to research, education, resource allocation, and health service planning.

Thus, it is critical to maintain high quality standards of clinical data and promote funding of health care research that addresses clinical data quality due to its direct impact on individual health outcomes as well as population health. This dissertation research is aimed at identifying current coding trends and other factors that could influence coding quality and productivity through two major emphases: (1) quality of coded clinical data; and (2) productivity of clinical coding. It has adopted a mix-method approach utilizing varied quantitative and qualitative data analysis techniques. Data analysis includes a wide range of univariate, bivariate, and multivariate analyses.

Results of this study have shown that length of stay (LOS), case mix index (CMI) and DRG relative weight were not found to be significant predictors of coding quality. Based on the qualitative analysis, history and physical (H&P), discharge summary, and progress notes were

identified as the three most common resources cited by Ciox auditors for coding changes. Also, results have shown that coding productivity in ICD-10 is improving over time. Length of stay, case mix index, DRG weight, and bed size were found to have a significant impact on coding productivity. Data related to coder's demographics could not be secured for this analysis. However, factors related to coders such as education, credentials, and years of experience are believed to have a significant impact on coding quality as well as productivity. Linking coder's demographics to coding quality and productivity data represents a promising area for future research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I have always been longing to this moment to arrive, to fill my heart with joy, to finally keep it as a precious memory in my mind and inside my heart. It has been such a long yet a pleasant journey. Through this journey, I have met so many people who made me the person I am today, including the nice strangers who I ran into at Starbucks.

I have been longing to this moment, to write the "preface" of my dissertation, to acknowledge all the great people who believed in me and brought this dissertation to life. First, I would like to express my sincerest thanks and deepest gratitude to my dissertation advisor and committee chair, Dr. Valerie Watzlaf, for her dedicated mentoring, never-ending support, and wonderful guidance in the last five years. Thank you for teaching me how to accept my weaknesses, learn from my mistakes, and grow and cherish every opportunity to be a better version of me. I would like to thank my dissertation committee for their guidance and support throughout this journey: Dr. Mervat Abdelhak, Dr. Melanie Hughes, Patricia Anania Firouzan, and Patty Sheridan.

I would like to express my heartfelt thanks, gratefulness, and gratitude for Batool Al Yousif & Moosa Alakrawi, my beloved parents, for their endless love, support, encouragement, and prayers. Also, I would like to thank my lovely sisters, Mariam and Khatoon, for always being by my side. Last, but not least, I am especially thankful to my beloved husband, Raid Al Jomaia. Thank you for your love, your care, and unlimited support. Thank you for all meals, for all the flowers, for lifting me up when I am down, and for all the jokes you tell to make me smile.

From the bottom of my heart, thank you to everyone who contributed to this dissertation.

1.0 INTRODUCTION

Coding constitutes one of the fundamental functions in the field of Health Information Management (HIM) (AHIMA, 2016). Coding can be defined as “the process of translating descriptions of diseases, injuries, and procedures into numeric or alphanumeric designations” (AHIMA, 2013). In this era of electronic health records (EHRs) and based on the need for electronic transactions, coders need not only to be familiar with the code assignment process but also with mapping among different clinical nomenclature and terminology (DeAlmeida, 2012; Giannangelo, 2011; Alakrawi, 2016).

Clinical coders, at least, should have the knowledge and skills that are needed to deal with the Health Insurance Portability and Accountability Act (HIPAA) code sets. HIPAA standard code sets include the following: International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM/PCS); Current Procedural Terminology (CPT-4); Code on Dental Procedures and Nomenclature (CDT); National Drug Codes (NDCs); and Healthcare Common Procedure Coding System (HCPCS) (AHIMA, 2016; CMS, 2016).

ICD-10-CM is the system used to collect morbidity statistics in the United States (CMS, 2016). It constitutes the basis for the U.S. reimbursement systems, particularly for the Inpatient Prospective Payment System (IPPS) developed by the Centers for Medicare and Medicaid Services (CMS). The IPPS is used by CMS to finance inpatient services rendered to Medicare and Medicaid beneficiaries (CMS, 2016). The United States implemented ICD-10-CM/PCS on October 1, 2015

(CMS,2016; Miller, 2016). ICD-10-CM/PCS includes both diagnoses and procedures code sets. Implementation of ICD-10-CM/PCS was crucial to replace the outdated ICD-9-CM coding system that had been in use since 1979 (Rode, 2013; Alakrawi, 2016).

However, the World Health Organization's (WHO) ICD-10 has been used since 1990 in the United States to collect mortality statistics; basically, to code death certificates and collect causes of death (NCHS, 2016; WHO, 2017). Based on the WHO's ICD-10, the United States has developed ICD-10-CM/PCS for purposes related to morbidity and public health. Furthermore, coded clinical data has a considerable impact on the health care industry for assessing clinical outcomes, conducting research, promoting education, and planning health services (Alakrawi, 2016; Avril & Bowman, 2012; Glenn, 2013; Linder, 2016; Rode, 2013; Walker, 2012). As mentioned earlier, coding is known to serve as the foundation of the reimbursement system in the United States. Therefore, there has been a rising demand to clinical data quality to meet reimbursement requirements (Alakrawi, 2016; Land, 2016). In addition, there has been an ever-increasing demand to improve ICD-10 coding productivity standards to maintain healthy revenue and cashflow (Linder, 2016; Martin, 2016; Stanfill, 2015).

There are many critical reasons to address the issue of clinical coding quality and productivity. Codes at the individual-level reflect the patients' health status and are used as a communication tool between different healthcare providers. Codes are also used in conjunction with other items for reimbursement of services rendered to patients during their episodes of care. At the public health level, clinical codes are used to collect mortality as well as morbidity statistics that are further used for assessing population health in addition to health services planning and monitoring (Alakrawi, 2016; CMS, 2016; NCHS, 2016).

There any many forces that can potentially influence coding quality and productivity. ICD-10-CM codes are used for patient safety and quality of care monitoring. Specifically, ICD-10-CM codes are used by acute care facilities for reporting of adverse events that could happen to patients during hospitalization. These codes are further used by governmental organizations such as the Centers for Medicare and Medicaid Services (CMS) and the Agency for Healthcare Research and Quality (AHRQ) for assessing patient safety, and quality of care through performance indicators used to compare hospital performance across the country. The results of these assessments are frequently released to the public so that healthcare consumers can make informed decisions about their own health and safety.

Clinical coded data are also used for public health reporting and health services planning. Particularly, ICD-10-CM data is used to collect population health statistics at the national and international levels. The Centers for Disease Control and Prevention (CDC) use coded clinical data in ICD-10-CM to identify the leading causes of death in the U.S. in addition to other measures of population health status. At the global level, the World Health Organization (WHO) utilizes data collected from all different countries for reporting of the leading causes of death in the globe. This type of reporting that is frequently performed is heavily dependent upon quality of data collected at the primary source. Consequently, health service planning and research priorities are set based on priorities identified through the aggregate coded data.

The compliance date of implementation of ICD-10-CM/PCS was October 1, 2015 in the United States and many organizations had been reluctant to meet the deadline. The American Medical Association (AMA) and its regional associations had tried to delay the implementation of ICD-10-CM/PCS until 2017 (Health Data Management, 2014). However, the American Health Information Management Association had reaffirmed its stance and commitment to the actual deadline. In 2012, “CMS estimated the cost to the healthcare industry of a one year delay to be as much as $6.6 billion, or approximately 30 percent of the $22 billion that CMS estimated had been invested or budgeted for ICD-10 implementation” (Butler, 2014; Butler, 2016).

Implementation of ICD-10-CM/PCS had motivated healthcare providers and organizations to focus on the quality as well as productivity of their coded data as coding became more complex under the new system.

Clinical documentation improvement (CDI) can also have a positive impact on coding quality as well as productivity. In fact, CDI programs could improve clinical documentation which can subsequently contribute to quality of the coded data. Furthermore, accurate and complete documentation can help reduce physician queries that are usually initiated by coders as they try to assign the appropriate codes based on the patient chart (Combs, 2016; Land, 2016).

Financial incentives in terms of payment maximization and efficient utilization of resources will have a significant impact in promoting coding quality and productivity. For example, “the American Recovery and Reinvestment Act of 2009 authorizes CMS to provide incentive payments to eligible professionals (EPs) and hospitals who adopt, implement, upgrade, or demonstrate meaningful use of certified electronic health record (EHR) technology” (HealthIT.gov, 2016; Houser & Meadow, 2017; Linder, 2016). Such financial incentives have contributed to higher coding quality and productivity standards through automation of coding workflow and continuous improvement of coding software applications.

Audit programs that look for compliance and coding issues have further contributed to an ever-increasing emphasis on coding quality and subsequently coding productivity. Conducting internal as well as external audits has been a major trend in health care. This is basically due to the health care organizations’ efforts to meet compliance requirements demanded by government auditors such as the CMS's Medicare Recovery Audit Contractors (RACs). In addition to meeting auditing standards, health care organizations should meet higher productivity standards for coding due to its direct link with reimbursements, claim submission, cashflow, and revenue cycle in general (Godbey-Miller, 2016; Martin, 2016)

Along with coding, healthcare providers are expected to comply with other federal laws and regulations. A brief discussion of some federal laws and regulations is provided in chapter 4.

It is inevitable to maintain high quality and productivity standards of coded clinical data and promote funding of health care research that addresses clinical coding due to its direct impact on individual health outcomes as well as population health. With the rapid adaptation of health information technology (HIT), there is a rising demand for effective and data-driven decision-making strategies (Houser & Meadow, 2017). Coded clinical data needed for such decision-making should be reliable and available to users at times of decision making. Therefore, this dissertation research aims at identifying current coding trends and other factors that could influence coding quality and productivity through two major emphases: (1) quality of coded clinical data; and (2) productivity of clinical coding. Figure 1 presents the conceptual framework of the literature review that will be followed in this dissertation research.

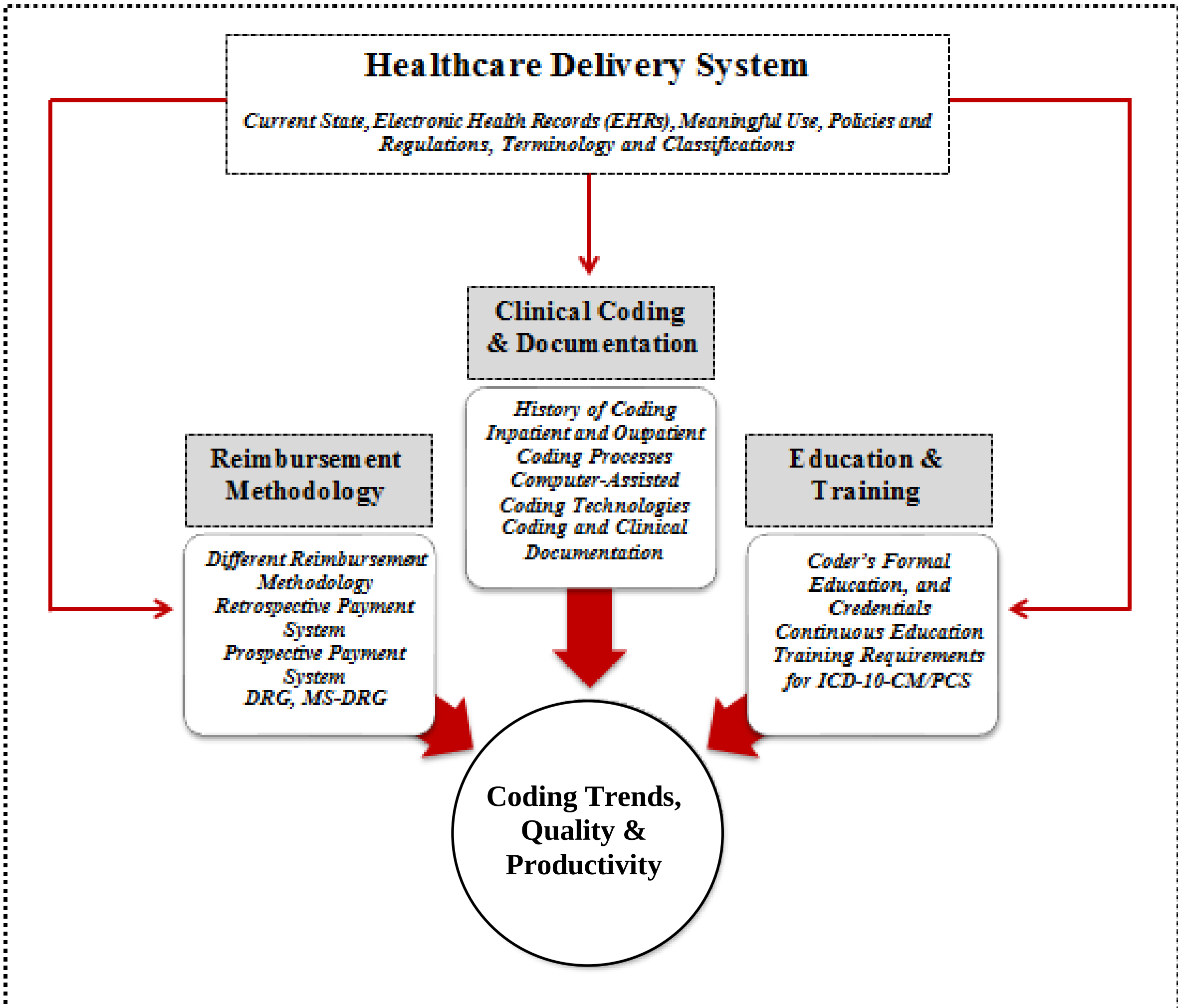


Figure 1: Conceptual Framework to Literature Review

This study aims at identifying determinants of coding quality and productivity through the following:

Specific Aim I: Identify factors that could influence coding accuracy:

(1) Length of stay (LOS)

(2) Case mix index (CMI)

(3) DRG relative weight

(4) MS_DRG categories that are more often impacted by coding discrepancies

(5) Coding errors at the major digit level versus the minor digit level

Specific Aim II: Identify documentation discrepancies that could influence coding quality.

Specific Aim III: Identify the impact of coding errors on CMI and hospital's payment.

Specific Aim IV: Identify individual and facility-related factors that could influence coding productivity:

(1) Length of stay (LOS)

(2) DRG relative weight

(3) Case mix index (CMI)

(4) Facility bed capacity (bed size)

(5) Teaching status

(6) Trauma status

Specific Aim V: Explore the relationship between coding productivity and coding quality

Specific Aim VI: Develop a predictive model to predict coding productivity and coding quality based on the individual and facility-related factors.

2.0 HEALTHCARE VOCABULARY, TERMINOLOGY, AND CLASSIFICATION

A very important aspect in this context is to discuss the differences between "vocabulary", "terminology", and "classification systems." In general, clinical vocabularies, terminologies, and classification systems, are a "structured list of terms which together with their definitions are designed to describe unambiguously the care and treatment of patients." (AHIMA, 2016, Alakrawi, 2016). They are used to cover diseases, procedures, diagnoses, findings, medications, and other items used to "support recording and reporting a patient's care at varying levels of detail, whether on paper or, increasingly, via an electronic medical record (EMR)." (AHIMA, 2016; De Lusignan, 2005). Figure 2 illustrates the levels of detail given by vocabularies, classification systems and terminologies (Alakrawi, 2016; HL7, 2009).

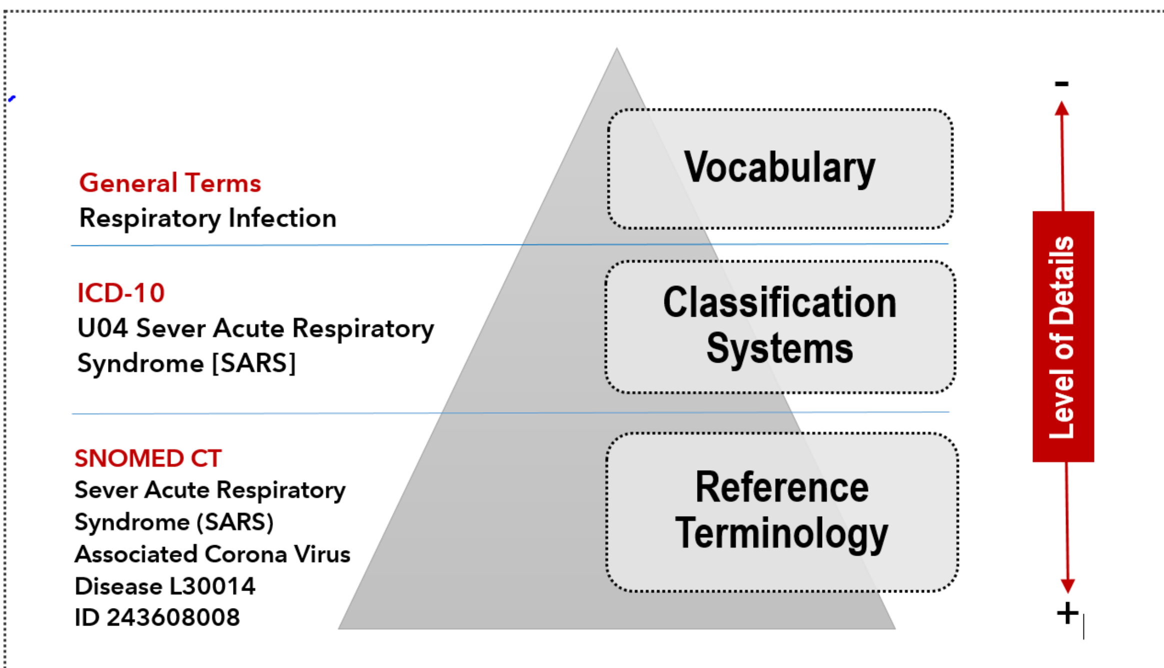


Figure 2: The level of detail given by a vocabulary, classification system, and terminology

Source: (HL7, 2009, De Lusignan, 2005; HISTDO, 2014)

Based on Figure 2, a vocabulary represents general terms about a certain concept (***Respiratory Infection and Inflammation***) with the lowest level of detail. These general terms are used for public communication and can also be adapted for specified fields of science and industry. A classification system can be used to communicate with higher level of detail regarding a certain concept. The general term ***Respiratory Infection*** can be further specified using International Classification of Diseases Tenth Revision (ICD-10):

Respiratory Infection -----------------> *U04 -Severe Acute Respiratory Syndrome [SARS]*

The highest level of detail and specificity can be provided when using a reference terminology such as SNOMED-CT. SNOMED-CT can provide further detail and specificity with respect to ICD-10 classification of ***Severe Acute Respiratory Syndrome [SARS]***:

U04 -Severe Acute Respiratory Syndrome [SARS] ----------->

Severe Acute Respiratory Syndrome (SARS)

Associated Coronavirus Disease L30041

ID 243608008

A vocabulary is "a collection of words or phrases with their meanings" and a classification is "a system that arranges or organizes like or related entities" (Alakrawi, 2016; De Lusignan, 2005; Ginanngelo, 2012). A terminology is "a set of terms representing a system of concepts" (De Lusignan, 2005; Ginanngelo, 2012; IHTSDO, 2016). Further, the ISO (ISO 17115) defines a clinical terminology as a "terminology required directly or indirectly to describe health conditions and healthcare activities". Effective communication of meanings across healthcare settings and disciplines is the main goal of developing healthcare terminologies (IHTSDO, 2016). Thus, different sets of healthcare terminology have been developed by healthcare professionals for use in their areas of clinical specialty (IHTSDO, 2016).

However, in this era of information exchange and ever-increasing use of electronic communication and the EHRs, a need arises for a more controlled and comprehensive set of terminologies that cover all concepts of healthcare (reference terminology) (Alakrawi, 2016; De Lusignan, 2005; IHTSDO, 2016). A reference terminology is defined by the ISO (ISO 17115) as "a terminology containing only concept names as determined by an authorized organization". In general, there are many reasons for needing a vocabulary, terminology, or classification system. Some of these reasons are presented in Table 1 (Giannangelo, 2012).

Table 1: Reasons for needing a vocabulary, terminology, or classification system

	Function	Reasons for needing a vocabulary, terminology, or classification system
1	**Access to complete and accurate clinical data**	• Facilitate electronic data collection at the point of care • Possess the ability to capture the detail of diagnostic studies, history, and physical examinations, visit notes, ancillary department information, nursing notes, vital signs, outcome measures, and any other clinically relevant observations about the patient • Allow many different sites and different providers the ability to send and receive medical data in an understandable and usable manner, thereby speeding care delivery and reducing duplicate testing and duplicate prescribing
2	**Links to medical knowledge**	• Provide organized systems of data collection and retrieval; • link published research with clinical care in order to improve quality of care through outcome measurement

Table 1 (continued)		
3	**Practitioner alerts and reminders and clinical decision support systems**	• Improve the quality of healthcare through the effective use of information found in other information management systems • Allow the computer to manipulate standardized data and find information relevant of individual patients for the purpose of producing automatic reminders or alerts • Permits retrieval of relevant data, information, and knowledge for generating patient-specific assessments or recommendations designed to aid clinicians in making clinical decisions • Provide data to consumers regarding costs and outcomes of treatment options
4	**Research and epidemiological studies and public health**	• Allow collection and reporting of health statistics and ensure a high-quality database for accurate clinical as well as statistical data • Provide data for use in public health monitoring
5	**Healthcare claims reimbursement and management**	• Provide data for use in designing payment systems, determining the correct payment for healthcare services, and identifying fraud and abuse • Make available information for use in improving performance (clinical, financial, and administrative)

2.1.1 HealthCare Terminology

Systematized Nomenclature of Medicine, SNOMED, is a standardized health care terminology which was originally developed from pathology-specific nomenclature called Systematized Nomenclature of Pathology (SNOP). SNOMED is a controlled medical terminology that encompasses diseases, clinical findings, etiologies, procedures, and health outcomes (Alakrawi, 2016; Cornet & Keizer, 2008; IHTSDO, 2016). It can be used by physicians, nurses, allied health professionals, veterinarians, and researchers.

SNOMED is defined by the International Health Terminology Standards Development Organisation (IHTSDO) as “comprehensive clinical terminology that provides clinical content and expressivity for clinical reporting which is comprised of concepts, terms, and relationships with the objective of precisely representing clinical information across the scope of health care” (IHTSDO, 2016). The ownership, maintenance, and distribution of SNOMED was originally the responsibility of the College of American Pathologists (CAP) but this responsibility was transferred to the IHTSDO in 2007 (IHTSDO, 2016).

However, it is useful in this context to discuss how a terminology differs from a classification system. First, terminologies and classifications systems are designed to serve different purposes; a clinical terminology such as SNOMED could be more useful in clinical applications and information retrieval, and research. SNOMED is considered as a global standard due to its wide acceptance and application world-wide which makes it a safe and accurate alternative for clinical communication between healthcare providers (Alakrawi, 2016; Bowman, 2014; IHTSDO, 2016).

In contrast, classification systems such as ICD-9-CM or ICD-10-CM/PCS are intended for classification of clinical conditions and procedures to be used for other applications including statistical reporting and reimbursement (Bowman, 2014; Butler, 2016). A classification system can be less-detailed than a clinical terminology. Per Bowman (2014), “If a researcher wants to know how many patients died with a diagnosis of heart attack last year, ICD-10 is enough. If they want more detail, such as what muscle of the heart was involved, they will need SNOMED CT.”

Nonetheless, SNOMED CT -the most current version of SNOMED- is available at no charge through the National Library of Medicine (NLM). The U.S. license for SNOMED was obtained by the NLM through the Unified Medical Language System (UMLS) project (NLM, 2014; UMLS, 2014). The first edition of SNOMED was published in 1974. However, this edition was based on the Systemized Nomenclature of Pathology (SNOP) that was published by CAP in 1965. Figure 3 provides a summary of the history of SNOMED and its evolution over time.

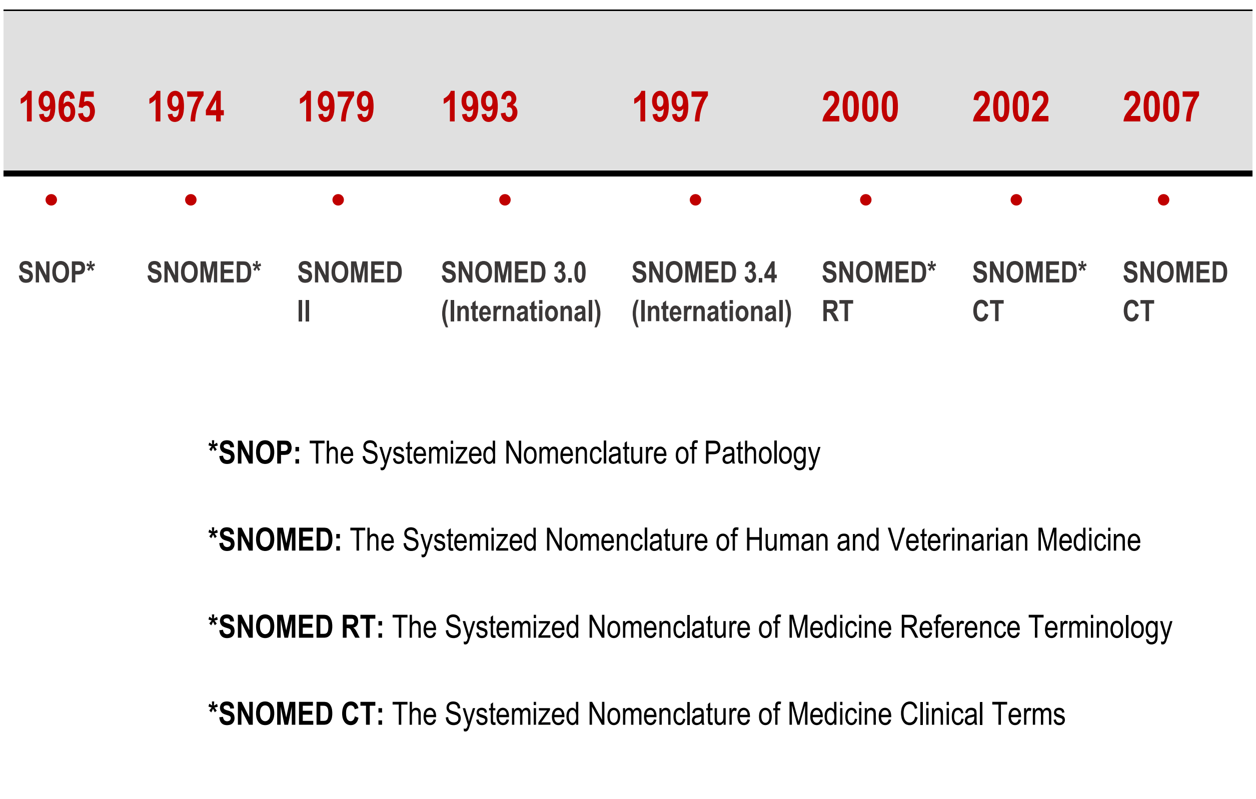


Figure 3: History of SNOMED

Regardless of its continuous evolvement, SNOMED-CT has not been fully utilized in clinical practice and applications (Duarte et al, 2014; IHTSDO, 2016; Lee et al, 2014). Cornet & Kate (2008) provides a literature review of published studies in which SNOMED has been used in different clinical applications. The authors analyzed the use of SNOMED over time as reflected in scientific publications. A Medline literature search using PubMed was performed to select papers in which SNOMED was either the primary or secondary object of the study (study period was from 1966-2006, therefore SNOP was included). Selected papers were further classified based on version of SNOMED, medical domain, time of coding, usage, and type of evaluation (Cornet & Keizer, 2008).

This analysis included 250 papers on SNOMED. However, in many cases it was difficult to determine which version of SNOMED was used. Pathology, nursing, and cancer were the most frequently mentioned medical domains when a specific medical domain is described. There were 163 papers in which SNOMED was the primary object of the study and 87 in which it played a secondary role. Two major subjects were identified for the primary role: (1) comparing SNOMED to other Terminology Systems (TS) - mostly in content coverage; and (2) using SNOMED to illustrate a TS theory. For secondary uses, SNOMED was utilized as an example in most of the cases (Cornet & Keizer, 2008).

Kate introduces a machine-learning method that can be utilized to convert clinical language text into structured representations using SNOMED CT. The author employed the Support Vector Machine (SVM) machine learning in combination with a new kernel specifically designed for this study. The aim of this study was to identify the relationship between clinical phrases and SNOMED-CT to enhance existing capabilities of natural language processing in clinical applications. Using existing datasets, the experimental results demonstrate that the trained system shows an increased performance on the relation-identification tasks- by measuring both recall and precision. The author identifies syntactic analysis of SNOMED CT as a possible area for future work (Kate, 2013)

In another study, Mikroyannidi et al (2012) provide an example of research in which SNOMED was incorporated into a framework to detect syntactic regularities as well as irregularities in ontology. This study specifically focuses on the Web Ontology Language (OWL) representation of SNOMED CT. It concluded that the tested framework can be utilized for quality assurance in ontology. However, application of SNOMED CT shows positive results that can be utilized in healthcare to support clinical application for administrative and direct care purposes (Allones et al, 2014; Daurte et al, 2014; IHTSDO, 2016; Lee et al, 2014; Mikroyannidi et al, 2012).

In general, the literature review reflects increasing utilization of SNOMED in clinical applications and across medical specialties –other than pathology. However, there are no indications of the use of SNOMED for direct care purposes, performance or productivity, and quality audit. Future research to address the effect of terminology systems on the care process and outcomes is needed (Alakrawi, 2016).

2.1.2 Healthcare Classifications

A classification is “a system that arranges or organizes like or related entities” (AHIMA, 2013). Classifications are used to support statistical data across the healthcare system. Thus, the WHO has developed different classification systems that can be integrated to describe different aspects of health. These classification systems can be of three types (Madden, 2008; WHO, 2016):

2.1.2.1 Reference Classifications

These classification systems cover the main parameters of healthcare as well as the healthcare system such as disease, functioning, disability, death, and healthcare interventions. The WHO reference classification systems are products of international agreement between the UN member states. They are used to describe the health experience or the health state of a given person at a particular point in time. Further, they can serve as models in development and revision of other classification systems. Examples are the International Classification of Diseases – 10th revision (ICD-10), and the International Classification of Functioning, Disability, and Health (ICF).

2.1.2.2 Derived Classifications

As the name implies, derived classifications are based upon one or more reference classifications. They are intended to be consistent with the references upon which they were developed and usually to provide additional details in specialized areas. Examples could include specialty-based adaptation of ICD or ICF such as the International Classification of Diseases for Oncology (ICD-O-3) and the ICF Version for Children and Youth (ICF-CY).

2.1.2.3 Related Classifications

These classifications describe important aspects of health or the healthcare system not covered by reference or derived classifications. An example is the International Classification of External Causes of Injury codes (ICECI).

2.1.3 WHO Family of International Classification (FIC)

Per Madden et al (2008), the WHO family is “a suite of classification products that may be used in an integrated fashion to compare health information internationally as well as nationally.” By using such classifications, compilations of consistent measures for comparing health systems within populations over time or between populations at a specific point in time- are facilitated at the national and international level (Madden, 2008; WHO, 2016)

2.1.3.1 Purpose of the WHO-FIC

The purpose of the WHO-FIC is to: (1) improve health through supporting health-related decision making, (2) provide a conceptual framework of health and health-related information domains, (3) provide a common language of communication, (4) facilitate comparison of data within and between countries, health disciplines, services and time, and (5) stimulate health research (WHO, 2016).

2.1.3.2 UN definition of the WHO-FIC

The WHO family of international classifications (WHO-FIC) is comprised of classifications that have been endorsed by the WHO to describe various aspects of the health and the healthcare system in a consistent manner. The classifications may be owned by the WHO or other groups. The purpose of the family is to assist in the development of reliable statistical systems at local, national, and international levels, with the aim of improving health status and health care. The WHO family includes reference, derived, and related classifications.

2.1.3.3 Scope of the WHO Family

The WHO-FIC is a conceptual framework of the healthcare system and factors influencing health. The reference classifications within the WHO-FIC cover the following dimensions: (1) diseases, (2) health problems, (3) body function, (4) body structure, (5) activity, (6) participation, (7) interventions, and (8) environment. More specialty-based or other health areas that are not covered in the reference classifications are included in either derived or related classifications (WHO, 2008). However, when an information gap is identified within the current classification systems, an inevitable need arises to either develop a new classification system or endorse an existing classification system into the WHO family. Figure 4 provides a schematic representation of the WHO-FIC along with some examples.

2.1.4 Use of Vocabulary, Terminology, and Classification Systems

Clinical vocabulary, terminology, and classification systems can be used in the EHR systems as well as administrative applications. Per Giannangelo, "collectively, vocabularies, terminologies, and classification systems provide the common medical language necessary for the future state" of eHIM; electronic, patient-centered, comprehensive, longitudinal, accessible, and credible (AHIMA, 2003; AHIMA, 2016; Giannangelo; 2012).

However, certain vocabulary, terminology, and classification systems are only appropriate for chosen applications or purposes such as documentation of clinical care, public health reporting, providing the data structure for EHRs, interoperability and health information exchange (HIE) (Houser & Meadow, 2017).

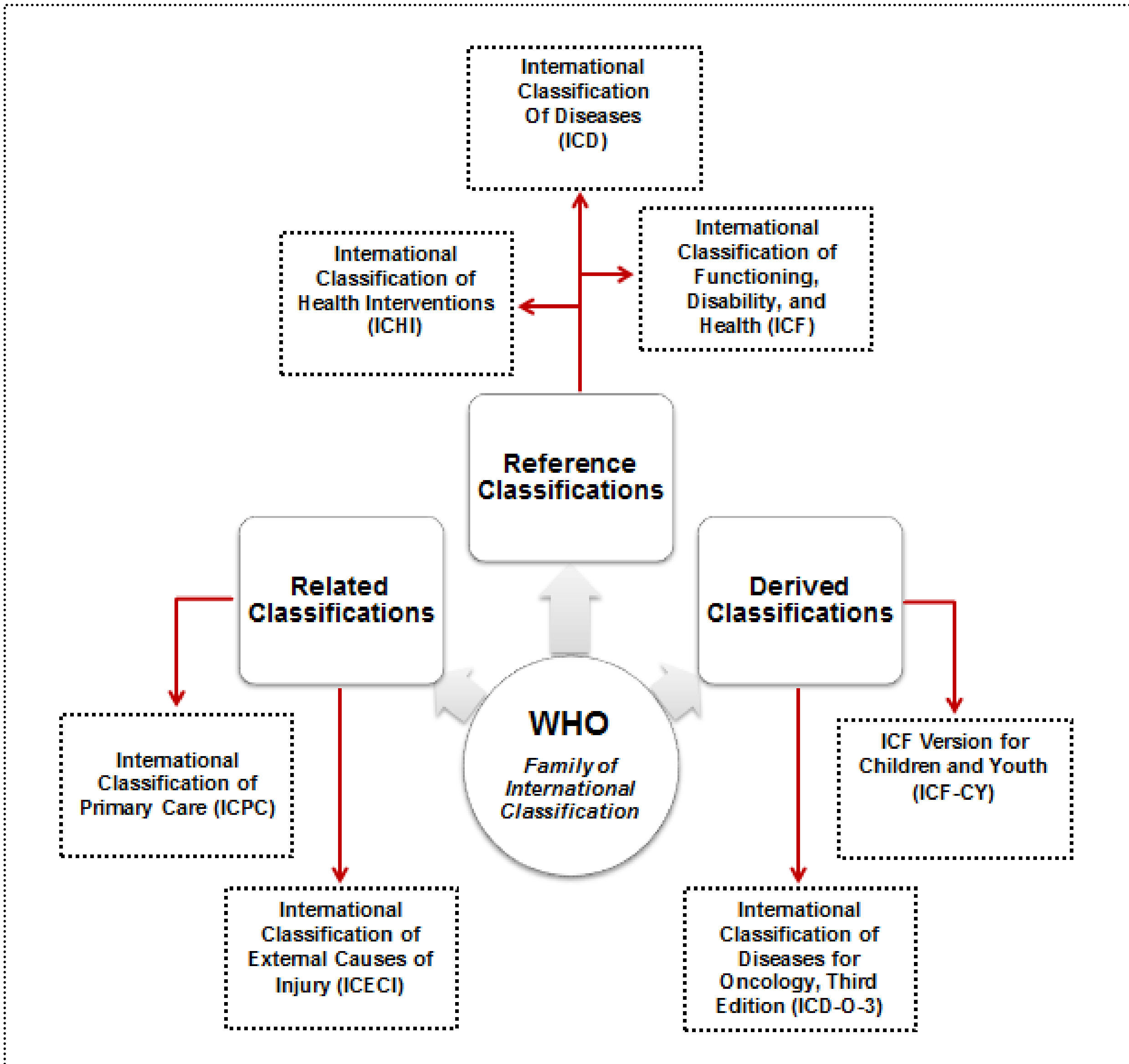


Figure 4: WHO-FIC with Examples

Source of information: (Madden et al, 2008)

The ultimate challenge is how to utilize such vocabularies, terminologies, and classification systems to implement interoperability standards for EHR systems and HIE. Table 2 presents the differences between vocabulary, terminology and classification systems based on the purpose and user (AHIMA, 2016; Giannangelo, 2012; Houser & Meadow, 2017). As illustrated in Table 2, healthcare vocabulary, terminology, and classification systems can be used by different users including consumers, healthcare providers, quality and utilization management personnel, researchers, and other administrative staff (accounting, billing, and coding personnel). Healthcare vocabularies are mainly used to facilitate communication between healthcare providers and consumers at the point of care for data collection purposes.

A more organized system of data collection and retrieval can be provided by utilizing healthcare terminology. This system can promote quality of care through providing a link between published research and clinical care. Furthermore, such systems can support integration of care through allowing effective exchange of clinical information between healthcare providers in different settings. Although terminologies such as SNOMED-CT can be utilized to support real time decision making and retrospective reporting for research and management, such utilization can be hindered by complexity of these systems (Alakrawi, 2016; IHTSDO, 2016).

Classification systems are utilized by a wider spectrum of users in healthcare. They can be used to provide data to consumers on costs, treatment options, and outcomes. Also, classification systems provide a less complex system for data collection and reporting that can be further used for research purposes. Information provided by such systems can be used to improve clinical, financial, and administrative performance through designing effective payment systems, identifying potential fraud and abuse, and ensuring accurate reporting (Alakrawi, 2016).

Table 2. Differences between vocabulary, terminology and classification systems based on chosen goal and users

	Users	Purpose
Vocabulary **Terminology**	*Consumers and Healthcare providers*	Facilitate data collection at the point of care with terms familiar to the user
	Healthcare providers	Capture the details of diagnostic studies, history and physical examinations, visit notes, nursing notes, outcome measures and any other clinically relevant information about the patient
	Healthcare providers and IS personnel	Allow exchange of medical data between different sites and different providers in an understandable and usable manner Allow effective use of information in other information management systems Allow manipulation of standardized data for generating alerts and reminders that are relevant to an individual patient Permits retrieval of relevant data, information, and knowledge to aid clinicians in making clinical decisions
	Data analysts, quality management and utilization management personnel	Provide an organized system of data collection and retrieval resulting in linkage of published research with clinical care, and ultimately improving quality of care through outcomes measurement

Table 2 (continued)		
Classification	*Consumers*	Provide data on costs and outcomes of treatment options
	Researchers and epidemiologists	Allow collection and reporting of health statistics
	Researchers and data analysts	Ensure high-quality database for accurate clinical as well as statistical data
	Accounting, coding, and billing personnel and payers	Provide data for designing payment systems and determining the correct payment for healthcare services
	Auditors and compliance personnel	Identify fraud and abuse
	Public health personnel	Provide data that are used in public health monitoring
	Management	Improve clinical, financial, and administrative performance through use of information

2.1.5 Clinical Terminology and Clinical Classification Systems: A Critique Using AHIMA's Data Quality Management DQM Model

Clinical classification systems and clinical terminologies represent two distinct sets of coding schemes that are used in healthcare. These concepts – clinical terminology and classification- are often used incorrectly and interchangeably. The purpose of this section is to try to make a distinction between clinical terminologies and clinical classification systems, identify how both sets of systems are utilized in healthcare settings, and acknowledge individual contributions of each system to providing data infrastructure for clinical as well as administrative data uses in the healthcare delivery system.

There are essentials elements that distinguish a clinical terminology from a classification system. Before jumping to a conclusion on which system is "best" to accommodate healthcare needs and data structure, a critique of both systems will be presented in the following section using American Health Information Management Association's (AHIMA) Data Quality Management Model.

The AHIMA's DQM Model will be utilized as a framework for assessment due to the following reasons: (1) AHIMA's DQM Model can provide a standard for comparison as well as an objective assessment of totally-different systems with varying scopes and applications; (2) AHIMA's DQM Model was developed to accommodate complexity of health care data by providing a way to quantify quality of this data and its attributes; and (3) There are no other relevant models that can replace the AHIMA's DQM Model in this capacity giving it is a long-established health information standard. SNOMED CT and ICD-10-CM/PCS will be utilized as examples for clinical terminologies and clinical classification systems, respectively.

2.1.5.1 AHIMA's DQM Model

Data Quality Management (Alakrawi, 2016; AHIMA's DQM practice brief, 2012) can be defined as "the business processes that ensure the integrity of an organization's data during collection, application (including aggregation), warehousing, and analysis (AHIMA, 2016; Giannangelo, 2009; Giannangelo, 2012; IHSDO, 2016). The purpose of DQM is continuous improvement of health data quality. DQM model consists of 10 characteristics to monitor data quality in 4 different domains including data application, collection, warehousing, and analysis. Table 3 provides a description of the four domains that constitute the AHIMA's DQM Model along with characteristics of data integrity that should be applied in each domain.

Accessibility

SNOMED CT contributes to semantic interoperability across a wide range of clinical applications between healthcare providers in different clinical settings and therefore can improve the capabilities of health information exchange (Duarte, 2014; Gøeg, 2014; Houser & Meadow, 2017) Semantic interoperability can be defined as "ensuring that precise meaning of exchanged information is understandable by any other system or application not initially developed for this purpose" (Gøeg, 2014). However, such high-level of information exchange is not quite feasible utilizing a classification system like ICD-10-CM/PCS that is too general to serve this purpose (Jensen, 2012) Therefore, SNOMED CT can greatly improve data accessibility as opposed to ICD-10-CM/PCS. In addition, applications that use SNOMED CT make the data accessible at the point of care, while ICD-10-CM/PCS data are accessible only after codes are assigned by the coders.

Table 3. DQM Domains and Characteristics with Definitions

DQM Domains and Definitions	
I. Application	The purpose for the data collection
II. Collection	The processes by which data elements are accumulated
III. Warehousing	Processes and systems used to archive data and data journals
IV. Analysis	The process of translating data into information utilized for an application
DQM Characteristic and Definitions	
1. Accessibility	Data items that are easily obtainable and legal to access with strong protections and controls built into the process
2. Accuracy	The extent to which the data are free of identifiable errors
3.Comprehensiveness	All required data items are included—ensures that the entire scope of the data is collected with intentional limitations documented
4. Consistency	The extent to which the healthcare data are reliable and the same across applications
5. Currency	The extent to which data are up-to-date; a datum value is up-to-date if it is current for a specific point in time, and it is outdated if it was current at a preceding time but incorrect later
6. Definition	The specific meaning of a healthcare-related data element
7. Granularity	The level of detail at which the attributes and values of healthcare data are defined
8. Precision	Data values should be strictly stated to support the purpose
9. Relevancy	The extent to which healthcare-related data are useful for the purposes for which they were collected
10. Timeliness	Concept of data quality that involves whether the data is up-to-date and available within a useful time frame; timeliness is determined by manner and context in which the data are being used

Accuracy

SNOMED CT is an automated clinical terminology scheme in which clinical representations are automatically encoded using a variety of coding applications that utilize Natural Language Processing NLP (Duarte, 2014; Stanfill, 2015). In fact, SNOMED CT is agnostic i.e. can capture all codes regardless of context. Therefore, incorrect data resulted from human errors are not probable as opposed to ICD-10-CM/PCS coding systems in which human judgement is an important element in the coding process. However, there is a higher risk of systematic errors in clinical applications as opposed to human errors which tend to be randomly distributed in most cases (AHIMA, 2014). Human judgment component of coding has also contributed to coding variations and issues with coded data accuracy. Complexity of resource grouping schemes as well as unclear documentation can lead to inaccurate coding (Drake, 2016; Land, 2016; Nouraei, 2013). Furthermore; accuracy requires familiarity with medical terminology, surgical techniques, and complex coding systems (Moar, 2012). For example, coding accuracy can vary greatly across medical specialties. Some specialties like otolaryngology encompass a wide-range of procedures which are performed in "close anatomical proximity" and that ultimately affect coding accuracy (Drake, 2016; Land, 2016; Nouraei, 2013). Similar results were found in different medical specialties; urology (Moar, 2012), neurosurgery (Beckley, 2009), and surgery (Naran, 2014).

Comprehensiveness

SNOMED CT has better clinical coverage than ICD-10-CM/PCS. The number of codes representing concepts in clinical findings alone is 100,000 concepts compared 68,000 diagnosis codes in ICD-10-CM (AHIMA, 2014; Alakrawi, 2016; IHSDO, 2016). Thus, we might need more than one code in ICD-10-CM to represent one concept in SNOMED CT. New concepts in SNOMED CT (post-coordinated expression) can be created which contributes to the system extensibility to cover all concepts related to the medical domain (IHSDO, 2016). In the other hand, ICD-10-CM/PCS is updated periodically to revise or add new diagnosis or procedure codes. Table 4 provides examples on comprehensiveness of both systems.

Consistency

Concepts in SNOMED CT are always consistent between different users and across all clinical applications (Duarte, 2014; IHSDO, 2016). In contrast, studies have shown issues with coding reliability that contributes to inconsistent code assignments between coders and across medical specilaities (Beckley, 2009; Land, 2016; Moar, 2012; Naran, 2014). In addition, ICD systems in general are influenced by coding conventions that are subject to interpretation by coders and which can vary across settings i.e. inpatients vs. outpatient clinical context (AHIMA, 2013; AHIMA, 2014; Butler, 2016). For examples, coding symptoms and signs such as "shortness of breath" can have different guidelines in acute-care hospitals and ambulatory care settings.

Currency

SNOMED CT in its current form was developed in 2007 (IHSDO, 2016) while the WHO's ICD-10 was first introduced in 1990s and has been used to collect mortality statistics in the US. However, the first field test of ICD-10-CM was conducted in 2003. Both systems are updated bi-annually to reflect contemporary medical knowledge and medical technology (CMS, 2015; IHSDO, 2016).

Definition

Due to its logical structure, SNOMED CT makes more sense and is easier to be understood by clinicians (Alakrawi, 2016; Duarte, 2014; El-Sappagh, 2014; Mikroyannidi, 2012). However, ICD-10-CM can be too impeded with coding conventions and sometimes clinically irrelevant details needed for reimbursement of healthcare services (initial encounter, delayed healing, NOS, NEC). These instructions are designed for professional coders and therefore make it hard for clinicians to adopt the system for direct care purposes (AHIMA, 2014; Stanfill, 2015). Table 5 provides examples on the different language used by both systems (data definition).

Table 4. Example on Comprehensiveness of Both Systems

SNOMED CT	ICD-10-CM
72854003 Aspiration pneumonia due to near drowning	J69.8 Pneumonitis due to inhalation of other solids and liquids Y21.8XXA Other drowning and submersion, undetermined intent (initial encounter)
283647006 Sewing needle in hand	S61.449A Puncture wound with foreign body of unspecified hand (initial encounter) W27.3XXA Contact with needle (sewing) (initial encounter)
275434003 Stroke in the puerperium	O99.43 Diseases of the circulatory system complicating the puerperium I63.9 Cerebral infarction, unspecified
15781000119107 Hypertensive heart AND chronic kidney disease with congestive heart failure	I13.0 Hypertensive heart and chronic kidney disease with heart failure and stage 1 through stage 4 chronic kidney disease, or unspecified chronic kidney disease N18.9 Chronic kidney disease, unspecified I50.9 Heart failure, unspecified
111570005 Anemia due to infection	B99.9 Unspecified infectious disease D64.89Other specified anemias

Granularity

SNOWMED CT is in general is more specific than ICD-10-CM/PCS (AHIMA, 2012). Furthermore, SNOMED CT has a unique characteristic that enables extensibility and creating of new concepts (post-coordinated expressions) by end-users (IHSDO, 2016). In contrast, less common diseases in ICD-10-CM are grouped together in “catch-all” categories e.g. J15.8 Pneumonia due to other specified bacteria which can lead to loss of information (Drake, 2016; Stanfill, 2015).

Precision

Concepts have the same values in SNOMED CT; studies have shown up to 93% precision of SNOMED CT for identifying clinical expressions (IHSDO, 2016; Lee, 2014; Skeppelstedt, 2011). However, the presence of some codes with unspecified (not specified in documentation) and other specified (present in medical record but no enough details in ICD to code it) can impact ability of the ICD system to collect data related to certain conditions such as rare conditions. Therefore, it is advised to take caution when utilizing administrative data for less common conditions such as Down Syndrome, eosinophilic esophagitis, congenital heart disease, genetic blood disorders, and surgery (Broberg, 2014; Nouraei, 2013; Rybnicek, 2014).

Table 5. Examples on the Different Language Used by Both Systems (Data Definition)

Clinical Expression	SNOMED CT	ICD-10-CM
Apert syndrome	20528009 Apert syndrome	Q87.0 Congenital malformation syndromes predominantly affecting facial appearance.
Hashimoto thyroiditis	21983002 Hashimoto thyroiditis	E06.3 Autoimmune thyroiditis
Feather picker's disease	11944003 Feather-pickers' disease	J67.8 Hypersensitivity pneumonitis due to other organic dusts
Airport malaria	240631007 Airport malaria	B54 Unspecified malaria
Adhesion of penis due to circumcision	435311000124103 Post-circumcision adhesion of penis	N99.89 Other postprocedural complications and disorders of genitourinary system
Family history of Sickle cell anemia	160321003 Family history of Sickle cell trait	Z83.2 Family history of diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism
Syphilitic parkinsonism	38523005 Syphilitic parkinsonism	A52.19 Other symptomatic neurosyphilis
Fragile X syndrome	205720009 Fragile X chromosome	Q99.2 Fragile X chromosome
Kabuki syndrome	313426007 Kabuki make-up syndrome	Q89.8 Other specified congenital malformations
Drug abuse - antidepressant	191928000 Abuse of antidepressant drug	F19.10 Other psychoactive substance abuse, uncomplicated

Relevancy

A clinical terminology such as SNOMED CT could be more useful in clinical applications, information retrieval, and research. SNOMED is regarded as a global standard due to its wide acceptance and application world-wide which makes it a safe and accurate alternative for clinical communication between healthcare providers (Alakrawi, 2016; Duarte, 2014; El-Sappagh; IHSDO, 2016; Mikroyannidi, 2012). In contrast, classification systems such as ICD-9-CM or ICD-10-CM/PCS are intended for classification of clinical conditions and procedures to be used for other applications including statistical reporting and reimbursement (Alakrawi, 2016; AHIMA, 2012; AHIMA, 2014, Duarte, 2014). Both systems are relevant to with respect to the purposes for which they were originally designed.

Timeliness

SNOMED CT is designed to be used at the point of care by clinicians while ICD-10-CM/PCS codes are usually assigned by professional coders after the patient's episode of care is complete (Alakrawi, 2016; IHSDO, 2016).

Figure 5 presents a model that was developed based on AHIMA's DQM to illustrate the fundamental differences between clinical terminologies (represented by SNOMED-CT) and clinical classification systems (represented by ICD-10-CM).

Figure 5: AHIMA's DQM Model; Comparing Data Quality of SNOMED-CT and ICD-10-CM

SNOMED CT *(Clinical Terminology)*		ICD-10-CM *(Classification System)*
Semantic interoperability enables sharing and exchange of information by different providers in different healthcare settings	***Accessibility***	Technical interoperability between coding applications and other local applications but no semantic interoperability to enable high level exchange of health information
SNOMED CT is originally designed to be used by computers. Data is automatically encoded and therefore errors in data entry caused by humans are eliminated	***Accuracy***	Coding is a semi-automated process at its best and therefore is susceptible to human errors. Coding conventions that require interpretations by coders are a major cause of coding variations.
SNOMED CT has more content coverage; 100,000 concepts in clinical findings. SNOMED can be expanded by creating new concepts (post-coordinated expressions)	***Comprehensiveness***	ICD-10-CM is limited to a set of codes that cannot be expanded. New medical conditions cannot be incorporated by end-users but rather through frequent updates of the system. Number of codes in ICD-10-CM is 68,000
Concepts has a unique numeric identifier, a unique description (FSN) and therefore, the same codes are generated for all users across different applications	***Consistency***	Coding is subjective to coding variability between coders. In addition, coding conventions can vary between inpatient and outpatient settings
SNOMED CT in its current form was developed in 2007 and it is updated biannually through IHTSO	***Currency***	WHO's ICD was used in 1990s and in 2003 the first field test of ICD-10-CM was conducted. Reviewed biannually
SNOMED-CT follows a logical structure which makes it easier for clinicians to understand. Every concept has a unique identifier and FSN which makes standard definitions of data elements that are not susceptible to interpretation	***Definition***	ICD-10-CM/PCS can be impeded with coding conventions and guidelines as well as irrelevant details that are important to coders but not clinicians. Also, some codes are not clearly defined
Greater granularity and specificity- every piece of information can be covered through pre-coordinated and post-coordinated expressions.	***Granularity***	Less specific than SNOMED CT which can lead to loss of important details; inability of ICD-10-CM systems to capture some details documented in the EHR
SNOMED-CT has shown higher precisions in information retrieval (up to 93%) due to its standardized structure.	***Precision***	ICD-9/10-CM have shown lower precision in identifying rare diseases and clinical conditions. Coding variability has significantly impacted precision of the ICD systems.
Relevant for its intended purpose. SNOMED-CT is an input system that is widely accepted which makes it suitable for standard health information sharing and information retrieval.	***Relevancy***	Statistically focused- expanded to include reimbursement. Relevant for its intended purpose: output system designed for general reporting and reimbursement when used for resource grouping.
Used at point of care by clinicians in different applications: clinical decision supports and in generating alerts and reminders.	***Timeliness***	Codes are usually entered after the episode of care is completed by coding professionals.

2.1.5.2 Users and Applications

Healthcare terminology and classification systems can be used by different users including consumers, healthcare providers, quality and utilization management personnel, researchers and other administrative staff (accounting, billing, and coding personnel). They are also used to facilitate communication between healthcare providers and consumers at the point of care for data collection purposes. A more organized system of data collection and retrieval can be provided by utilizing healthcare terminology. This system can promote quality of care through providing a link between published research and clinical care. Furthermore, such systems can support integration of care through allowing effective exchange of clinical information between healthcare providers in different settings. Although terminologies such as SNOMED CT can be utilized to support real time decision making and retrospective reporting for research and management, such utilization can be hindered by complexity of these systems. Classification systems are utilized by wider spectrum of users in healthcare. They can be used to provide data to consumers on costs, treatment options, and outcomes. Also, classification systems provide a less complex system for data collection and reporting that can be further used for research purposes. Information provided by such systems can be used to improve clinical, financial, and administrative performance through designing effective payment systems, identifying potential fraud and abuse, and ensuring accurate reporting.

ICD-10-CM/PCS

The ICD coding system was originally created to code death certificates but its use was expanded to encompass a wide variety of statistical reporting. In fact, ICD-10 has been used since the 1990s to collect mortality statistics around the world. The WHO defines coding as "the translation of diagnoses, procedures, co-morbidities and complications that occur over the course of a patient's encounter from medical terminology to an internationally coded syntax" (WHO, 2016). In this definition, the WHO acknowledges the capability of the ICD system that is used for clinical coding and classification to enable international comparisons with respect to mortality as well as morbidity statistics.

ICD-9-CM had been used from 1978 - 2015 as the foundation of the reimbursement system in the United States and used by the Center for Medicare and Medicaid Services for inpatient and ambulatory resource grouping. Medicare Severity Diagnosis Related Group (MS-DRG) constitutes the foundation of Medicare's Inpatient Prospective Payment System (IPPS) used to reimburse acute care and short term hospital for services rendered to Medicare beneficiaries. ICD-9-CM was replaced by ICD-10-CM/PCS in October 1, 2015 and it will continue to serve as a base for healthcare reimbursement. For outpatient encounters, reporting diagnosis codes in ICD-10-CM is required to establish medical necessity (Alakrawi, 2016; CMS, 2016).

Also, ICD-10-CM is now used in place of ICD-9-CM for public health reporting i.e. reporting the leading cause of death and morbidity on the national level. ICD-10-CM/PCS can also be used to assess clinical outcomes and improve quality of care provided for individual patients. For example, ICD-10-CM/PCS data is utilized for clinical documentation improvement (CDI) initiatives to educate physicians on clinical documentation in the EHR systems.

However, the process of clinical classification itself is prone to variation due to the complex coding schemes and conventions that are subject to interpretation by coders which makes it difficult for clinicians to assign the codes by themselves. Thus, ICD-10 in general and ICD-10-CM/PCS lacks standardization needed for electronic communication and clinical documentation.

SNOMED CT

SNOMED CT can provide a unified language that can be used as a standard for communication between healthcare providers and across clinical applications. SNOMED CT can contribute greatly to semantic interoperability in healthcare applications (Alakrawi, 2016; Duarte, 2014; IHSDO, 2016). Its standardized logical structure as well as its wide acceptance makes it more suitable than other terminologies or classification systems for high-level information sharing and information retrieval (IHSDO, 2016). Thus, SNOMED CT can be used for health information exchange HIE and clinical documentation in the EHRs. SNOMED CT is an automated system which makes it convenient to be used at the point of care for generating clinical alerts and reminders, serves as a part of the Clinical Decision Support (CDS) System, and link providers to medical knowledge and current publications that can be used for outcome measurement. Furthermore, due to its fully-automated scheme, SNOMED CT can be used for health care research, and in automated identification of patients for clinical trials due to its extensive granularity and content coverage (Alakrawi, 2016; Della Mea, 2014; IHSDO, 2016). In addition to its higher specificity, SNOMED CT has a unique feature that enables extensibility of concepts by end users which can foster reliable communication between healthcare providers, across medical specialties, and health information exchange at national as well as international levels. SNOMED CT has become one of the federal requirements for health information technology HIT; the Centers for Medicare and Medicaid Services (CMS) mandates the use of SNOMED CT to code the problem list for Meaningful Use (MU) stage 2 (Alakrawi, 2016; Della Mea, 2014; IHSDO, 2016).

2.1.5.3 Clinical Documentation into the EHR

However, SNOMED CT is not superior to ICD-10-CM/PCS as both coding schemes provide the necessary data structure needed to support healthcare clinical and administrative processes.

Clinical terminology systems as well as clinical classification systems were originally designed to serve different purposes and consequently different users requirements. ICD-10-CM/PCS is a more of an output system that is designed for general reporting purposes, public health surveillance, administrative performance monitoring, and reimbursement of healthcare services. In contrast, SNOMED CT was developed to serve as a standard data infrastructure for clinical applications which requires more specificity. A classification system can be less-detailed than a clinical terminology (Alakrawi, 2016; Chavis, 2013). Therefore, "less specificity" of ICD-10-CM/PCS is an intrinsic feature rather than a "malfunction"; SNOMED CT is too detailed therefore to replace ICD-10 in this context (Alakrawi, 2016; AHIMA, 2014). In fact, both systems complement each other and contribute to providing quality data for different domains of the healthcare system. For example, "If a researcher wants to know how many patients died with a diagnosis of heart attack last year, ICD-10 (WHO's) is enough. If they want more detail, such as what muscle of the heart was involved, they will need SNOMED CT" (Chavis, 2013). Therefore, both can be used in research and education depends on which degree of specificity is required by circumstances: SNOMED is a better choice for identifying rare diseases while ICD-10-CM/PCS is more efficient for general reporting such as collecting the top causes of mortality and morbidity at the national level. Furthermore, ICD-10-CM/PCS will be needed to constitute the foundation of the reimbursement system.

3.0 CLINICAL CODING PROCESSES

3.1 CLINICAL CODING

Clinical coding can be defined as "the translation of diagnoses, procedures, co-morbidities and complications that occur over the course of a patient's encounter from medical terminology to an internationally coded syntax" (WHO, 1994). Clinical coding was initially intended for causes of mortality reporting (Land, 2016; Nouraei et al, 2013). However, coded clinical data has a significant impact on the health care industry for assessing clinical outcomes, monitoring quality of care, conducting research, promoting education, resource allocation, planning health services, and benchmarking (Alakrawi, 2016; AHIMA, 2016; CDC, 2009; Giannangelo, 2012; Nouraei et al, 2013).

Particularly, coding impacts public health reporting since it is used to determine the leading causes of mortality and morbidity in the U.S. Also, it is the major factor in the promotion of funding for different diseases and healthcare services in general (Alakrawi, 2016; CDC, 2014; CMS, 2016). Therefore, accurate coding for public health reporting solely depends on coding or data collection at the baseline (individual patient's encounter). However, coded data are generally under-utilized in healthcare because of a lack of familiarity and issues related to data accuracy and availability (Land, 2016; Nouraei et al, 2013). Further discussion of these issues is provided in chapter 5.

Coding is known to be the foundation of the reimbursement system in the United States, which creates an increasing demand to improving medical coding to meet compliance requirements. It is part of the fundamental functions in the field of HIM. However, in this era of

EHRs and based on the need for electronic transaction, coders need not only to be familiar with the code assignment process but also with mapping among different clinical nomenclatures and terminologies (Alakrawi, 2016; AHIMA, 2013; AHIMA & AMIA, 2007).

According to McBride (2006), “Data mapping involves "matching" between a source and a target, such as between two databases that contain the same data elements but call them by different names. This matching enables software and systems to meaningfully exchange patient information, reimbursement claims, outcomes reporting, and other data.” Data mapping can be classified into unidirectional and directional mapping where “unidirectional mapping goes from the source to the target. Bidirectional maps translate in both directions” (McBride, 2006; NLM, 2016).

The National Library of Medicine (NLM) with participation from the National Center for Health Statistics (NCHS) is working on a project to Map SNOMED CT concepts to ICD-10-CM codes through I-MAGIC (Interactive Map-Assisted Generation of ICD Codes) (CDC, 2015; CMS, 2016; NLM, 2016). Per NLM (2015), the purpose of mapping is to “is to support semi-automated generation of ICD-10-CM codes from clinical data encoded in SNOMED CT” to fulfill the requirements of healthcare. Therefore, SNOMED CT cannot replace ICD-10-CM/PCS and both systems complement each other and equally contribute to quality data structure for the entire healthcare system.

In fact, the WHO joint with the International Health Terminology Standards Development Organisation (IHTSDO) has been working on similar projects that will enable mapping between SNOMED CT and ICD-10 (the WHO’s version) and ICD-11 as well (Alakrawi, 2016; Chavis, 2013; NLM, 2016). However, due to substantial differences between both coding schemes, it is not always possible to have one-to-one map. However, these mapping projects further emphasize the importance of future data infrastructure that encompasses both system characteristics to utilize the maximum benefits of information technology in healthcare.

Thus, clinical coders, at least, should have the knowledge and skills that are needed to deal with the HIPAA code sets. HIPAA standard code sets include the following: International Classification of Diseases, tenth Revision, Clinical Modification (ICD-10-CM); Current Procedural Terminology (CPT); Code on Dental Procedures and Nomenclature (CDT); National Drug Codes (NDCs); and Healthcare Common Procedure Coding System (HCPCS) (AMA, 2014; CMS, 2017).

3.2 CODING CLINICAL EXPRESSIONS USING SNOMED CT AND ICD-10-CM/PCS

The two sets of systems were designed to serve different purposes and therefore are intended to satisfy different user requirements. SNOMED CT is designed for input into Electronic Health Record (EHR) systems and other clinical applications while ICD-10-CM/PCS is basically designed for providing outputs in terms of reports and statistics. Therefore, each system has a unique hierarchical structure to serve the purposes for which it was originally intended (AHIMA, 2013; Glenn, 2013; IHSDO, 2016).

Figure 6 represents a brief description of how to code the clinical expression "pain in right leg" using a clinical terminology (SNOMED CT) and a classification system (ICD-10-CM). Also, more examples can be found in Table 6.

Coding in SNOMED is totally different than conventional coding using ICD-10-CM/PCS. In fact, the process of "coding" using SNOMED CT differs from ICD-10-CM/PCS. Coding using

SNOMED CT is always automated: end users cannot view the codes assigned by the system (AHIMA, 2013; Glenn, 2013; IHSDO, 2016). For this reason, SNOMED-CT is being used by software developers and EHR vendors to facilitate communication between different applications through creating a standard language. In fact, we can think of SNOMED-CT as a programing language; users utilize applications that use SNOMED-CT without knowing what is it that in the background (IHSDO, 2016).

SNOMED CT	ICD-10-CM
Composed of a wide set of concepts and relationships that connect these concepts together to fully cover the presented clinical expression. Each concept is represented by a unique numeric identifier and a Fully Specified Name (FSN), which is a unique description of that specific concept. SNOMED CT is designed for clinical applications and therefore clinical expressions are automatically coded in the background without user intervention. In order to code the clinical expression "pain in the right leg", a user needs to input the clinical phrase and SNOMED CT will generate the following code: 287048003 "Pain in the right leg" = "pain" + "right" + "leg".	A classification system organized into chapters as well as categories and sub-categories in each chapter. ICD-10-CM coding has not been fully automated yet so the process of coding requires a degree of human intervention. To code the same clinical condition "pain in the right leg", a coder is required first to search the alphabetic index and follow a specific set of coding conventions and instructions to assign the correct code from the tabular list. The corresponding code for "pain in the right leg" is M79.604. However, with increasing use of technology, Computer Assisted Coding (CAC) applications can be used to connect suggested codes to text entries in EHR system.

Figure 6: Coding Natural Language Clinical Phrases Using SNOMED CT and ICD 10-CM

Table 6. Examples of Natural Language Expressions Coded in SNOMED-CT and ICD-10-CM

Natural language clinical phrase	SNOMED-CT	ICD-10-CM
Pain in right leg	287048003 Pain in right leg	M79.604 Pain in right leg
Metabolic acidosis	59455009 Metabolic acidosis	E87.2 Acidosis
Respiratory acidosis	12326000 Respiratory acidosis	E87.2 Acidosis
Diverticulitis of sigmoid colon	427910000 Diverticulitis of sigmoid colon	K57.32 Diverticulitis of large intestine without perforation or abscess without bleeding
G6PD anemia	62403005 Glucose-6-phosphate dehydrogenase deficiency anemia	D55.0 Anemia due to glucose-6-phosphate dehydrogenase [G6PD] deficiency
Polyp in cervix	65576009 Polyp of cervix	N84.1 Polyp of cervix uteri
Otitis media in the right ear	194289001 Acute right otitis media	H66.91Otitis media, unspecified, right ear
E. coli pneumonia	51530003 Pneumonia due to Escherichia coli	J15.5 Pneumonia due to Escherichia coli
Ovale malaria	19341001 Ovale malaria	B53.0 Plasmodium ovale malaria
Vitamin A deficiency	72000004 Vitamin A deficiency	E50.9 Vitamin A deficiency, unspecified

For example, SNOMED-CT has been combined with NLP to improve EHR capabilities. In this case, SNOMED could identify where a condition exists or not or when it is ruled out due to the unlimited set of concepts and attributes that could further clarify a certain case. If such capabilities are enabled, SNOMED-CT can be used for generating alerts and reminders as well as a part of the decision support system to spot such contradicting notes and improve the quality of patient care (Alakrawi, 2016; IHSDO, 2016).

In contrast, ICD-10-CM/PCS coding is performed by professional coders who manually assign codes to patients' diagnoses and procedures (AHIMA, 2014). With the advancement of technology, coders have been using special encoders and Computer Assisted Coding (CAC) applications. CAC applications can facilitate accurate and efficient coding by automatically suggesting codes based on the clinical documentation in the EHR system (AHIMA, 2013; Houser & Meadow, 2017; Godbey-Miller, 2016). Thus, ICD-10-CM/PCS coding is semi-automated at best and always requires a degree of human intervention to either assign or validate selected codes.

3.3 INPATIENT V. OUTPATIENT CODING

The International Classification of Diseases, Tenth Revision, Clinical Modification, and Procedure Classification System (ICD-10-CM/PCS) is used for inpatient coding while Current Procedural Terminology (CPT) codes are used for outpatient coding (AHIMA; 2017; CMS, 2017). Further, Healthcare Common Procedure Coding System (HCPCS) is used for Medicare and Medicaid in addition to CPT. ICD classification systems are published by the WHO (NCHS, 2014; WHO,2012; AHIMA;2017).

However, the National Center for Health Statistics (NCHS) is the public agency responsible for maintaining and coordinating activities related to ICD classification in North America. On the other hand, The American Medical Association (AMA) is responsible for publishing CPT codes. Although inpatient and outpatient coding utilizes different classification systems, the main difference between them is the procedure code (Arner, 2007; Chavis, 2013; CMS, 2016; Linder, 2016).

Also, different resource grouping schemes are used in inpatient and outpatient settings. Resource grouping is simply grouping of conditions that are estimated to consume similar level of resources. This grouping is used for reimbursement particularly in the Prospective Payment Systems (PPS) where reimbursement is established before the services are rendered to patients (CMS, 2016; Giannangelo, 2012).

Medicare Severity Diagnosis Related Group MS-DRG constitutes the foundation of Medicare's Inpatient Prospective Payment System (IPPS) used to reimburse acute care and short term hospital for services rendered to Medicare beneficiaries. In contrast, Ambulatory Payment Classification (APC) is utilized as the unit of payment under the Outpatient Prospective Payment System (OPPS) used for reimbursement of hospital outpatient services rendered to Medicare beneficiaries. Furthermore, Resource Based Relative Value Scale (RBRVS) is Medicare's payment methods for physician's services rendered to its patients. Table 7 summarizes the difference between inpatient and outpatient coding (CMS, 2016).

Table 7. Difference between inpatient and outpatient coding

	Inpatient	**Outpatient**
Diagnosis Code	ICD-10-CM	ICD-10-CM
Procedure Code	ICD-10-PCS	CPT
Resource Grouping	MS-DRG	APCs/RBRVS

3.4 INPATIENT CODING CLASSIFICATION SYSTEMS; ICD-9-CM, ICD-10, ICD-10-CM/PCS, AND ICD-11

According to the CDC**, "**The ICD has been revised periodically to incorporate changes in the medical field. To date, there have been 10 revisions of the ICD. The years for which causes of death in the United States have been classified by each revision are illustrated in figure 7. (CDC, 2016):

1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
1900-09	*1910-20*	*1921-29*	*1930-38*	*1939-48*	*1949-57*	*1958-67*	*1968-78*	*1979-98*	*1999-present*

Figure 7: Revisions of ICD

The International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) was built on the WHO's Ninth Revision, International Classification of Diseases (ICD-9). ICD-9-CM used to be the official system of assigning codes to diagnoses as well as procedures related to utilization of acute care in the United States (1979-2015). The ICD-9 (the WHO edition) was used to code and classify mortality data from death certificates until 1999. After that, ICD-9 was replaced by ICD-10 to serve the same purpose (CDC, 2016).

However, ICD-9-CM had been used to collect morbidity statistics since 1979 until 2015 when ICD-10-CM/PCS was implemented in the U.S. The ICD-9-CM consists of: (1) a tabular list which is a numerical list of the disease code numbers in tabular form; (2) an alphabetical index to the diseases; and (3) a classification system for surgical, diagnostic, and therapeutic procedures (alphabetic index and tabular list of medical procedures) (CDC, 2016, CMS, 2016; WHO, 2016).

Now, ICD-10-CM constitutes the basis for the IPPS developed by CMS to pay for services rendered to Medicare and Medicaid beneficiaries (Linder, 2016). Further, NCHS and the CMS are the U.S. governmental agencies are responsible for overseeing all changes and modifications to the ICD-10-CM (CMS, 2016). The United States has transitioned to ICD-10-CM/PCS in October 1, 2015, to replace ICD-9-CM Volume 2 and 3, diagnoses and procedures code sets respectively.

ICD-10-CM/PCS is undergone periodic revision. This revision is necessary to enable a scientific update of the coding scheme as well as interoperability of ICD-10 with electronic health applications (AHIMA; 2014; Houser & Meadow, 2017; Rode, 2013; WHO, 2012). A major interoperability issue here is how to make ICD-10 (WHO's version) compatible with the Systematized Nomenclature of Medicine (SNOMED CT) and other terminologies and ontologies used for building clinical applications (IHSDO, 2016; Mahajan, 2013; WHO, 2012).

ICD-10-CM is the United States' clinical modification of the WHO' ICD-10. The NCHS has developed ICD-10-CM for morbidity purposes. On the other hand, ICD-10-PCS was developed by 3M Health Information Systems based on a 3-year contract with Healthcare Financing Administration (HCFA), now CMS, in 1995.

Major changes of ICD-10-CM include the following: (1) E codes are no longer separated but incorporated in the main classification; (2) injuries are grouped by body parts instead of categories; (3) expanded excludes notes; (4) combination codes have been created; (5) laterality has been added (as a concept); and (6) greater specificity in code assignment (AHIMA 2014; AHIMA 2013, Boyed et al, 2013; Land, 2016; Walker, 2012). Comparison between ICD-9-CM and ICD-10-CM/PCS is provided in Table 8. Also, Figure 8 provides a coding scenario in ICD-10-CM/PCS that could emphasize some of these differences.

Table 8. Comparison between ICD-9-CM and ICD-10-CM/PCS

	9th Revision	10th Revision	
	ICD-9-CM	**ICD-10-CM**	**ICD-10-PCS**
Maintenance	National Center for Health Statistics (NCHS)	National Center for Health Statistics (NCHS)	Centers for Medicare and Medicaid Services (CMS)
Structure	Hierarchal structure: • All codes within the same category have common traits (first three digits) Greater specificity can be added with each additional character beyond the 3-digit category	Has the same hierarchal structure of ICD-9-CM: All codes within the same category have common traits (first three digits) Greater specificity can be added with each additional character beyond the 3-digit category	Multi-axial structure
Number of Codes	Diagnoses: 13,500 Procedures: 4,000 Max for diagnosis codes: 5-digit • Max for procedure codes: 4-digit	Diagnoses: 70,000 Max for diagnosis codes: 7-digit	Procedures: 72,000 Procedure codes:7-digit

Table 8 (continued)			
Number of Chapters	17 chapters (diagnoses): conditions are classified per etiology (cause of disease) or by anatomical site (body system)	21 chapters (diagnoses): this includes chapters' rearrangement, additions and revisions in addition to extensive changes to the mental and behavioral disorders, injury and poisoning, and external causes of morbidity and mortality.	
Type of Codes	Mostly numeric with some alphanumeric codes (E,V, and Morphology codes)	Alphanumeric coding scheme to provide more categories for diseases and health related conditions	Alphanumeric
	E and V codes are considered as supplementary classifications	Incorporated into the classification and not separated into supplementary classifications	
V and E codes	Lacks laterality	The concept of laterality (right-left) has been added	Laterality added as opposed to procedure coding in ICD-9-CM (volume 3)

Table 8 (continued)			
Laterality	Lacks detail	Greater specificity in code assignment (for example, diabetes, family history)	The current structure of ICD-10-PCS support greater specificity as opposed to ICD-9-CM's volume 3
Specificity	Sequencing multiple codes is necessary	Combination codes have been created to resolve issues related to code-sequencing	
Multiple conditions	Grouped by categories of injuries: • Fractures (800-829) Sprains and strains (840-848)	Grouped by body parts: • Injuries to the head (S00-S09) Injuries to the neck (S10-S19)	
Grouping of injuries	Volume 3 of ICD-9-CM was used to code medical procedures (it does not reflect the rapid changing in surgical technology)	ICD-10-CM only contains diagnosis codes. Procedures are coded using ICD-10-PCS.	
Procedures Coding		Sophisticated multi-axial system used to code procedures. It has a seven-character alphanumeric code structure. Each character essentially has many possible values in this coding scheme.	

The following coding Scenario is provided in figure 8 with relevant ICD-10-CM/PCS codes: "***A 16-year-old was at the gymnasium participating in a dancing contest to raise money for her class. She slipped on the gym floor and fell, injuring her left ankle. After she visited the emergency room, it was determined that she did not fracture her ankle. However, she did suffer a sprain to that ankle***"

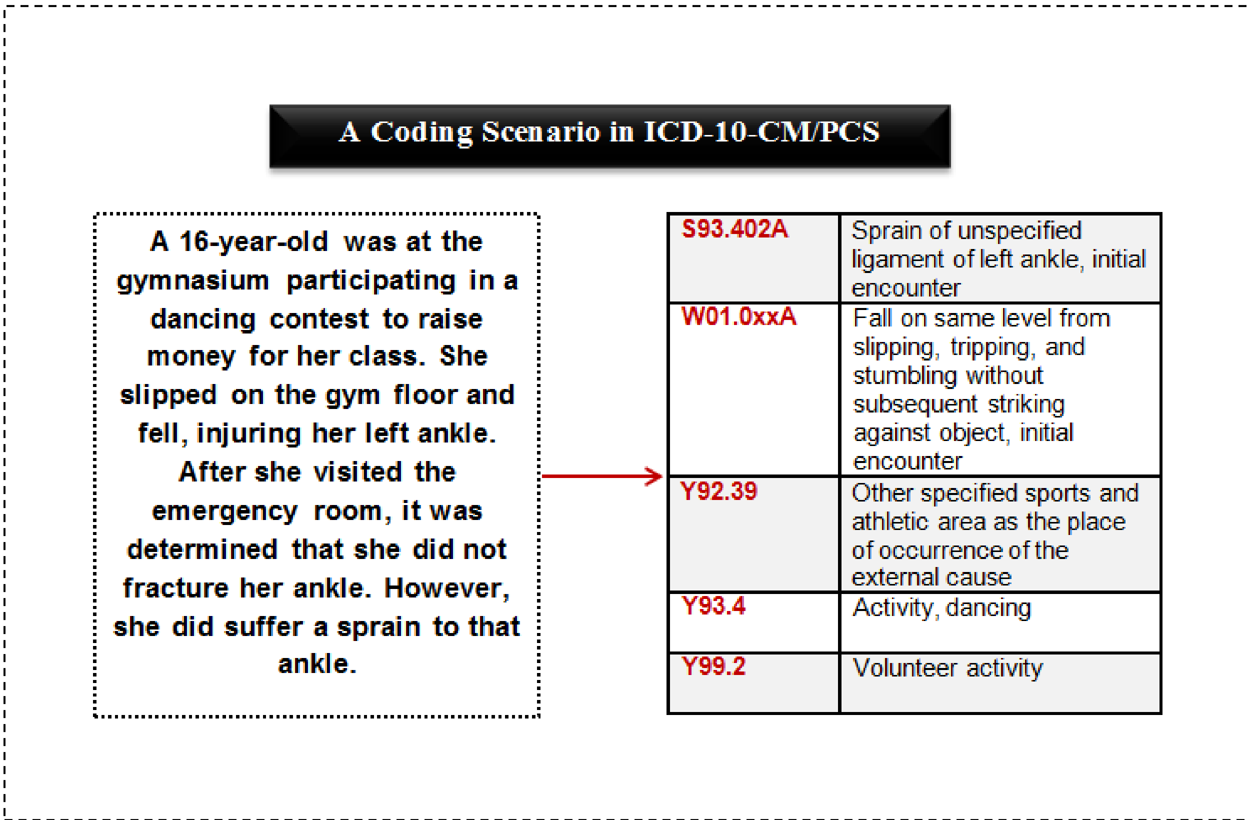


Figure 8: A coding Scenario Using ICD-10-CM/PCS

ICD-10-CM/PCS as illustrated above is superior to ICD-9-CM with respect to specificity, laterality, and detail surrounding the causes of injury such as the type of activity and place of occurrence. However, the WHO is currently working on the 11th revision of the ICD. The Beta draft of ICD-11 was made available online in May 2012 for interested stakeholders and individuals to make comments or a proposal to change, participate in field trials, and assist in translating (Reed, 2010; Stanfill; 2016). Figure 9 illustrates ICD-11 timeline.

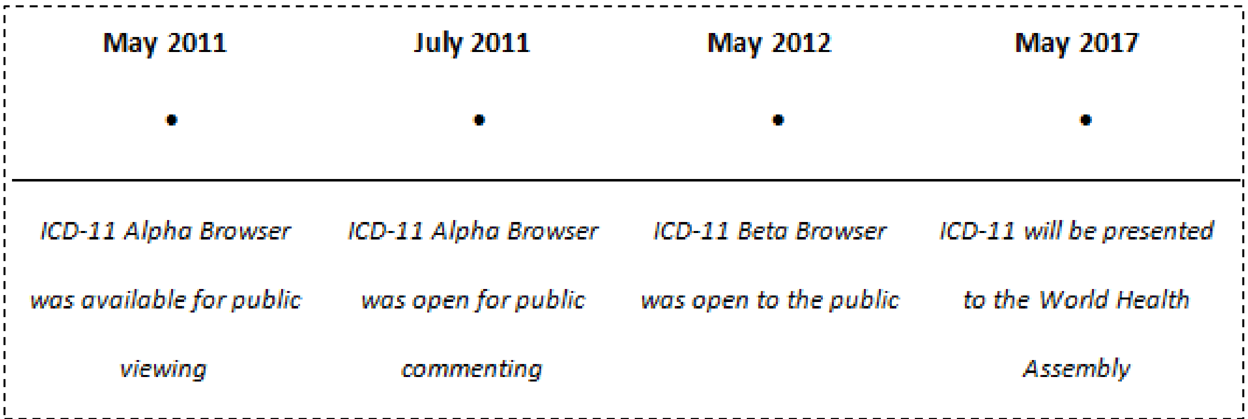


Figure 9: ICD-11 Timeline

4.0 REIMBURSEMENT METHODS IN HEALTHCARE

Different reimbursement methods exist in the U.S. healthcare delivery system. In general, reimbursement is determined based on the following factors: (1) health care setting, (2) health care provider, and (3) third party payer. However, encoders and groupers could have an impact on reimbursement as well (AHIMA, 2016; Cade, 2012).

Reimbursement methodologies can be simply classified into two distinct categories: Prospective and Retrospective payment systems (CMS, 2016). In the Retrospective Payment Systems (RPS), reimbursement is established <u>after</u> the healthcare services are rendered while in the Prospective Payment Systems (PPS), reimbursement is established <u>before</u> healthcare services are rendered (Alakrawi, 2016; Cade, 2012; DeAlmeida, 2012). The CMS utilizes different reimbursement methodologies for different types of healthcare facilities (CMS, 2016). Below is a discussion of two PPSs used by the CMS: Inpatient Prospective Payment System (IPPS) and Outpatient Prospective Payment System (OPPS).

4.1 INPATIENT PROSPECTIVE PAYMENT SYSTEM (IPPS)

According to CMS, a Prospective Payment System (PPS) is “a method of reimbursement in which Medicare payment is made based on a predetermined, fixed amount. The payment amount for a particular service is derived based on the classification system of that service; for example, diagnosis-related groups [DRGs] for inpatient hospital services” (CMS, 2016). Further, CMS uses separate PPSs for reimbursement to acute inpatient hospitals, home health agencies, hospice, hospital outpatient, inpatient psychiatric facilities, inpatient rehabilitation facilities, long-term care hospitals, and skilled nursing facilities (CMS, 2016).

The DRGs were originally developed at Yale University in 1975. The purpose of this project was to enable grouping of patients with similar conditions and treatments for comparative studies. In 1983, the DRGs were adopted by Medicare as the basis for the IPPS and have been modified by many agencies and companies since then. Different DRG systems are now used by different payers. However, the two main DRG systems in use today are the Medicare-Severity Diagnosis Related Group (MS-DRG) and All Patient Refined DRGs (APR-DRGs) developed by 3M.

In general, DRGs are designed based on codes. However, there are other factors that should be considered: (1) diagnosis codes (ICD-10-CM); (2) procedure codes (ICD-10-PCS); (3) patient age; (4) patient sex; and (5) discharge disposition. Sequencing of codes on the claims has a significant impact on determining proper DRGs for each patient. DRGs are assigned using software applications that are called DRG groupers. However, DRGs were grouped manually using decision trees when they were first developed in the 1980s.

To assign an MS-DRG, a case should be classified into one of 25 Major Diagnostic Categories (MDC). These MDCs are usually classified based on body systems with some exceptions. Then, it should be determined whether this specific case is medical or surgical because surgical cases usually require more resources.

In many cases, paints have other conditions that could influence their care. These conditions can be classified into Complications and Comorbidities (CC) or Major Complications and Comorbidities (MCC) simply based on their severity. Each individual DRG has a pre-determined relative weight that reflects the amount of resources used in treating patients with that

DRG. DRGs with a relative weight of 1 suggest average resource consumption. DRGs with relative weights less than 1 suggest less than average resource consumption while DRGs with relative weights greater than 1 suggest more than average resource consumption in treating patients with these DRGs. Table 9 provides some examples of medical and surgical MS-DRGs.

Table 9. Examples of MS-DRGs with different weights (FY 2016)

MS-DRG		MDC	Type	MS-DRG Title	Weight
1	020	01	Surgical	INTRACRANIAL VASCULAR PROCEDURES W PDX HEMORRHAGE W MCC	9.4201
2	032	01	Surgical	VENTRICULAR SHUNT PROCEDURES W CC	1.9875
3	042	01	Surgical	PERIPH/CRANIAL NERVE & OTHER NERV SYST PROC W/O CC/MCC	1.8655
4	123	02	Medical	NEUROLOGICAL EYE DISORDERS	0.6697
5	152	03	Medical	OTITIS MEDIA & URI W MCC	1.0141
6	158	03	Medical	DENTAL & ORAL DISEASES W CC	0.8482

Source of information: (CMS, FY 2015 Proposed Rule Tables, 2014)

As shown in Table 9, MS-DRG 032 (Ventricular Shunt Procedures W CC) has a relative weight of 1.9875. This suggests that more than average resources are used in treating patients with this condition. However, MS-DRG 032 is a surgical case with other complications or comorbidities which can justify the more than average resource consumption. In contrast, MS-DRG 158 (Dental and Oral Diseases W CC) is a medical case that requires less than average resource consumption since the relative weight for that specific MS-DRG is 0.8482. To determine the hospital's payment for each case, the DRG relative weight is multiplied by the hospital base rate. Table 10 provides the total payment amount for Hospital A where hospital rate is assumed to be $3,000.

Table 10. Total payment for each case based on hospital rate

MS-DRG		MDC	Type	MS-DRG Title	Weight	Hospital Payment
1	020	01	Surgical	INTRACRANIAL VASCULAR PROCEDURES W PDX HEMORRHAGE W MCC	9.4201	$28260.3
2	032	01	Surgical	VENTRICULAR SHUNT PROCEDURES W CC	1.9875	$5962.5
3	042	01	Surgical	PERIPH/CRANIAL NERVE & OTHER NERV SYST PROC W/O CC/MCC	1.8655	$5596.5
4	123	02	Medical	NEUROLOGICAL EYE DISORDERS	0.6697	$2009.1
5	152	03	Medical	OTITIS MEDIA & URI W MCC	1.0141	$3042.3
6	158	03	Medical	DENTAL & ORAL DISEASES W CC	0.8482	$2544.6

The average relative weight for all DRGs in a certain hospital is what constitutes the case mix index for that hospital. Thus, case mix is based on DRGs which are originally assigned based on ICD codes. It is a financial indicator of reimbursement; any change in the case mix index of a certain hospital could be attributed to a change in either patient population or coding. Therefore, case mix is frequently monitored to assess the financial health and quality of coding.

In addition, CMI has become an indicator of hospital disease severity in the United States (Mendez et al, 2013). Yang and Reinke (2006) conducted a study to evaluate different CMIs in capturing disease severity. They concluded that DRG-based CMIs are the most valid CMIs in capturing disease severity. However, CMI can be affected by documentation, coding practices, hospital, and patients' characteristics (AHIMA, 2008; Friesner et al, 2007; Hvenegaard et al, 2009; Martin, 2016; Mendez et al, 2013; Rosenbaum et al, 2014; Steinbusch et al, 2006)

In 2007, Friesner et al. conducted a study that evaluates the use of CMI as an indicator of resource utilization and patient illness severity using a panel of Washington state hospitals. Friesner and colleagues concluded that using a single CMI might not be appropriate for comparing small or mid-size hospitals but is appropriate when comparing large hospitals that treat a wide variation of conditions (Butler, 2016).

Other hospital variables could have an impact on CMI as a marker of disease severity (Hvenegaard, 2009; Martin, 2016; Mendez et al, 2013). In 2009, Hvenegaard and colleagues conducted a study to develop a model to predict hospital cost based on CMI and other routinely collected characteristics. A major study finding is that CMI is a robust factor in predicting financial performance and adding other factors such as age, gender, and socioeconomic characteristics does not seem to affect the cost significantly.

Furthermore, a study conducted by Mendez and colleagues (2013) to evaluate the impact of hospital variables on average CMI suggested that "between 1996 and 2007, average CMI declined by 0.4% for public hospitals, while rising significantly for private for-profit (14%) and non-profit (6%) hospitals." However, after introducing the MS-DRG system in 2007, the CMI increased for all types of hospitals but remained lowest in public hospitals. Also, trauma centers have higher CMI compared to non-trauma centers.

Documentation practices and coding accuracy can have an impact on CMI (AHIMA, 2008; Combs, 2016; Land, 2016; Mendez et al, 2013, Rosenbaum, 2014). For instance, lower CMI can be attributed to diminished financial support to clinical documentation improvement (CDI) in public hospitals (Mendez et al, 2013). In 2014, Rosenbaum and colleagues conducted a study to evaluate the effect of CDI and education on CMI. For this study, they created a new metric to measure the subsequent documentation improvement (normalized CMI) and compare it with the traditional CMI after conducting the educational intervention. This study reported an increase in CMI and suggested that documentation accuracy and quality are significant factors that impact the hospital CMI. Another important factor that could impact CMI is coding.

Coding practices such as quality and productivity might have a significant influence on MS-DRG assignment as well as CMI (AHIMA, 2008; Combs, 2016; Martin, 2016; Rosenbaum, 2014; Steinbusch et al, 2007). Comparing different CMI systems, Steinbursch and colleagues (2007) suggested that "there are fewer opportunities for up-coding to occur in case-mix systems that do not allow for-profit ownership and in which the coder's salary does not depend on the outcome of the classification process." Therefore, "the US case-mix system tends to be more open to up-coding than the Australian system". With respect to time, this is a higher probability of up-coding when registration is initiated at the beginning of the care process (Land, 2016; Steinbusch et al, 2007).

4.2 OUTPATIENT PROSPECTIVE PAYMENT SYSTEM (OPPS)

APCs or Ambulatory Payment Classifications is the United States government's method of paying for facility outpatient services for the Medicare beneficiaries (CMS, 2016). A part of the Federal Balanced Budget Act of 1997 made the CMS create a new Medicare "Outpatient Prospective Payment System" (OPPS) for hospital outpatient services (CMS, 2012; Cade, 2012). APCs are an outpatient prospective payment system applicable only to hospitals. The total number of APCs is 850. A case is first coded using HCPCS/CPT codes then grouped to a relevant APC category; a single patient can have many different APCs. The provider receives payment for services for each APC. However, it is important to note that "not every CPT code will have a corresponding APC, and some APCs will have multiple CPT codes associated with them" (CMS, 2016; Martin, 2016; Shaeffer & Wash, 2000; Stanfill, 2016; Wirtzer, 2012).

Physicians are reimbursed through other methodologies such as the Resource Based Relative Value System RBRVS (CMS, 2016; Cade; 2012; Linder, 2016). RBRVS is Medicare's payment method for services provided by physicians to Medicare beneficiaries. The coding system used for this payment method is HCPCS/CPT and each CPT and HCPCS code has a payment amount. Particularly, each code has a Relative Value Unit (RVU) that accounts for the physician's work, practice expense, and malpractice insurance. All RVUs are adjusted by Geographical Practice Cost Index (GPCI). The sum of the adjusted RVUs is then multiplied by the corresponding Medicare fee schedule amount to determine to total payment.

CPT's Evaluation and Management (E&M) services account for the majority of services rendered For Medicare patients (CMS, 2016; Wirtzer, 2012). Many studies have shown that E&M coding "exhibits poor reliability" even if performed by professional coders (Martin, 2016; Morsch et al, 2007; Stoner et al, 2007; Flanagan & Santos, 2009). In 2009, Flanagan and Santos conducted a study to identify sources of outpatient coding variations. This study has identified two sources of outpatient coding variations or inconsistencies. First, CMS guidelines for E&M coding "allow a large range of interpretation, requiring several ad hoc decisions (CMS, 2016; Combs, 2016; Flanagan & Santos, 2009; Stanfill, 2016). These ad hoc decisions -required to be taken by providers, carriers, institutions, auditors, and individual coders- represent one source of coding variation in outpatient settings (Flanagan & Santos, 2009). In addition to "inference", documentation represents another source of E&M coding variation. E&M coding inconsistency were attributed to issues related to history, physical exam and complexity of the coding scenarios.

4.3 FEDERAL LAWS AND REGULATIONS

Compliance with federal laws and regulation is an important element that drives quality and efficiency of operations in health care including clinical coding. Non-compliance with such regulations can present a serious disadvantage to any healthcare facility in term of liability and financial loss (disciplinary actions and fines by the federal government) (William & Cabin, 2014). Below are some of the federal regulations that have influenced clinical coding and auditing processes.

(1) False Claim Act is a federal law that imposes liability on federal contractors who defraud governmental programs. Claims under the law have typically involved health care, military, or other government spending programs, and dominate the list of the largest pharmaceutical settlements (Hill et al, 2014).

(2) Health Insurance Portability and Accountability Act of 2002 (HIPAA): includes provisions to increase as well as stabilize funding for activities directed at reducing health care fraud and improper payments in federal health programs (CMS, 2014):

   a. Creating Health Care Fraud and Abuse Control (HCFAC) program
   b. Establishing Medicare Integrity Program (MIP); reduce improper payment in Medicare.

(3) Improper Payments Information Act of 2002: Congress requires federal agencies to estimate and report an annual amount of improper payments for all programs and activities (CMS, 2016; Stockdale: 2009).

(4) Medicare Prescription Drug, Improvement, and Modernization Act of 2003: Congress authorized the RAC demonstration program for Part A and B of Medicare. The RACs were contracted to identify overpayments and underpayments based on a contingency fee. However, many criticize that this contingency fee incentivize RACs to aggressively seek overpayments in particular (AHIMA: 2010; CMS, 2016; Stockdale: 2009).

(5) Tax Relief and Health Care Act of 2006: RACs was authorized as a permanent program and extended to all states (CMS, 2016).

(6) Patient Protection and Affordable Care Act of 2010: Medicare was required to expand the Recovery Audit Contractor (RAC) program to the Medicare Part C (Medicare Advantage) and Part D (Prescription Drug Benefit) programs (CMS, 2016).

5.0 CLINICAL CODING VALIDATION, CODED DATA QUALITY, AND PRODUCTIVITY-DRIVEN PRACTICES

5.1 INTERNAL AUDIT PROGRAMS

The role of the internal auditor is "to independently and objectively analyze, review, and evaluate existing procedures and activities as well as reports and recommended changes to management on various operations of the organization." (Forman, 2013) Of course, internal audit can protect against the impact of noncompliance in health care organizations. However, the internal auditors primarily focus on non-financial operational audits.

Per Kusserow (2014), internal auditors have been subjected to pressure from two forces: (1) outsourcing of audit in all business sectors including health care; and (2) the ever-increasing overlap between compliance and audit as organizational functions. Accordingly, there are three major approaches to audit including: (1) outsourcing; (2) merging of the internal audit and compliance functions; (3) coordination and cooperation of the two functions.

5.1.1 Outsourcing approach

Outsourcing of internal audit function is common in smaller and mid-size health care organizations. Kusserow suggested that "the smaller an organization, the more likely it is to either eliminate or outsource internal auditor" due to unaffordability of such internal audit programs or services to small facilities such as physician offices (Kusserow, 2014). However, smaller facilities can overcome the problem of affordability by outsourcing services as well as internal audit to the same firm. A major limitation of this practice is the potential conflict of interests when the accounting or consulting firm that provides these services is the one in charge of conducting internal audit. To avoid such conflict of interest, it is advisable to hire another outsourcing firm to conduct external reviews from time to time (Boomershine et al, 2017; Forman, 2013).

5.1.2 Merging approach

In this approach, the internal audit and compliance functions are merged into one function to monitor and "ensure organizational compliance with all applicable laws, regulations, standards, policies and procedures as well as addressing high-risk areas" (Boomershine et al, 2017; Forman, 2013; Kusserow, 2014). Therefore, this merged unit: (1) has access to the organization's records, resources, and personnel; (2) perform independent reviews; and (3) monitor compliance. This practice is common in smaller and mid-sized facilities. However, in larger organizations, this can lead to persisting tension between the two functions with respect to fighting over resources and managerial attention.

5.1.3 Coordination approach

For many health care organizations, the best option is to "promote cooperation and coordination between the compliance officer and internal auditor functions" (Forman, 2013; Kusserow, 2014; Pitsikoulis & Doty, 2016). In this approach, "the internal auditor focuses on documents, operations, and controls" while the compliance officer focuses on compliance with rules and regulations and effective communication that builds trust between management and employees" (Kusserow, 2014). Coordination can enhance operational efficiency; this can be achieved by development of annual audit plans, ensuring compliance with high-risk areas, and including compliance-related audit. An example of this audit-compliance approach is to include compliance elements in the audit process. For example, “does medical record documentation support coding and billing” (Kusserow, 2014). However, this is the core component of the Clinical Documentation Improvement initiatives and programs in health care organizations (Boomershine et al, 2017; DeAlmeida, 2012; Pitsikoulis & Doty, 2016; Stegman, 2011; Stanfill, 2015).

Per Isenberg (2006), health plan administrators have started outsourcing chart audits. Particularly, the audit of outpatient charts (E/M codes, documentation, and level of billing). Further, practitioners should pay attention to level 4 and 5 E/M codes (CMS, 2016; Forman, 2013; Isenberg, 2006).

However, providers are advised to conduct external coding audits (at least once a year) to assess their coding quality as well as productivity (Wilson & Dunn, 2009; Stanfill, 2015; Boomershine et al, 2017). According Brownfield & Didier (2009), “external audits can objectively analyze operations, detect holes in the system, and uncover deficiencies that an internal audit program may miss. This outside review helps strengthen future internal audits by discovering how and why internal audits may have overlooked findings.”

5.2 THE RECOVERY AUDIT PROGRAM AND MEDICARE

5.2.1 Background on Medicare

Medicare is considered the nation's largest health insurance program for people 65 or older, and people with certain disabilities. Medicare insurance consists of four parts A, B, C, and D (Stockdale, 2009):

(1) **Part A** is Medicare (Hospital Insurance) that primarily covers inpatient hospital services, skilled nursing facility services, home health services, and hospice services (Stockdale, 2009; CMS, 2016);

(2) **Part B** is Medicare (Supplementary Medical Insurance) that covers other medical services such as physician visits, outpatient hospital care, laboratory services, and durable medical equipment (CMS, 2016).

(3) **Part C** is (Medicare Advantage MA) plan; it an optional plan that provides beneficiaries the benefits of Part A, B, and D (CMS, 2016).

(4) **Part D** is Medicare (Prescription Drug Plan PDP) which is a private insurance for drug coverage.

Part A and B constitute the fee-for-service portion of the program (Original Medicare) while Part C and D constitute the private insurance portion of the program (Stockdale, 2009). The Centers for Medicare and Medicaid Services (CMS) is the federal agency responsible for administering Medicare that was authorized under Title XVIII of the Social Security Act (Stockdale, 2009). CMS contracts with a variety of private entities to perform daily operations of the program such as claim payment, fraud detection, quality of care supervision (CMS, 2016; Stockdale, 2008; Stockdale, 2009; William & Cabin, 2014).

Further Medicare's contractors can be responsible for other administrative functions such as provider enrollment in Medicare, physician education on proper billing, appeals, improper payment recovery (CMS, 2014; Stockdale, 2009; William & Cabin, 2014). According to CMS, an improper payment "is any payment that should not have been made or that was made in an incorrect amount" which can include: (1) duplicate payments; (2) payments to ineligible recipients; (3) payments for ineligible services; or (4) payments for services not received. In Medicare, improper payments include both overpayment and underpayment to providers (CMS, 2016).

Due to its size, scope, and decentralized administrative structure, Medicare is at high risk of improper payments and fraud (Government Accountability Office, 2009). However, Stockdale suggests that improper payment cannot be considered as a measure of fraud although they sometimes could be fraudulent (CMS; 2016; Martin, 2016; Stockdale, 2009).

5.2.2 RAC Audit Program

The RAC was established initially as a 3-year demonstration program under section 306 of Medicare Prescription Drug, Improvement, and Modernization Act (CMS, 2016; Land, 2016; Wilson, 2009). Its primary purpose was to test the cost-effectiveness of using contract auditors to detect and correct underpayment and overpayments in the Medicare program for both Medicare as Secondary Payer (MSP) and non-MSP or Claim situations.

The demonstration program was converted to an ongoing part of the Medicare Integrity Program under section 302 of the Tax Relief and Health Care Act of 2006 (AHIMA, 2016). A June 2008 RAC evaluation report found that the claims program, by far the largest of the two programs, had corrected in pre-appeal findings $1.3 billion in errors in 2½ years with 96% being overpayments (CMS, 2016).

To identify improper payment, RACs are instructed to use two types of review processes (AHIMA, 2009; CMS, 2016; Martin, 2016; Stockdale, 2009):

1. **Automated review:** In this case, there is no human review of the claims or medical records. Alternatively, automated systems are used to automatically check claims within the claim processing system for evidence of improper payment or mistakes. Automated review is used when two conditions are met: (1) there is certainty that the service is not covered by Medicare; and (2) there is a written Medicare policy (Stockdal, 2009). However, when there is no policy by Medicare, RACs are required to perform complex review (CMS, 2016).

2. **Complex review:** this involves human review of the medical record as well as additional documentation supporting the claim (CMS, 2016; Stockdal, 2009). RACs must use a complex review when there is a high probability that the claim encloses overpayment.

In general, the following claims can be considered improper by RACs: (1) claims that are incorrectly coded; (2) claims that have incorrect payment amounts; (3) claim for services not covered by Medicare; (4) claim for services that are already provided (AHIMA, 2009; CMS, 2016; Linder, 2016; Stockdal, 2009; William & Cabin, 2014).

RACs can review all aspects of the supporting medical records. They are further advised to look for appropriate medical literature and clinical judgment when making complex claim demonstrations (AHIMA, 2009; CMS, 2016; Stockdal, 2009). However, CMS does not require RACs to hire nurses or certified coders for the record review. However, all Medicare contractors –including RACs- are required to hire one full-time medical director to supervise the claims review process (CMS, 2016; Stockdal, 2009). Table 11 represents each of the RAC contractors with the states that it covers.

Table 11. States Covered by Each RAC Contractor

Region	Contractor	States
A	Diversified Collection Services (DCS) with subcontractor PRG Shultz	Connecticut, Delaware, District of Columbia, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont
B	CGI Technologies and Solutions with subcontractor PRG Shultz	Illinois, Indiana, Kentucky, Michigan, Minnesota, Ohio, Wisconsin
C	Cannolly Consulting Associates with subcontractor Viant Payment Systems	Alabama, Arkansas, Colorado, Florida Georgia, Louisiana, Mississippi, New Mexico, North Carolina, Oklahoma, South Carolina, Tennessee, Texas Virginia, West Virginia
D	Health Data Insights with subcontractor PRG Shultz	Alaska, Arizona, California, Hawaii, Idaho, Iowa, Kansas, Missouri, Montana, Nebraska, Nevada, North Dakota, Oregon, South Dakota, Utah, Washington, Wyoming

Source of information: Centers for Medicare and Medicaid CMS

5.3 ADMINISTRATIVE DATA QUALITY & PRODUCTIVITY

Payment-by-results is a prospective payment method of payment for hospital care based on the actual volume and complexity of services rendered by hospitals in the United Kingdom. In that payment system, Healthcare Resource Groupings HRGs is the equivalent to Diagnosis Related Groups (DRGs) in the IPPS in the U.S. (CMS, 2016; Haliasos, et al, 2009).

A study of clinical coding audit in otolaryngology that was based in the United Kingdom suggested that coding variability cannot be eliminated but rather improved by ongoing education and audit programs (Nouraei et al, 2013). In this study, an audit of 3131 randomly-selected otolaryngology patients resulted in 13% change in the primary diagnosis (n=420), and 13% change in primary procedure (n=417). Further, in 44% of the cases (n=1420), there was at least one change of the original coding and in 16% (n=514), there was a change in the Health Resource Groupings (HRG). Primary diagnosis and primary procedure are what referred to as principle diagnosis and principle procedure in the U.S, respectively.

Another study was concerned about the effect of inaccurate coding on the departmental activities concluded that coding inaccuracies provide a distorted picture of departmental activities in addition to contributing to major financial disadvantages. This study was conducted in the department of neurosurgery in a hospital in the United Kingdom. Audit was performed by physicians due to “sub-specialism” of neurosurgery (Haliasos, et al, 2009).

A common coding discrepancy expected to result in a considerable financial impact was miscoding procedures (Moar & Rogers, 2011). Some studies have shown errors in coding 7-16% of the procedures (Dalal & Roy, 2009; Kwaja et al, 2009; Nouraei et al, 2009; Moar & Rogers, 2011). Some specialties like otolaryngology encompass a wide-range of procedures which are performed in “close anatomical proximity” and that ultimately affect coding accuracy in addition to coding productivity (Nouraei et al, 2013). In fact, neurology and otolaryngology specialties were associated with increased coding time and similar results were found in different medical specialties; urology (Beckley & Nouraei, 2009), neurosurgery (Haliasos et al, 2010), and surgery (Townley et al, 2011; Moar & Rogers, 2012; Naran et al, 2014).

However, it should be noted that coding accuracy requires familiarity with medical terminology, surgical techniques, and complex coding systems (Moar & Rogers, 2011). Coders familiar with medical terminology, and anatomy were found to be more efficient in performing complex coding (Moar & Rogers, 2011). Also, this study suggested that more experienced coders who are familiar with medical terminology and surgical techniques were more efficient in coding complex cases compared to their less experienced counterparts (Moar & Rogers, 2011).

Per Moar and Rogers (2011), “staff who record the codes are not familiar with technicalities and clinical staff are not familiar with coding protocols”. In this study, coding inaccuracies were found in all audited cases (n=21). Other studies (Dalal & Roy, 2009; Fillit et al, 2002; Kwaja et al, 2009; Nouraei et al, 2009) further assure that “clinicians are no better at providing codes than administrative staff” (Moar & Rogers, 2011). However, the same study found that professional coders who were familiar with anatomy and medical terminology are the most efficient when it comes to code assignment (Moar & Rogers, 2011).

Measuring coding time to assess productivity is not a new concept (Endicott, 2015; Martin, 2016, Stanfill, 2016). However, measuring coding time for clinicians performing coding is not a prevalent practice (Nouaei, 2013). More research has been conducted on coding productivity to meet the demand of administrative efficiency in the health care industry (Godbey-Miller, 2016). Thus, many health care organizations have created new coding metrics to monitor coding productivity of their coders (Boomershine, 2016; Godbey-Miller, 2016).

In fact, factors related to medical coders' background as well as education and training can further influence coding quality and productivity. A coder's survey conducted by the American Academy of Professional Coders (AAPC) in 2010 concluded that: (1) most coders are paid by the hour, and wages vary based on background;(2) coders in general report a positive relationship with physicians; (3) facilities prefer certified coders, (4) practice physicians regularly perform coding duties; (5) compliance risks are the biggest issue for coders; and (6) more coders work in physician practices than any other setting (AAPC, 2010). Based on this survey, coding variations can emerge because of varying educational background, relationship with physicians, and to which extent medical coders are involved in administrative tasks such as billing and compliance. In fact, coders' demographics in this survey were found to have a profound impact on coding productivity as well as coding quality.

Another survey conducted by HCPro in 2012 suggested that coders are usually involved in non-coding related tasks mostly including abstracting (79%). The other duties include: appealing denials, release of information, incomplete record management, chart assembly, RAC-related tasks, DRG, data set completion and others (AHIMA, 2008; AHIMA, 2013; HCPro, 2011). However, 10% of coders reported they spend 18 hours or more on non-coding related tasks compared to 27% who spend 3-5 hours per week performing such tasks (HCPro, 2011). This survey found out that coder's involvement in more administrative tasks can have a negative impact on coding productivity.

In a recent study conducted by the AHIMA Foundation, coders' demographics were found to have a great impact on coding quality as well as productivity. A study published by the AHIMA Foundation has suggested that coder's credentials have a significant impact on coding accuracy and productivity (AHIMA, 2016). The same study suggested that education and years of experience are important determinants of coding productivity (AHIMA, 2016). However, this study was based on anecdotal information provided by the professional coders who agreed to be interviewed for this study. Those coders provided their demographic information along with their perceptions of changes in coding accuracy, and coding productivity after ICD-10 implementation.

Furthermore, coding audit is believed to have a significant impact on coding productivity and quality (Linder, 2016). In most cases, physicians were the ones responsible for performing the audit for their medical specialties. Some of the studies, however, did not indicate who was responsible for the coding audit in their facilities. In 2010, The American Academy of Professional Coders (AAPC) announced a new Audit Services Division that provides full-service health care compliance and corporate integrity audits. The AAPC validates each audit, focusing on the areas of the organization that have the largest risk potential. The following services are provided by auditors: (1) insurance audit appeal; (2) coding and billing accuracy; (3) account receivable audits; (5) compliance audit; (6) ICD-10 readiness (AAPC, 2011).

However, the RAC program represents the largest clinical coding audit project in the United States. In 2009, the RAC demonstration program identified five key areas for improper payment: (1) Excisional Debridement; (2) Lysis of Adhesions; (3) Wrong Principal Diagnosis; (4) Coagulopathy; and (5) DRGs Designated as CC or MCC with Only One Secondary Diagnosis (Wilson, 2009). Many studies have recommended health care organizations to perform periodic coding audits to increase their coding quality, and productivity of their professional coders (Combs, 2016, Martin, 2016; Land, 2016).

The CMS's RAC program would present a valuable source of information on the impact of coding audit programs on hospitals across the nation. Unfortunately, such information is not yet provided by the CMS. Based on the American Hospitals Association (AHA), "the AHA created

RACTRAC—a free, web-based survey—in response to a lack of data provided by CMS on the impact of the RAC program on America's hospitals" (AHA, 2015).

Based on the AHA report on RACs (Quarter 1st , 2016), outpatient billing errors accounts for approximately 35% of the automated denials while inpatient coding errors (MS-DRG) accounts for only 2% of these denials (AHA, 2016). However, the most commonly cited reason for complex denial is inpatient coding (75%). Furthermore, MS-DRGs as well as other inpatient coding errors had the highest dollar impact on hospitals nationwide during the first quarter of 2016. Inpatient claim denials represent 44% of all denials nationwide. Interestingly, around 16% of reporting hospitals have claims denied for DRG validation converted into full medical necessity denials when the determination was appealed.

Complexity of resource grouping schemes can lead to inaccurate coding (Nouraei et al, 2013). A change in the resource grouping category usually results in greater financial impact than a change in coding within the same category (Moar & Rogers, 2011). Further, unclear documentation especially with respect to coexisting morbidities and complications can impact resource grouping accuracy (Moar & Rogers, 2011).

Per Stegman (2011) and based on the results from the RAC findings, the high-risk DRGs are (1) MS-DRGs 207 *(RESPIRATORY SYSTEM DIAGNOSIS W VENTILATOR SUPPORT 96+ HOURS)* and 208 (*RESPIRATORY SYSTEM DIAGNOSIS W VENTILATOR SUPPORT <96 HOURS)*; (2) MS-DRGs 166 *(OTHER RESP SYSTEM O.R. PROCEDURES W MCC),* 167 *(OTHER RESP SYSTEM O.R. PROCEDURES W MCC)*, and 168 *(OTHER RESP SYSTEM O.R. PROCEDURES W/O CC/MCC)*; (3) MS-DRGs 853 *(INFECTIOUS & PARASITIC DISEASES W O.R. PROCEDURE W MCC),* 854 *(INFECTIOUS & PARASITIC DISEASES W O.R. PROCEDURE W CC),* 855 *(INFECTIOUS & PARASITIC DISEASES W O.R. PROCEDURE W/O*

CC/MCC); (4) MS-DRGs 813 *(COAGULATION DISORDERS)*. These represent the following respectively: respiratory system diagnosis with vent support, closed biopsy of lung, procedure for infections and parasitic diseases, and coagulopathy. However, the “most identified improper payments due to the coding/DRG assignments were in cases where only one complication/comorbidity (CC) or major complication/comorbidity (MCC) were coded without clinical validation.” (AHIMA, 2014).

Furthermore, complexity of resource grouping found to have a significant impact on coding productivity. Complexity of resource grouping was also associated with increased coding time as coders might need more time in coding complex cases that are usually associated with higher DRGs (Linder, 2016, Moar & Rogers, 2011; Stanfill, 2016). A recent productivity study that was conducted in Rochester Regional Health has shown that more time is required to code complex cases suggesting a positive relationship between DRG weight and coding time (Linder, 2016). In fact, other studies concluded similar results with a 10% average decrease in productivity (Linder, 2016; Watzlaf et al, 2016, Alakrawi et al, 2017).

Therefore, coding audits should be incorporated as an integral part of coding workflow (Combs, 2016). Audit should be established as an ongoing process that requires collaboration between clinicians, coders, and auditors on a regular rather than ad-hoc basis (Nouraei et al, 2013). Nouraei and colleagues suggested that conducting second audit cycles can help reduce variability in coding accuracy as well as in coding time. However, the same study found that reduction in coding variability was significant for the primary procedure and secondary diagnoses but not for primary diagnosis when conducting a secondary review (Nouraei et al, 2013).

Martin (2016), suggests that “chronic conditions may be hardest for coding professionals to determine whether a code should be assigned” and that time might be wasted on additional diagnoses that have no influence on reimbursement. Also, Linder (2016) suggested that inpatient coding productivity is the slowest to improve due to many co-morbidities that require additional diagnosis codes. Clear documentation of coexisting morbidities and complications is critical to payment maximization. These conditions however, should be documented using specific medical terminology rather than general terms. Further, coding should be linked to databases to enable “easier data capture and retrospective audit” (Moar & Rogers, 2011).

Furthermore, clinical documentation is believed to have a significant impact on productivity (AHIMA, 2014; Boomershine, 2016; Endicott, 2015; Nourei, 2013). Clinical documents represent the main channel of communication between different caregivers. Effective communication is required for improving quality of care, ensuring efficient utilization of resources, and maintaining access to more health benefits (DeAlmeida et al, 2014; Stanfill, 2016).

Also, patient’s safety could be compromised if documentation of clinical episodes was not reliable or available to clinicians for clinical decision making in a timely manner. In fact, provision of subsequent healthcare can ultimately be very costly if data required for clinical decision making was not reliable or available (Bower-Jernigan et al, 2014).

In addition to its significant impact on patients’ care, clinical documentation is considered a critical factor in determining coding quality as well as coding productivity (Land, 2016). In fact, clinical documentation improvement (CDI) programs are believed to have a positive influence on coding quality and productivity (Bower-Jernigan et al, 2014; Combs; 2016). In contrast, documentation deficits could lead to more coding errors as well as increased time in coding patients’ charts (Combs, 2016). Also, Combs (2016) suggested that clinical documentation can have a direct influence on coding productivity. Therefore, it is important to constantly monitor the following CDI metrics and assess their impact on coding productivity: query rate, response time, and revenue impact (Combs, 2016).

Other studies have been conducted to examine coded administrative data quality in Canada and United States suggested issues with coded data accuracy and consistency and other issues related to coding productivity (Awad et al, 2014; Broberg et al, 2014; Gologorsky et al, 2014, Jensen et al, 2014; Nouraei et al, 2014; Rybnicek et al, 2014; Sacks et al, 2014).

In 2014, Goldinvaux and colleagues from Yale School of Medicine conducted a cross-sectional study to evaluate the ability of ICD-9-CM to identify preoperative anemia in patients undergoing spinal fusion. This study examined data for 260 patients at an academic medical center. Only 3.8% (n=10) received ICD-9-CM code for anemia and 7 of these cases were miscoded. According to this study, administrative data are compiled based on ICD-9-CM codes that are generated based on provider input and professional coders' abstraction for reimbursement purposes. Therefore, this data could be "prone to omission of details and may not accurately represent the entire patient population" (Goldinvaux et al, 2014).

Also, Broberg and colleagues (2014) performed a study to evaluate the accuracy of ICD-9-CM data for detection and categorization of adult congenital heart disease (ACHD) patients using EHR data. An EHR algorithm for ACHD was developed and applied to 740 patients. The sensitivity and specificity for this algorithm were 99 and 88%, respectively. However, of 411 non-ACHD patients, 49 were incorrectly categorized as ACHD based on ICD-9-CM codes. Of 329 ACHD patients, 326 were correctly categorized and the ACHD defect subtype was correct in 80% of the patients. This study suggested that ICD-9-CM data can be utilized in identification of ACHD patients based on its excellent sensitivity and good specificity values. However, using this data for identification of the defect subtype is less robust since the "accuracy of sub-type categorization varied greatly by defect group” (Broberg et al, 2014). Furthermore, less familiarity with ACHD will likely result in decreased coding accuracy and increased use of non-specific codes such as “other congenital heart disease”.

However, accuracy of coded data varies greatly across medical specialty (Broberg et al, 2014; Jensen et al, 2014; Rybnicek et al, 2014). Based on their study, Rybnicek and colleagues suggested that administrative coding is specific but not sensitive for identifying eosinophilic esophagitis (EoE). In this study, all diagnostic and procedures codes of EoE patients were obtained using University of Carolina data warehouse (2008-2011). Specificity and sensitivity were calculated based on data for 308,372 patients: 99% and 37% respectively. Consequently, using ICD-9-CM data for identifying EoE cases “will still miss number of cases, but those identified in this manner are highly likely to have the disease” and therefore coded data can be used as an effective tool to study EoE patients in large-scale administrative databases (Rybnicek et al, 2014).

Furthermore, a study published in 2014 by Jensen, Cookes and Davis suggested that there are many potential pitfalls of using administrative coded data (ICD-9-CM) in analyses related to epidemiology, clinical effectiveness, risk assessment, healthcare utilization, and making informed decisions with respect to clinical care and health policy. In addition to confirming previous findings of false positive miscoding errors, this study highlights findings related to false negative miscoding errors and subsequent implications of these miscoding errors on data accuracy and conclusions that can be drowned from this data (Jensen et al, 2014). It is advised to take caution when utilizing administrative data for less common conditions such as Down Syndrome, eosinophilic esophagitis, congenital heart disease, genetic blood disorders, and surgery (Broberg et al, 2014; Jensen et al, 2014; Nouraei et al, 2014; Rybnicek et al, 2014; Sacks et al, 2014).

Likewise, productivity can be influenced by medical specialty (Nouraei et al, 2014, Combs, 2016; Linder, 2016).

Administrative coded data have always been criticized of inaccuracy because it is usually collected by coders who have no direct contact with the care process (Chang, 2015). Per Chang, "supporters of administrative databases have noted that there are already extensive processes in place to ensure the accuracy of administrative coding." Chang suggested that inaccuracies related to administrative databases are more likely limited to diagnosis coding rather than procedure coding since there is a higher impact of procedure coding in reimbursement. Inaccuracies in diagnostic coding are likely to be randomly distributed and therefore less likely to bias any findings (Chang, 2015).

Sacks and colleagues (2014) conducted a retrospective review to evaluate hospital readmissions in surgical patients using administrative coded data. This review includes all consecutive patients discharged from general surgery services at a tertiary care, university-affiliated hospital (2009-2011). This study reported "significant limitations of the Hospital-Wide All-Cause Unplanned Readmission Measure developed by CMS" (Sacks et al, 2014).

These varying practices have significantly contributed to coding variation as well as coding discrepancies. Such practice variations can influence two major aspects of clinical coding: (1) quality, accuracy of coded data, and amount of coding errors; and (2) time required to perform clinical coding as a clinical task.

5.4 COMPUTER-ASSISTED CODING (CAC)

The advancement of Information Technology (IT) has led to a new generation of software applications that would inevitably enhance the efficiency of operations and reduce the cost of direct and indirect health care services including coding. Although medical coding represents one of the areas in which information technology has not been fully utilized, few attempts have shown some success in employing technology to improve coding operations (Land, 2016; AHIMA, 2013). Such efforts progressed gradually from the attempts of introducing early encoders in the 1980s to the revolutionary use of NLP in Computer Assisted Coding (CAC) software and automated coding systems (AHIMA, 2013).

Computer-Assisted coding (CAC) is defined by the American Health Information Management Association (AHIMA) as the: "... use of computer software that automatically generates a set of medical codes for review, validation, and use based upon clinical documentation provided by healthcare practitioners." There are currently two available CAC models: natural language processing (NLP) and structured input (SI). Both models require human intervention to a certain level. The function of NLP in CAC is to convert words into codes to generate a set of suggested codes to be reviewed, validated, or edited by coders.

With the advancement of informatics and technology, along with the federal incentives to adopt EHRs represented by The American Reinvestment and Recovery Act (ARRA/HITECH), and the cost of healthcare increasing year after year, healthcare decision makers are driven to pay greater attention to coding as it plays a critical role in reimbursement, research, and public health reporting. Computer-Assisted Coding (CAC) has been shown to increase productivity, improve accuracy, and promote consistency of coding in addition to ultimately reducing overall cost (AHIMA, 2014; Houser & Meadow, 2017; Tully and Charmichael, 2012; Stanfill; 2016). Health IT vendors have provided some excellent CAC solutions that have been implemented in different hospitals across the country (AHIMA, 2013; Linder, 2016).

However, encoders have been used in healthcare for more than 20 years. Encoders can be defined as a "tool used to automate the coding process that is similar to using a code book to assign codes. Encoders are computer software programs that usually prompt the coder to evaluate documentation and coding rules during the process of assigning a code" (AHIMA, 2005; Stanfill, 2016). In general, using encoders is intended to decrease variability in the code assignment process and therefore increasing accuracy of clinical coding (AHIMA, 2005; Houser & Meadow, 2017; Stanfill, 2015).

Aside from encoders, two different types of CAC can be recognized in major health systems; structured input and CAC using NLP. While the former is mainly menu-driven, the latter utilizes narrative-text form which is promising as more qualitative data are needed to serve the purpose of quality of healthcare and patient's safety (Alakrawi, 2016; Salmasian, 2013). Emergence of CAC applications has played a major role in shaping the health care industry in general and coding in particular. CAC has added another layer of technology reliance (AHIMA, 2013) and therefore has initiated a major shift related to the role of clinical coders in the automated coding workflow environment. Instead of manual assignment of codes, coders have been more engaged in reviewing and validation of codes proposed by the CAC system (AHIMA, 2013; Bronnert et al, 2011; Houser & Meadow, 2017).

The AHIMA's CAC Industry Outlook and Resource Report (2011) provides valuable resources for all HIM professionals on current coding practices, envisioned changes in these practices, and required steps to enable a smooth transition to automated coding workflow using CAC technologies. In fact, Bronnert and colleagues (2011) identify specific gaps in the formal educational system and other training programs in addition to the skills and competencies needed to fill those gaps.

However, there are some major challenges to full adaptation of CAC systems. CAC applications work very well for outpatient coding. However, reliability of CAC applications in inpatients coding has not given comparable results (AHIMA, 2013; Linder, 2016). Inpatient coding is more complex than outpatient coding; coding complexity increases with inpatient hospitalization where more documents needed to be reviewed. Revenue generated by inpatient services is greater than what is generated by outpatient services; acute care accounts for 51% of Medicare spending compared to other healthcare services (CMS, 2012; CMS, 2016).

Regardless, rapid adaptation of health information technology in general, and CAC in particular, made it possible for healthcare providers to integrate remote coding in the conventional coding workflow and operations. Remote coding has contributed greatly to fill in coding-related vacancies across the United States (AHIMA, 2013). Also, CAC provides a direct link between coding assignment and clinical documentation by (1) enabling tractability to the source documents used in the codes assignment; and (2) providing coding audit trails (AHIMA, 2013). Coding errors are consistent in CAC applications as opposed to coders; by "machine learning", such applications can be trained over time to enable the correct code assignment (Alakrawi, 2016; Salmasian, 2013).

CAC cannot eventually replace human coders in inpatient settings but could raise the level of coding function to an analyst, enhance the overall coding workflow, improve coding quality by providing a direct link to documentation, and foster a transition to ICD-10-CM/PCS that requires a more sophisticated information technology architecture (AHIMA, 2013; DeAlmieda, 2012; Martin, 2016). Nevertheless, the field of clinical coding has been rapidly evolving because of the increasing complexity in the field of HIM (AHIMA, 2014). The next generation of the CAC applications will reflect these changes in the fields of HIM.HIT). The role of such applications will be shifted from coding to clinical validation (AHIMA, 2014).

Per CMS, “Clinical validation is an additional process that may be performed along with DRG validation. Clinical validation involves a clinical review of the case to see whether the patient truly possesses the conditions that were documented in the medical record. Recovery Auditor clinicians shall review any information necessary to make a prepayment or post-payment claim determination. Clinical validation is performed by a clinician (RN, CMD or therapist). Clinical validation is beyond the scope of DRG (coding) validation, and the skills of a certified coder. This type of review can only be performed by a clinician or maybe performed by a clinician with approved coding credentials.” (AHIMA, 2014). The concept of clinical coding validation will be discussed in the following chapter.

6.0 STUDY SIGNIFICANCE

Cost containment and quality of care have always been major challenges to the health care delivery system in the United States. Health care organizations utilize coded clinical data for health care monitoring, and reporting that includes a wide range of diseases and clinical conditions along with adverse events that could occur to patients during hospitalization. Governmental organizations- such as the Centers for Medicare & Medicaid Services (CMS), and the Agency for Healthcare Research and Quality (AHRQ)- also utilize coded clinical data for assessing patient safety, and quality of care through performance indicators used to compare hospital performance across the country. The results of these assessments are frequently released to the public to aid health care consumers in making informed decisions related to treatment options, and health care utilization.

Also, coded clinical data can have a major impact on population health since it is used to determine the leading causes of mortality and morbidity in the United States. Thus, it is a critical factor for promoting fund for healthcare services and research. Furthermore, it has other uses in research, education, resource allocation, and health service planning.

Thus, it is very critical to maintain high quality standards of clinical coded data and promote funding for health care research that addresses clinical coding, due to its direct impact on individual health outcomes as well as population health. With the rapid adoption of health information technology (HIT), there has been a rising demand for effective and data-driven decision-making strategies. Coded clinical data needed for such decision-making should be reliable and available to users in a timely fashion.

Clinical coding can be influenced by many factors such as the clinical documentation within the health record, and education and training of the professional coders. Clinical documents represent the main channel of communication between different caregivers. Effective communication is required for improving quality of care, ensuring efficient utilization of resources, and maintaining access to more health benefits (DeAlmeida et al, 2014; Stanfill, 2016).

Also, patient's safety could be compromised if documentation of clinical episodes was not reliable or available to clinicians for clinical decision making in a timely manner. In fact, provision of subsequent healthcare can ultimately be very costly if data required for clinical decision making was not reliable or available (Bower-Jernigan et al, 2014).

In addition to its significant impact on patients' care, clinical documentation is considered a critical factor in determining coding quality as well as coding productivity (Land, 2016). In fact, clinical documentation improvement (CDI) programs are believed to have a positive influence on coding quality and productivity (Bower-Jernigan et al, 2014; Combs; 2016). In contrast, documentation deficits could lead to more coding errors as well as increased time in coding patients' charts (Combs, 2016).

Furthermore, attributes related to the professional coders' background can have significant influences on coding quality, consistency, and productivity. A coder's survey conducted by the American Academy of Professional Coders (AAPC) in 2010 concluded that: (1) most coders are paid by the hour, and wages vary based on background;(2) coders in general report a positive relationship with physicians; (3) facilities prefer certified coders, (4) practice physicians regularly perform coding duties; (5) compliance risks are the biggest issue for coders; and (6) more coders work in physician practices than any other setting (AAPC, 2010).

Based on this survey, variations of coding practices were attributed to coders' education, years of experience, relationship with physicians, and coders' involvement in other administrative tasks such as billing and compliance. In fact, coders' demographics in this survey were found to have a profound impact on coding productivity as well as coding quality.

Another survey that was conducted by HCPro in 2012 suggested that coders are usually involved in non-coding related tasks, mostly including abstracting (79%). The other duties include: appealing denials, release of information, incomplete record management, chart assembly, RAC-related tasks, DRG, data set completion and others (AHIMA, 2008; AHIMA, 2013; HCPro, 2011). However, 10% of coders reported they spend 18 hours or more on non-coding related tasks compared to 27% who spend 3-5 hours per week performing such tasks (HCPro, 2011). This survey found out that coder's involvement in more administrative tasks can have a negative impact on coding quality and consistency. In addition, it found that coding productivity decreased as coders were constantly distracted by other tasks.

Clinical data quality might suffer because of variations in coding practices. Furthermore, coding productivity can also be affected due to these coding variations. Therefore, this dissertation research aimed at identifying current coding trends. and other factors that could influence coding quality and productivity through two major emphases: (1) quality of coded clinical data; and (2) productivity of clinical coding.

It will also examine the relationship between coding quality and coding productivity which represents a major strength of this study as no previous research has been conducted in this area. Previous studies have only tried to establish a link between coding quality and productivity based on qualitative data rather than quantities evidence.

Thus, data analytics will be performed on coding quality and productivity data sets to explore the most common trends related to clinical data quality, and coding productivity. Specifically, major factors that influence coding quality as well as productivity will be identified using different data analytics and statistical techniques and mixed research designs.

To summarize, the significance of this study lies in three major premises: (1) this dissertation research focuses on coding, a critical function that is underutilized in health care research; (2) it applies a new approach utilizing quantitative and qualitative methods along with statistics and data analytics techniques to identify factors that could influence clinical coding quality and productivity; and (3) it tries to establish a connection between coding quality and productivity, a topic that has never been addressed based on real data analysis.

7.0 SPECIFIC AIMS AND RESEARCH QUESTIONS

This study aims at identifying determinants of coding quality and productivity, error patterns, and current trends of professional coding practices. This can be achieved through the following specific aims:

Specific Aim I: Identify factors that could influence coding accuracy:

1. Length of stay (LOS)
2. Case mix index (CMI)
3. DRG relative weight
4. MS_DRG categories that are more often impacted by coding discrepancies
5. Coding errors at the major digit level versus the minor digit level

Specific Aim II: Identify documentation discrepancies that could influence coding quality.

Specific Aim III: Identify the impact of coding errors on CMI and hospital's payment.

Specific Aim IV: Identify individual and facility-related factors that could influence coding productivity:

1. Length of stay (LOS)
2. DRG relative weight
3. Case mix index (CMI)
4. Facility bed capacity (bed size)
5. Teaching status
6. Trauma status

Specific Aim V: Explore the relationship between coding productivity and coding quality

Specific Aim VI: Develop a predictive model to predict coding productivity and coding quality based on the individual and facility-related factors.

7.1.1 Specific Aim I: Identifying factors that could influence coding accuracy:

7.1.1.1 Length of Stay (LOS)

Extended Average Length of Stay (ALOS) can be indicative of more complex cases being treated. Thus, coders tend to review more documents to assign the right codes and maintain correct sequencing of these codes when working with more complex cases. However, this can lead to higher probability of coding errors associated with increased ALOS.

7.1.1.2 Case Mix Index (CMI)

Related to length of stay is the hospital Case Mix Index (CMI). According to CMS, "A hospital's CMI represents the average diagnosis-related group (DRG) relative weight for that hospital. It is calculated by summing the DRG weights for all Medicare discharges and dividing by the number of discharges." Therefore, it is expected to some extent that there is a correlation between coding quality and CMI. Particularly, higher CMI is associated with greater coding quality: higher CMI can be translated into higher reimbursement which can trigger external audit activities.

7.1.1.3 DRG Relative Weight

Based on the preceding discussion, it is expected to find a significant correlation between coding quality and DRG relative weight.

7.1.1.4 MS_DRG categories

Some MS-DRG categories, Major Diagnostic Categories (MDCs), are more affected by coding errors than others.

7.1.1.5 Digit Level

Coding errors are more likely to occur at the minor digit level than at the major digit level: the more specific the case, the more difficult it is to select the appropriate code among different related alternatives. Therefore, as the degree of specificity in the coding assignment process increases, the coding errors increase. This results in more coding errors at the major digit level versus the minor digit level. Table 12 provides a summary of Specific Aim I.

Table 12. Summary of Specific Aim I

Factor		Null Hypothesis	Research Hypothesis
1	Length of stay (LOS)	There is no correlation between coding quality and ALOS (H_0): $\rho = 0$	There is a significant correlation between coding quality and ALOS (H_1) $\rho > 0$
2	Case mix index (CMI)	There is no correlation between CMI and coding quality (H_0): $\rho = 0$	There is a positive correlation between CMI and coding quality (H_1) $\rho > 0$

Table 12 (continued)			
3	DRG relative weight	There is no correlation between coding quality and DRG relative weight (H$_0$): $\rho = 0$	There is a significant correlation between coding quality and DRG weight (H$_1$) $\rho > 0$
4	MS-DRG category	There is no difference in coding quality across MDCs. (H$_0$): $\mu_1 = \mu_2 = \mu_3 = \mu_{4....\neq}\,\mu_n$	Coding quality varies across different MDCs (H$_1$): $\mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_{4....\neq}\,\mu_n$
5	Digit level	There is no difference in coding error at the major and minor digit levels (H$_1$) $\mu 1 = \mu 2 = \mu 3$	There are more coding errors at 4th and 5th digit levels (H$_1$) $\mu 1 \neq \mu 2 \neq \mu 3$

7.1.2 Specific Aim II: Identifying documentation discrepancies that could influence coding quality

Documentation discrepancies can also have a significant impact on coding quality. Documentation issues can be related to lack of documentation, conflicting or incomplete documentation, and overlooking some documents in the process of code assignment. Therefore, it is important to identify documentation-related issues that can lead to inappropriate code assignment.

7.1.3 Specific Aim III: Identifying impact of coding errors on CMI and hospital's payment

Coding errors can have an impact on facility's payment and consequently its case mix index. Assigning inaccurate codes can lead to inaccurate DRG assignment which can eventually lead to erroneous calculation of the facility's CMI. Table 13 provides a summary of Specific Aim III.

Table 13. Summary of Specific Aim III

Factor		Null Hypothesis	Research Hypothesis
1	Facility payment	There is no correlation between coding quality and payment (H_0): $\rho = 0$	There is a significant correlation between coding quality and payment (H_1) $\rho > 0$
2	Case mix index (CMI)	There is no correlation between coding quality and CMI (H_0): $\rho = 0$	There is a significant correlation between coding quality and CMI (H_1) $\rho \neq 0$

7.1.4 Specific Aim IV: Identifying individual and facility-related factors that could influence coding productivity

Coding productivity can be influenced by factors that are related to individual patients as well as to the healthcare facility or system in which coding takes place. Individual and facility related factors include the following: (1) length of stay (LOS) (2) DRG relative weight; (3) case mix index (CMI); (4) facility bed capacity (bed size); (5) teaching status; and (6) trauma status.

7.1.4.1 Length of Stay (LOS)

Extended Average Length of Stay (ALOS) can be indicative of more complex cases being treated. Thus, coders will be demanded to review more documents to assign the right codes and maintain correct sequencing of these codes which can increase coding time. Furthermore, surgical cases where patients undergo surgical procedures are expected to have increased coding time compared to cases with surgeries are performed.

7.1.4.2 Case Mix Index (CMI)

Related to length of stay is the hospital Case Mix Index (CMI). A hospital's CMI represents the average diagnosis-related group (DRG) relative weight for that hospital. It is calculated by summing the DRG weights for all Medicare discharges and dividing by the number of discharges." Therefore, it is expected to some extent that there is a correlation between coding time and CMI. Particularly, higher CMI is associated with greater coding quality: higher CMI can be translated into higher reimbursement which can trigger external audit activities.

7.1.4.3 DRG Relative Weight

Based on the preceding discussion, it is expected to find a significant correlation between the coding time and DRG relative weight.

7.1.4.4 Bed Size

Facility size can be measured by bed count. Healthcare facilities with greater capacity have higher volume of patients with wide variety of conditions (usually more complex cases) which can have a positive impact on coding quality. Furthermore, healthcare facility with higher bed count tend to hire more experienced coders to maintain quality of clinical coding that in turn has a significant influence on revenue cycle and “cash-flow”. However, quality can be compromised when an increasing demand for productivity is created by a higher-volume of patients. Therefore, the direction of the relationship between coding quality and facility size is inconclusive.

7.1.4.5 Teaching Status

Teaching facilities are expected to have increased coding time compared to non-teaching facilities.

7.1.4.6 Trauma Status

Trauma centers are expected to have increased coding time due to complexity of cases treated in these facilities. Table 14 provides a summary of Specific Aim IV.

Table 14. Summary of Specific Aim IV

Factor		Null Hypothesis	Research Hypothesis
1	Length of stay (LOS)	There is no correlation between coding time and ALOS (H_0): $\rho = 0$	There is a positive correlation between coding time and ALOS (H_1) $\rho > 0$
2	DRG relative weight	There is no correlation between coding time and DRG relative weight (H_0): $\rho = 0$	There is a positive correlation between coding time and DRG weight (H_1) $\rho > 0$

Table 14 (continued)			
3	Case mix index (CMI)	There is no correlation between CMI and coding time (H_0): $\rho = 0$	There is a positive correlation between CMI and coding time (H_1) $\rho > 0$
4	Bed size	There is no difference in coding time between healthcare facilities with different bed counts. (H_0): $\mu 1 = \mu 2 = \mu 3$	Coding time is affected by facility size (bed count) (H_1) $\mu 1 \neq \mu 2 \neq \mu 3$
5	Teaching status	There is no difference between teaching and non-teaching facilities in coding time (H_0): $\mu 1 - \mu 2 = 0$ ($\mu 1 = \mu 2$)	Coding time is different between teaching and non-teaching facilities (H_1): $\mu 1 - \mu 2 \neq 0$ ($\mu 1 \neq \mu 2$)
6	Trauma Status	There is no difference between trauma and non-trauma facilities in coding time (H_0): $\mu 1 - \mu 2 = 0$ ($\mu 1 = \mu 2$)	Coding time is different between trauma and non-trauma facilities (H_1): $\mu 1 - \mu 2 \neq 0$ ($\mu 1 \neq \mu 2$)

7.1.5 Specific Aim V: Explore the relationship between coding productivity and coding quality

Coding productivity and quality can be perceived as conflicting values when it comes to clinical coding. However, coders do not have to sacrifice quality for quantity. Therefore, the correlation between coding quality and coding productivity will be examined to identify the type of relationship between both variables. Table 15 provides a summary of Specific Aim V.

Table 15. Summary of Specific Aim V

Factor		Null Hypothesis	Research Hypothesis
1	Coding productivity	There is no correlation between coding quality and coding productivity (H_0): $\rho = 0$	There is a significant correlation between coding quality and coding productivity (H_1) $\rho > 0$

7.1.6 Specific Aim VI: Develop predictive models to predict coding productivity and coding quality based on the individual and facility-related factors.

A predictive model will be developed to predict coding quality and productivity based on all significant factors examined in this research study.

8.0 METHODOLOGY

8.1 STUDY DESIGN

This study is a descriptive study that examined coding quality and productivity databases provided by Ciox Health. Quantitative as well as qualitative methods were utilized to answer the research questions.

First, data was tabulated and organized into graphs using descriptive statistics. Using descriptive statistics is an important step to (1) explore the distribution of all variables across the selected sample (normal vs skewed); (2) account for any missing data in the subsequent analysis (pairwise vs listwise analysis); and (3) determine the types of tests to be used i.e. parametric vs. non-parametric tests. Second, bivariate analysis was performed in addition to tests of significance to look for significant correlations and relationships between different variables. Linear and multiple regression was used to develop a predictive model of coding quality as well as productivity.

Periodic reports were provided to Ciox Health and conference calls were conducted on a monthly-basis to discuss the progress of this research. Based on study findings, recommendations were provided to Ciox health on how to improve coding quality and productivity. Furthermore, a predictive model was developed to predict coding quality and productivity based on significant predictors.

8.2 SAMPLING DESIGN

This study utilizes nonprobability sampling design (convenient sampling). However, large sample size can account for sampling bias that could be produced using nonprobability sampling. Samples will be representative of the general population of the United States. Representativeness is critical to establish external validity requirements needed for generalizing the study findings. Samples are representative with respect to the following: (1) demographically representativeness (age, gender, conditions); (2) geographically representativeness; and (3) diseases and procedures.

8.3 SAMPLE SIZE

The accuracy data includes a total of 106 audit reviews, including 57 facilities, were conducted by Ciox Health in 2011, 2012, and 2013. The total number of inpatient cases reviewed is 1,010 cases (13,713 ICD-9-CM codes). In contrast, ICD-10 productivity data includes a total number of 323,112 cases for a 10-month period (October 2015-July 2016).

8.4 DATA COLLECTION

The accuracy data contains coded ICD-9-CM data that has been audited to improve quality and accuracy of coding and billing practices of Ciox’s clients. Upon audit, clients are given feedback on their coding accuracy level with respect to: (1) ICD-9-CM codes assignment (can include inpatient, outpatient, and long-term care data); (2) MS-DRG grouping; and (3) the reasons for codes or DRG changes to learn the best coding practices. The productivity data contains data on coding time that was automatically recorded by Ciox professional coders.

All databases are de-identified and therefore do not contain any patient identifying information. The IRB office at the University of Pittsburgh has approved this dissertation research and Data Use Agreement (DUA) was determined by the University of Pittsburgh Office of Research.

8.4.1 Accuracy Data Set

Accuracy data was received in the form of individual reports. Thus, the data was re-structured, organized, and tabulated in a data set. This data set includes the following data items: (1) case ID; (2) facility ID; (3) code ID; (4) length of stay; (5) CMI; (6) DRG relative weight; (7) DRG description; (8) DRG relative weight; and (9) payment amount; and (10) ICD-9-CM codes. DRG variables, ICD-9-CM codes, and payment were redundant (measured for both coders and auditors). In addition, accuracy scores were assigned to all codes based on the following ranked agreement:

5	All digits are captured by codes assigned
4	One digit is different between the codes assigned
3	Two digits are different between the codes assigned
2	Three digits are different between the codes assigned
1	>3 digits are different between the codes assigned
0	0= Not coded (added by reviewer)

Finally, a total accuracy score was assigned to each case. The following formula was used to calculate the overall accuracy score:

$$\frac{\text{Sum of accuracy scores of individual codes}}{\text{sum of highest possible score of individual codes}} * 100$$

Using this formula as opposed to the conventional accuracy rate formula has a major advantage in accounting for depth of the coding. In other words, the number of codes in each case is considered when measuring accuracy.

8.4.2 Productivity Data Set

The productivity data set used data compiled by Ciox for a 10-month period (October 2015- July 2016). This data was selected specifically so researchers could focus on coding productivity after ICD-10 was in use for a longer period. The data was analyzed, organized, and influential outliers were removed from the final analysis. Influential outliers include length of stay greater than 365 and coding time greater than 10 hours. Productivity is defined as the time required to code a patient record measured in minutes. The productivity data set includes the following data items: (1) case ID; (2) facility ID; (3) LOS; (4) CMI); (5) DRG; (6) DRG relative weight; (7) DRG description; (8) MDC; (9) MDC description; (10) teaching status; (11) trauma level; (12) coding time (in minutes); (13) coding complete data.

8.5 DATA ANALYSIS

This is an exploratory study and therefore starting with descriptive statistics is important to identify the current state of coding and issues related to its quality and practice. Univariate and bivariate analyses was performed on Ciox data followed by regression analyses. Quantitative analysis was performed using SAS version 9.4 and SPSS 17 while qualitative analysis was performed using Nvivo Qualitative Analysis Software. Outliers as well as missing data was adjusted for during this stage as well:

Handling missing data: data was examined to determine whether data is missing completely at random (MCAR), missing at random (MAR), or not missing at random (NMAR):

(1) MCAR: a missing value is unrelated to any other value

(2) MAR: a missing value is related to other observed value(s)

(3) NMAR: a missing value is related to other missing values

In case of missing data, listwise deletion, or complete case analysis was performed. If the percentage of missing data is high, multiple imputation (MI) was performed.

Dealing with outliers: Data analysis was performed with as well as without extreme values (outliers). Univariate outliers do not usually represent a significant problem in the analysis. However, outlier diagnostics was performed to determine whether regression outliers are influential.

Furthermore, recommendations have been provided on how to improve coding quality and productivity and models to predict both variables were developed based on the data analysis.

9.0 RESULTS

In this section, results of this dissertation research are presented with respect to the following sections:

I. Identifying factors that could influence coding accuracy.

II. Identifying documentation discrepancies that could influence coding quality.

III. Identifying the impact of coding errors on CMI and hospital's payment.

IV. Identifying individual and facility-related factors that could influence coding productivity.

V. Exploring the relationship between coding productivity and coding quality

VI. Developing a predictive model to predict coding productivity and coding quality based on the individual and facility-related factors.

9.1 IDENTIFYING FACTORS THAT COULD INFLUENCE CODING ACCURACY

In this section, univariate and bivariate analyses were conducted to identify the influence of LOS, CMI, and DRG relative weight on coding accuracy. Descriptive statistics revealed interesting patterns related to coding errors and DRG changes. However, no significant correlations were found between coding accuracy and LOS, CMI, and DRG relative weight.

9.1.1 Descriptive Statistics

The total number of cases analyzed is equal to 1,010 cases with a total number of 13,713 codes. Quality data is ICD-9 audited data that includes 12, 938 diagnosis codes (94.35%) and 775 procedure codes (5.65%). The average LOS in this sample is 3.85 days with a standard deviation of ± 4.07 days (LOS ranges from 1 to 60 days). Also, Average CMI in this sample is 1.10 with a standard deviation of ± 0.19 points. The DRG relative weight ranges from .43 to 7.80 with a mean of 1.12 and standard deviation of ± .83. Interestingly, differences in payment (based on coding audit) ranges from -$28,587.74 to +$26,060.16. Table 16 represents distribution of LOS, CMI, DRG Relative Weight and Payment Difference.

Table 16. Descriptive Statistics (Accuracy Data)

	N	Mean	Standard Deviation	Minimum	Maximum
LOS	1,010	3.85	4.07	1.00	60.00
CMI	1,010	1.04	0.19	0.88	1.74
DRG Relative Weight	1,010	1.12	0.83	0.43	7.80
Payment Difference	1,010	$-34.91	$1603.62	$-28587.74	$26060.16

As shown in Table 17 and figure 10, the top 10 principle diagnoses in this sample include the following: pneumonia, atrial fibrillation, septicemia, urinary tract infection, chest pain, chronic obstructive bronchitis, transient cerebral ischemia, syncope and collapse, congestive heart failure, and acute on chronic diastolic heart failure.

Table 17. Most Frequent Principle Diagnoses

ICD-9-CM Code		Description	Number of Cases (%)
1	486	Pneumonia, organism NOS	93 (9.21%)
2	427.31	Atrial fibrillation	62 (6.14%)
3	389	Septicemia NOS	58 (5.74%)
4	599.0	Urinary tract infection NOS	52 (5.15%)
5	786.59	Chest pain NEC	33 (3.27%)
6	491.21	Chronic Obstructive bronchitis with acute exacerbation	29 (2.87%)
7	435.9	Transient cerebral ischemia NOS	22 (2.18%)
8	780.2	Syncope and collapse	22 (2.18%)
9	428.0	Congestive heart failure NOS	20 (1.98%)
10	428.33	Acute on chronic diastolic heart failure	19 (1.88%)
Total			410 (40.59%)

Also, the top 10 secondary diagnoses are the following: pnuemonia, urinary tract infection, hyposmolality, acute kidney failure, congestive heart failure, acute respiratory failure, end stage renal disease, dehydration, atrial fibrilation, and chronic airway obstruction. Table 18 and figure 11 represent distributions of the most frequent secondary diagnoses.

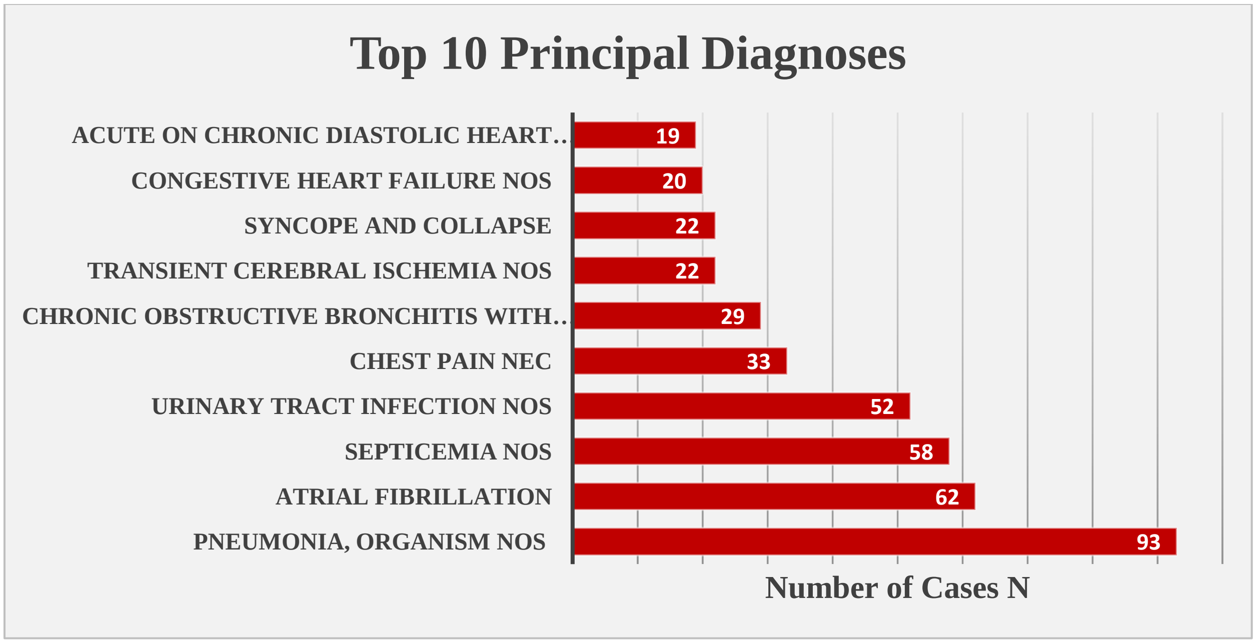


Figure 10: Top 10 Principal Diagnoses

Table 18. Most Frequent Secondary Diagnoses

ICD-9-CM Codes and Descriptions			N (%)
1	486	Pneumonia, organism NOS	37 (3.66%)
2	599.0	Urinary tract infection NOS	31 (3.07%)
3	276.1	Hyposmolality	30 (2.97%)
4	584.9	Acute kidney failure NOS	20 (1.98%)
5	428.0	Congestive heart failure NOS	19 (1.88%)
6	518.81	Acute respiratory failure	19 (1.88%)
7	585.6	End stage renal disease	17 (1.68%)
8	276.51	Dehydration	14 (1.39%)
9	427.31	Atrial fibrillation	13 (1.29%)
10	496	Chronic airway obstruct NEC	13 (1.29%)
Total			213 (21.09%)

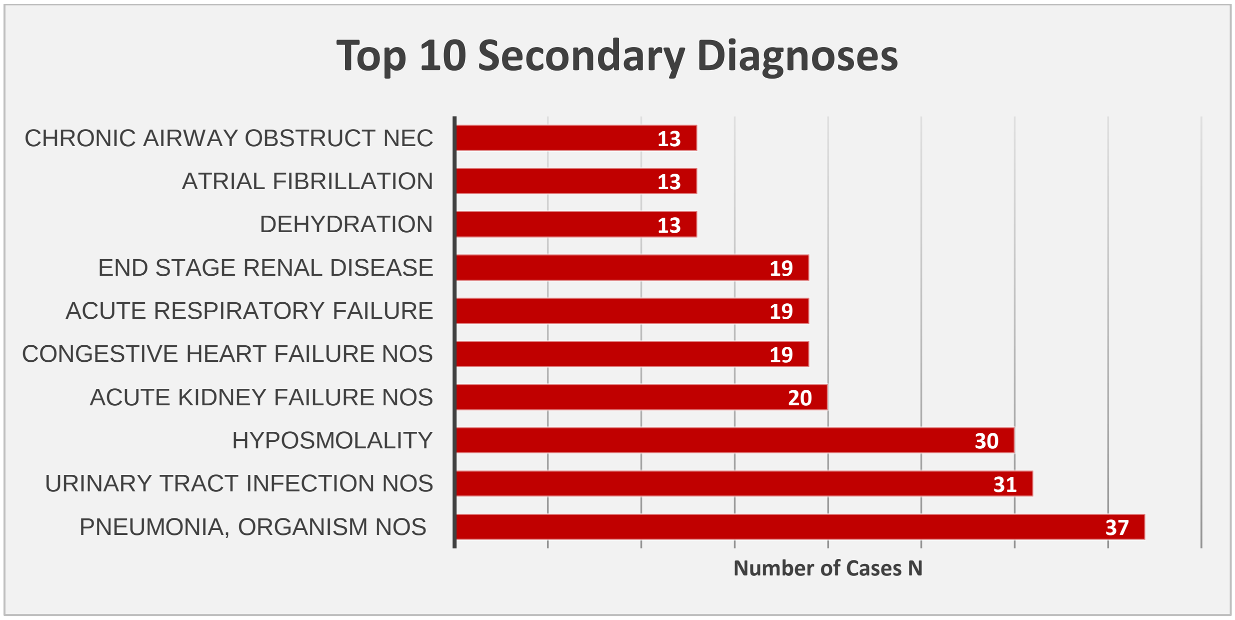


Figure 11: Top 10 Secondary Diagnoses

Furthermore, the following are the five most frequent principal procedures: implantation or replacement of intracranial neurostimulator, vaccination, cervical fusion, diagnostic ultrasound of heart, and heart countershock (table 19 and figure 12).

Table 19. Most Frequent Principal Procedures

ICD-9		Description	Number of cases (%)
1	02.93	Implantation or replacement of intracranial neurostimulator lead(s)	23 (2.28%)
2	99.55	Vaccination NEC	19 (1.88%)
3	81.03	Other cervical fusion of the posterior column, posterior technique	13 (1.29)
4	88.72	Diagnostic ultrasound of heart	9 (0.89%)
5	99.62	Heart countershock NEC	6 (0.59%)
Total			70 (6.93%)

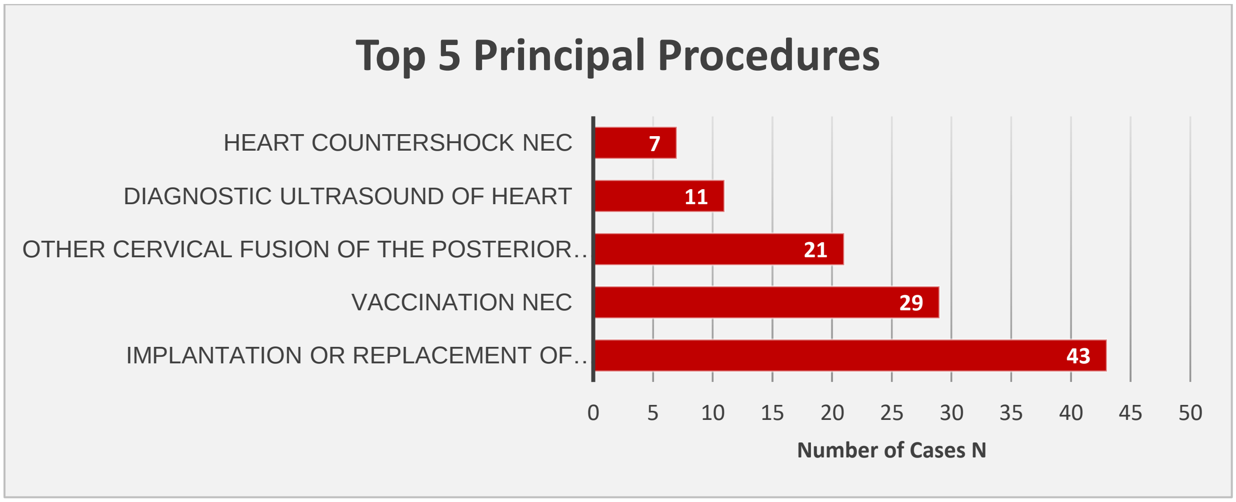


Figure 12: Top 5 principal procedures

Finally, table 20 and figure 13 represent the 10 most frequent DRGs in this sample. The most frequent DRG in this list is 0.5617 for chest pain. This DRG also represents the lowest weight in this list.

Table 20. Most Frequent DRGs

DRG	Description	Weight
0313	CHEST PAIN	0.5617
0392	ESOPHAGITIS, GASTROENT & MISC DIGEST DISORDERS W/O MCC	0.7375
0885	PSYCHOSES	0.9209
0194	SIMPLE PNEUMONIA & PLEURISY W CC	0.9996
0871	SEPTICEMIA OR SEVERE SEPSIS W/O MV 96+ HOURS W MCC	1.8803
0312	SYNCOPE & COLLAPSE	0.7339
0309	CARDIAC ARRHYTHMIA & CONDUCTION DISORDERS W CC	0.8098
0069	TRANSIENT ISCHEMIA	0.7449
0690	KIDNEY & URINARY TRACT INFECTIONS W/O MCC	0.7810
0552	MEDICAL BACK PROBLEMS W/O MCC	0.8533

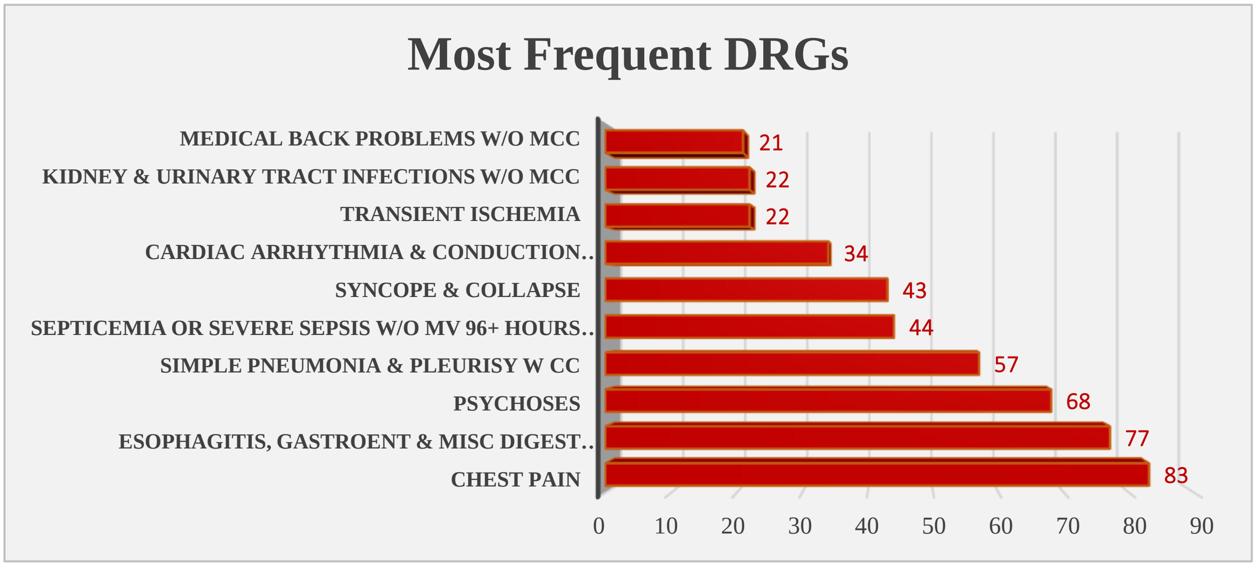


Figure 13: Most frequent DRGs

Distribution of Coding Errors Based on Number of Digits

Accuracy scores were assigned to individual codes based on the following scale:

5	All digits are captured by codes assigned
4	One digit is different between the codes assigned
3	Two digits are different between the codes assigned
2	Three digits are different between the codes assigned
1	>3 digits are different between the codes assigned
0	0= Not coded (added by reviewer)

A total of 1,058 coding errrors were found in this sample which represents 7.72% of the total codes. The accuracy rate in this sample is around 93% which is considered relatively high. The most common error was missing codes which accounts for 46.88% of the entire coding errors. Errors where one digit is different account for approximately 9% while errrors where two digits are different accounts for 7%. Table 21 and figure 14 represent the distribution of coding errors based on the different number of digits cited by the coding auditors.

Table 21. Number of different digits between coders & reviewers

Number of different digits between coders & reviewers			
1	One digit is different between the codes assigned	93	8.79%
2	Two digits are different between the codes assigned	73	6.90%
3	Three digits are different between the codes assigned	34	3.21%
5	>3 digits are different between the codes assigned	104	9.83%
6	Not coded (added by reviewer)	496	46.88%
7	Deleted by reviewer	258	24.39%
Total		1058	100.00

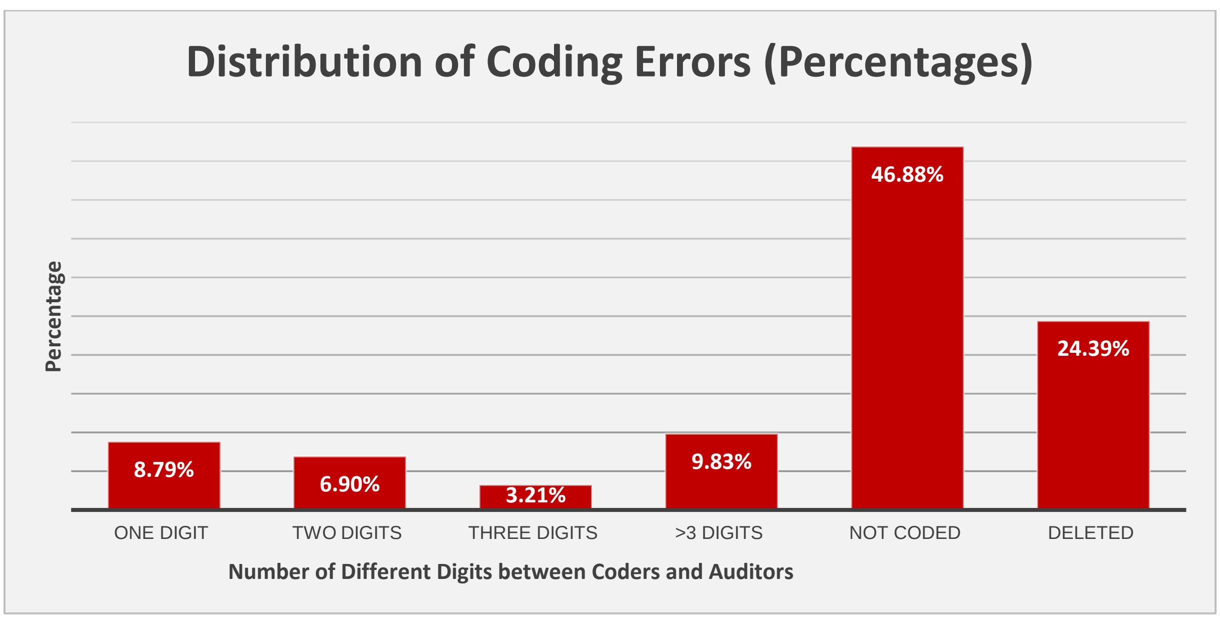


Figure 14: Distribution of Coding Errors (%)

However, when examining errors at the major vs minor digit categories, errors at the fifth-digit categories account for approximately 25% while errors at the fourth-digit and three-digit categories account for around 20% and 55%, respectively. Figure 15 illustrates percentages of errors at the third, fourth, and fifth digit levels.

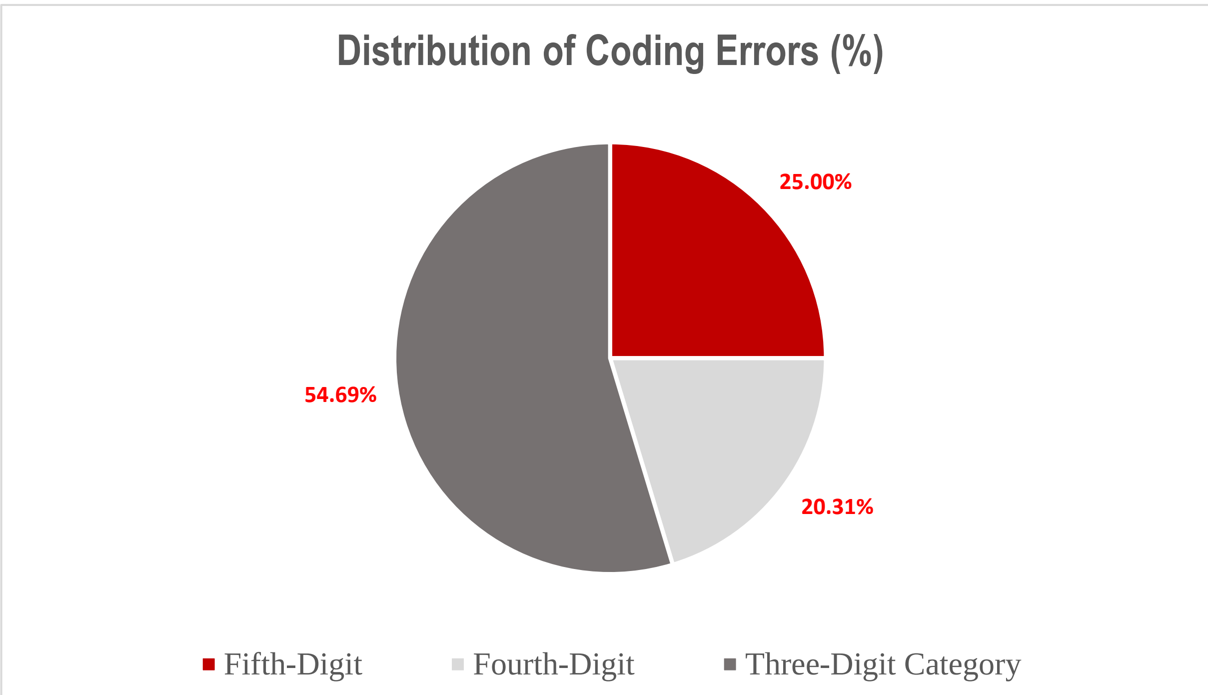


Figure 15: Distribution of Coding Errors Based on 3rd, 4th, and 5th Digits

Coding Accuracy Observations

The total number of cases analyzed in this sample is 1,010 cases with a total number of 13,713 codes. The total number of cases with errors is equal to 940 cases while the total number of cases with any type of errors is equal to 70 cases. The accuracy rate was calculated using two methods:

1. **Method I:**

$$\frac{\text{Total Number of Cases with Errors}}{\text{Total Number of Cases}} * 100$$

<u>Based on this method, accuracy rate = (70/1,010)*100= 93.07%</u>

2. **<u>Method II:</u>**

$$\frac{\text{Sum of Accuracy Scores of Individual Codes}}{\text{Sum of Highest Possible Score of Individual codes}} * 100$$

<u>When using the second method, the average accuracy rate increased to 94.98%.</u>

The total number of cases with DRG changes is 52 cases of which 18 cases were changed due to principle diagnosis. DRG change due to secondary or additional diagnoses was observed in 23 cases. Furthermore, sequencing errors lead to DRG change in 11 cases. The total payment difference is equal to $ -34,461.6 in deficit. Table 22 provides a summary of the descriptive analysis.

Accuracy Rate by DRGs

Table 23 represents the accuracy rate for the 10 most frequent DRGs: (1) CHEST PAIN; (2) ESOPHAGITIS, GASTROENT & MISC DIGEST DISORDERS W/O MCC; (3) PSYCHOSES; (4) SIMPLE PNEUMONIA & PLEURISY W CC; (5) SEPTICEMIA OR SEVERE SEPSIS W/O MV 96+ HOURS W MCC; (6) SYNCOPE & COLLAPSE; (7) CARDIAC ARRHYTHMIA & CONDUCTION DISORDERS W CC; (8) TRANSIENT ISCHEMIA; (9) KIDNEY & URINARY TRACT INFECTIONS W/O MCC; and (10) MEDICAL BACK PROBLEMS W/O MCC.

Table 22. Accuracy Statistics

1	Total number of cases analyzed	1010
2	Total number of codes	13713
3	Number of cases with no errors	940 (93.07%)
4	Number of cases with errors	70 (6.93%)
5	Number of cases with DRG change	52 (5.15%)
6	DRG change due to principle diagnosis	18 (1.78%)
7	DRG change due to secondary (additional) diagnoses	23 (2.28%)
8	DRG change due to sequencing error	11 (1.09%)
9	DRG change due to principle or secondary procedure	0
10	Number of cases with change in principle diagnosis	37 (3.66%)
11	Number of cases with change in secondary diagnosis	41 (4.06)
12	Number of cases with change in principle procedure	7 (0.69%)
13	Number of cases with change in secondary procedure	2 (0.20%)
14	Overall Payment Difference (deficit)	-34,461.6

DRG Changes based on Coding Audit

In most cases, coding audit did not result in changes in DRG assignment. Changes in principle diagnosis as well as sequencing were the most common reasons for changes in DRG assignment. Table 24 represents the most frequent DRG changes due to principle diagnosis.

Table 23. Accuracy Rate based on the Most Frequent DRGs

DRG	Description	Weight	Accuracy
0313	CHEST PAIN	0.5617	98.65
0392	ESOPHAGITIS, GASTROENT & MISC DIGEST DISORDERS W/O MCC	0.7375	99.75
0885	PSYCHOSES	0.9209	100
0194	SIMPLE PNEUMONIA & PLEURISY W CC	0.9996	100
0871	SEPTICEMIA OR SEVERE SEPSIS W/O MV 96+ HOURS W MCC	1.8803	99.98
0312	SYNCOPE & COLLAPSE	0.7339	100
0309	CARDIAC ARRHYTHMIA & CONDUCTION DISORDERS W CC	0.8098	97.99
0069	TRANSIENT ISCHEMIA	0.7449	99.09
0690	KIDNEY & URINARY TRACT INFECTIONS W/O MCC	0.7810	98.97
0552	MEDICAL BACK PROBLEMS W/O MCC	0.8533	100

Table 24. Most Frequent DRG Changes Due to Principle Diagnosis

DRG (Weight)	Payment	Description	DRG (Weight)	Payment	Description
871 (1.8803)	11657.86	SEPTICEMIA OR SEVERE SEPSIS W/O MV 96+ HOURS W MCC	291 (1.5174)	9407.88	HEART FAILURE & SHOCK W MCC
193 (1.4893)	9233.66	SIMPLE PNEUMONIA & PLEURISY W MCC	190 (1.186)	7353.2	CHRONIC OBSTRUCTIVE PULMONARY DISEASE W MCC
291 (1.5174)	9407.88	HEART FAILURE & SHOCK W MCC	689 (1.1784)	7306.08	KIDNEY & URINARY TRACT INFECTIONS W MCC
455 (5.8705)	34937.93	COMBINED ANTERIOR/POSTERIOR SPINAL FUSION W/O CC/MCC	502 (1.067)	6350.19	SOFT TISSUE PROCEDURES W/O CC/MCC
689 (1.1784)	7306.08	KIDNEY & URINARY TRACT INFECTIONS W MCC	194 (0.9996)	6197.52	SIMPLE PNEUMONIA & PLEURISY W CC
988 (1.8567)	11511.54	NON-EXTENSIVE O.R. PROC UNRELATED TO PRINCIPAL DIAGNOSIS W CC	349 (0.8075)	5006.5	ANAL & STOMAL PROCEDURES W/O CC/MCC

Table 24 (continued)					
812 (0.792)	4910.4	RED BLOOD CELL DISORDERS W/O MCC	760 (0.7892)	4893.04	MENSTRUAL & OTHER FEMALE REPRODUCTIVE SYSTEM DISORDERS W CC/MCC
872 (1.1339)	7030.18	SEPTICEMIA OR SEVERE SEPSIS W/O MV 96+ HOURS W/O MCC	690 (0.787)	4879.4	KIDNEY & URINARY TRACT INFECTIONS W/O MCC
422 (1.3006)	7740.44	HEPATOBILIARY DIAGNOSTIC PROCEDURES W/O CC/MCC	866 (0.7594)	4519.52	VIRAL ILLNESS W/O MCC
312 (0.7339)	4334.66	SYNCOPE & COLLAPSE	948 (0.701)	4140.34	SIGNS & SYMPTOMS W/O MCC

9.1.2 Bivariate Analyses

Pearson and Spearman correlation coefficients (*r*, *rho*) were calculated for the relationship between variables. Table 25 demonstrates zero-order correlations among variables.

The strongest correlation was found between DRG Weight and CMI (r (1008) = 0.259, $p < .01$), indicating that a significant linear relationship exists between both variables. This is considered as a positive and weak correlation.

Table 25. Zero-order correlations among the variables (Accuracy)

		(1)	(2)	(3)	(4)
(1)	Coding Accuracy	-			
(2)	Length of Stay (LOS)	-0.016	-		
(3)	Case Mix Index (CMI)	-0.022	-0.046	-	
(4)	DRG Relative Weight	-0.023	.183**	.259**	-
	N	1,010	1,010	1,010	

****** Correlation is significant at the 0.01 level (2-tailed)

However, it was expected to find a significant correlation in this context since CMI and DRG weight are not conceptually independent. Also, another positive but weak correlation was found between LOS and DRG weight (r (1010) = 0.183, $p < .01$), indicating a significant linear relationship between the two variables. Obviously, higher DRGs are associated with longer hospital stay.

Furthermore, there are weak and negative relationships between coding accuracy and LOS (r (1,010) = -0.016, $p = 0.610$); coding accuracy and CMI (r (1,010) = -0.022, $p = 0.485$); and coding accuracy and DRG weight (r (1,010) = -0.023, $p = 0.532$). However, all of them were not found to be statistically significant. It should be noted here that accuracy score was calculated for each case using the following formula:

$$\frac{\text{Sum of Accuracy Scores of Individual Codes}}{\text{Sum of Highest Possible Score of Individual codes}} * 100$$

9.2 IDENTIFYING DOCUMENTATION DISCREPANCIES THAT COULD INFLUENCE CODING QUALITY

Clinical documentation can have a significant influence on coding quality. Accurate and complete documentation would contribute to clinical data integrity and help coders through code assignment process without having to constantly query physicians about documentation (Combs, 2016). Therefore, documentation-related issues that could influence coding quality will be explored and identified through quantitative and qualitative data analyses. The following aspects will be discussed in this section: (1) unspecified codes rate; (2) physician query rate; (3) most frequent coding errors by ICD-10 chapters; (4) source documents most frequently used to identify coding errors; and (5) coding errors related to coding guidelines.

9.2.1 Unspecified codes rate

The unspecified codes rate is one of the most current coding metrics used to evaluate the impact of clinical documentation on coding quality (combs, 2016). Assigning unspecified codes would be coder’s last resort if no further information can be obtained from the patient chart. The unspecified codes rate is calculated by dividing the total number of unspecified codes (numerator) by the total number of codes in each sample (denominator). The total number of unspecified codes assigned in this sample (N=13,713) is 2,027 which means that the unspecified code rate is 14.78%. This can indicate high standards of clinical documentation in this sample as Combs (2015) suggested that an unspecified code rate should not exceed 20%.

9.2.2 Physician Query Rate

Physician query rate is also considered one of the new coding metrics that is used to evaluate the quality of clinical documentation (Combs, 2016). It has been suggested that the need to query physicians decreases with complete and clear clinical documentation. Physician query rate is calculated by dividing the total number of cases in which coders had to query physicians (numerator) by the total number of cases in each sample (denominator). The physician query rate for this sample (N=1010) is 19/1010= 1.88 %. Low physician query rate in this sample can indicate high clinical documentation standards in this sample (Butler, 2016).

9.2.3 Most Frequent Errors By ICD-10 Chapters

In this section, ICD-10-CM chapters most commonly related to coding errors were identified using General Equivalence Mappings (GEMs). In some cases, one-to-one mapping could not be obtained. Therefore, all potential targets were included per NLM and IHSDO mapping rules (IHSDO, 2014; NLM; 2015).

Table 26 and figure 16 show the distribution of coding errors by ICD-10 chapters with the number of errors in each chapter. *Diseases of the circulatory system* (chapter 9), *Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified* (chapter 18), *Diseases of the digestive system* (11), *Diseases of the musculoskeletal system and connective tissue* (chapter13), and *Diseases of the respiratory system* (chapter10) were identified as the top 5 ICD-10-CM chapters associated with coding errors (table 27).

Table 26. Distribution of Errors by ICD-10 Chapters

	ICD-10-CM Chapter	N	Percent	Cumulative Percent
1	Certain infectious and parasitic diseases	11	5.9	5.9
2	Neoplasms	2	1.1	7.0
3	Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism	2	1.1	8.1
4	Endocrine, nutritional and metabolic diseases	15	8.1	16.2
5	Mental, Behavioral and Neurodevelopmental disorders	6	3.2	19.5
6	Diseases of the nervous system	1	0.5	20.0
8	Diseases of the ear and mastoid process	4	2.2	22.2
9	Diseases of the circulatory system	38	20.5	42.7
10	Diseases of the respiratory system	16	8.6	51.4
11	Diseases of the digestive system	24	13.0	64.3
13	Diseases of the musculoskeletal system and connective tissue	18	9.7	74.1
14	Diseases of the genitourinary system	11	5.9	80.0
18	Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified	28	15.1	95.1
19	Injury, poisoning and certain other consequences of external causes	9	4.9	100.0
Total		*185*	*100.0*	*100.0*

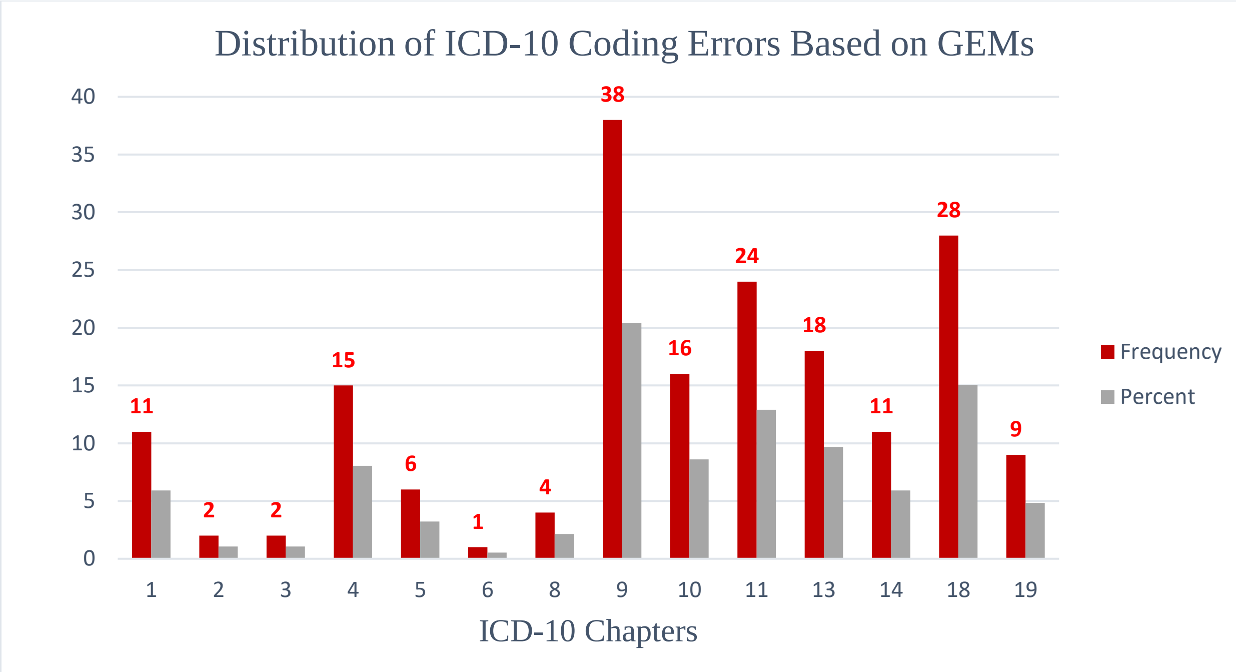


Figure 16: Distribution of ICD-10 Coding Errors Based on GEMs

Table 27. Most Frequent Errors: Top Five ICD-10 Chapters

	ICD-10 Chapter	Frequency	%
1	Diseases of the circulatory system (9)	38	20.54
2	Symptoms, signs, and abnormal clinical and laboratory findings, not elsewhere classified (18)	28	15.14
3	Diseases of the digestive system (11)	24	12.97
4	Diseases of the musculoskeletal system and connective tissue (13)	18	9.73
5	Diseases of the respiratory system (10)	16	8.65

9.2.4 Source Documents Used to Identify Coding Errors

A critical aspect of the coding audit process is providing feedback to coders in case of disagreement between coders and auditors regarding code assignment. Coding audit is conducted for two major reasons: (1) quality improvement of coded data; and (2) training and education for coders.

Therefore, auditors usually explain their reasoning behind any coding changes. They guide coders to look at certain coding guidelines or to look at specific documents that were used to change the codes initially assigned by coders.

In this section, qualitative analysis was performed on the audited data to identify documentation-related issues that influenced accuracy of coding. All cases with coding discrepancies which include auditor's comments to coders were identified. A transcript of all comments was developed and analyzed using Nvivo. Automated coding was performed to identify themes with higher relative weights. In addition, manual coding was performed to look at specific issues related to documentation and coding change. Based on this qualitative analysis, different documentation-related issues that could have a significant impact on coding accuracy were identified.

The following are the most common documents cited by auditors for coding change: (1) history & physical; (2) discharge summary; (3) progress notes; (4) consulting notes; (5) anesthesia record; (6) operative report; (7) emergency record; and (8) psychiatric evaluation. An auditor might use a single document to justify coding changes or different documents combined.

Coders were instructed to look at certain documents especially when there is a change in the principle diagnosis code. Other changes include changes in secondary diagnoses, sequencing errors, and adding more codes related to health status based on documentation. Table 28 and figure 17 show most cited documents used by auditors for coding change.

Table 28.. Source Documents used to Change the Assigned Codes (Nvivo)

Word	Length	Count	Percentage (%)
H&P	3	87	0.2175
Discharge summary	17	81	0.2025
Progress Notes	14	79	0.1975
Consultation	12	75	0.1875
Anesthesia	10	74	0.185
Operative	9	74	0.185
ED Record	9	69	0.1725
Psychiatric Evaluation	22	30	0.075

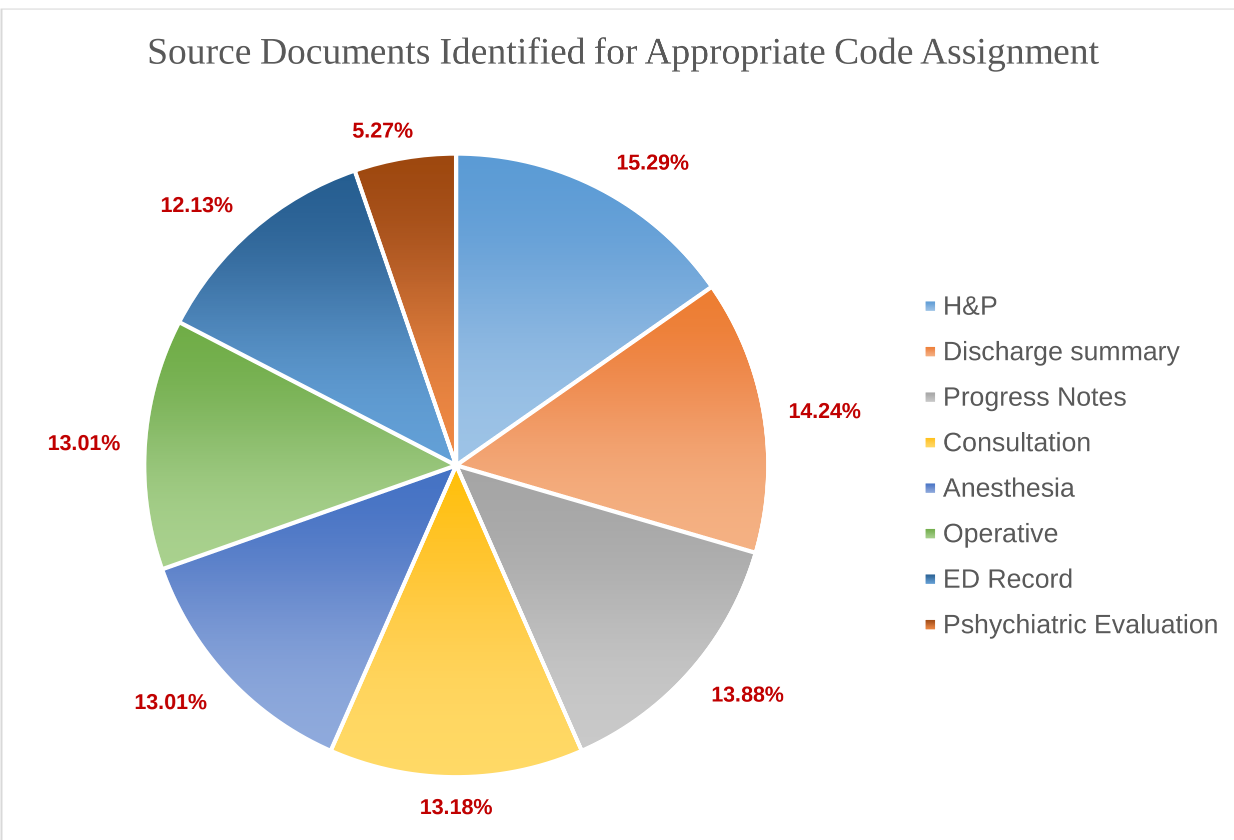


Figure 17: Source Documents Used for Coding Changes

Examples of cases where specific documents were identified to guide the coders through coding changes are provided in APPENDIX A.

9.2.5 Most Frequent Errors Related to Coding Guidelines

The following are the most frequent errors related to coding guidelines that were cited by auditors for coding change: (1) general coding guidelines; (2) symptoms & signs; (3) principle diagnosis; (4) secondary diagnoses; (5) V codes; (6) procedures; (7) combination codes; (8) MCC; (9) CC; and (10) POA. Table 29 shows most cited guidelines used for coding change. Also, figure 18 shows distribution of errors by coding guidelines.

Table 29. Most Frequent Errors Related to Coding Guidelines (Nvivo)

Word	Length	Count	Percentage (%)
Guidelines	10	53	0.19
Symptoms & Signs	16	49	0.18
Principle	9	46	0.16
Secondary	9	39	0.14
V code	6	39	0.14
Procedure	9	28	0.10
Combination code	16	17	0.06
MCC	3	13	0.05
CC	2	12	0.04
POA	3	9	0.03
Place	5	9	0.03

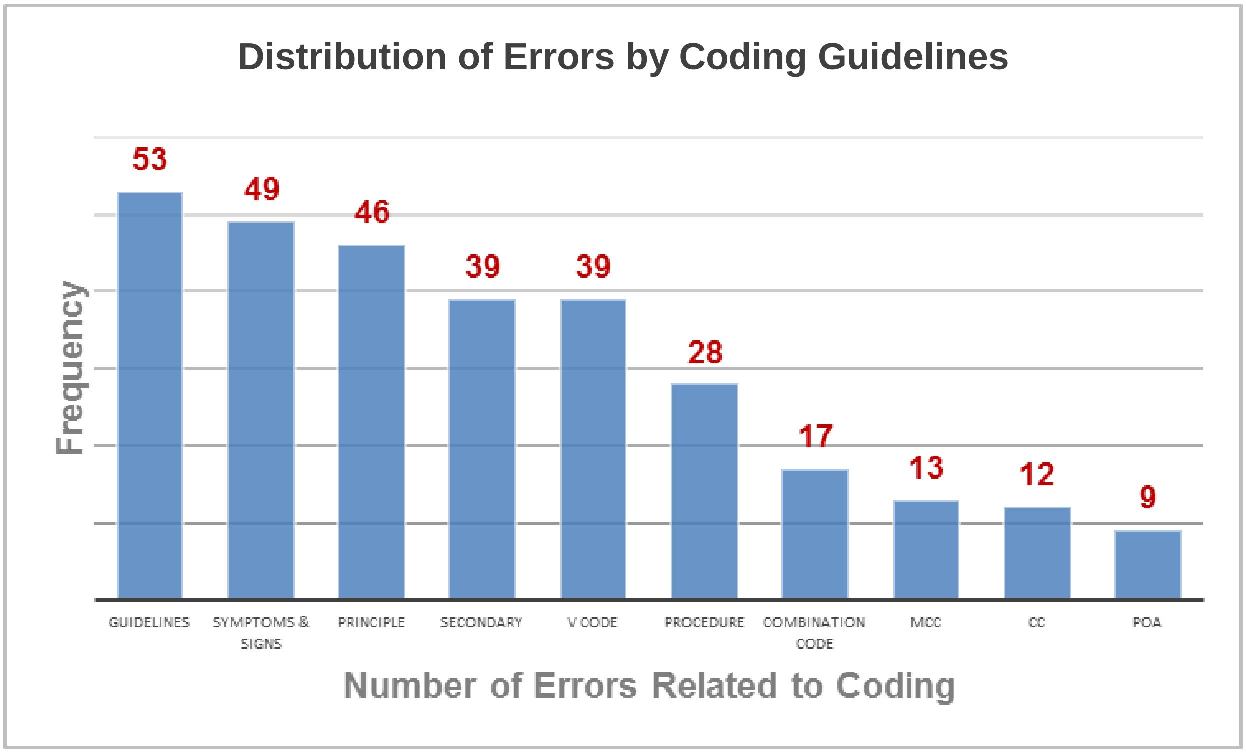


Figure 18: Distribution of Errors by Coding Guidelines

Examples of cases where relevant guidelines were identified to guide the coders through coding changes are provided in APPENDIX A. In addition, examples of recommended changes per documentation are provided in APPENDIX A.

9.3 IDENTIFYING IMPACT OF CODING ERRORS ON CMI AND HOSPITAL'S PAYMENT

Table 30 represents a list of the top 5 sequencing errors that resulted in DRG changes along with their impact on facility payment. To identify the impact of coding errors on hospital payment as well as CMI, two paired-samples *t* tests were calculated.

A paired-samples *t* test was calculated to compare the mean CMI (before audit) to the mean CMI (after audit). The mean CMI before coding audit is 1.048 (sd=.19), and the mean CMI after coding audit is 1.043 (sd=.19). No significant difference in CMI was found based on the coding audit (t (16) = .861, p= 0.19).

Also, another paired-samples *t* test was calculated to compare the mean facility payment (before audit) to the mean facility payment (after audit). The average hospital payment before coding audit is $6774.07 (sd= 4976.65), and after coding audit is $6739.17 (sd= 5207.39). No significant difference in payment was found based on the coding audit (t (16) = -.608, p=0.11).

Table 30: Most Frequent Sequencing Errors and Their Impact on Payment

#	DRG	Weight	Payment	Description	DRG	Weight	Payment	Description	ICD-9-F	Description	ICD-9-R	Description	Impact on Payment
1	0949	0.9372	$5,810.64	AFTERCARE W CC/MCC	194	0.9996	$6,197.52	SIMPLE PNEUMONIA & PLEURISY W CC	V5862	Long-term use antibiotic	486	Pneumonia, organism NOS	$386.88
									486	Pneumonia, organism NOS	V5862	Long-term use antibiotic	$0.00
2	0689	1.1784	$7,306.08	KIDNEY & URINARY TRACT INFECTIONS W MCC	194	0.9996	$6,197.52	SIMPLE PNEUMONIA & PLEURISY W CC	599.0	Urinary tract infection NOS	486	Pneumonia, organism NOS	($1,108.56)
									486	Pneumonia, organism NOS	599.0	Urinary tract infection NOS	$0.00
3	0872	1.0988	$6,489.89	SEPTICEMIA OR SEVERE SEPSIS W/O MV 96+ HOURS W/O MCC	564	1.4459	$8,539.98	OTHER MUSCULOSKELETAL SYS & CONNECTIVE TISSUE DIAGNOSES W MCC	381.2	MRSA septicemia	997.62	Infection amputation stump	$2,050.09
									997.62	Infection amputation stump	381.2	MRSA septicemia	$0.00
4	0988	1.8567	$11,511.54	NON-EXTENSIVE O.R. PROC UNRELATED TO PRINCIPAL DIAGNOSIS W CC	349	0.8075	$5,006.50	ANAL & STOMAL PROCEDURES W/O CC/MCC	250.80	Diabetes with other specified manifestations, type II or unspecified type, not stated as uncontrolled	566	Anal & rectal abscess	($6,505.04)
									566	Anal & rectal abscess	250.80	Diabetes with other specified manifestations, type II or unspecified type, not stated as uncontrolled	$0.00
5	0392	0.7241	$4,489.42	ESOPHAGITIS, GASTROENT & MISC DIGEST DISORDERS W/O MCC	641	0.6988	$4,332.56	MISC DISORDERS OF NUTRITION,METABOLISM, ,FLUIDS/ELECTROLYTES W/O MCC	558.9	Noninfectious gastroenteritis NEC	276.51	Dehydration	($156.86)
									276.51	Dehydration	558.9	Noninfectious gastroenteritis NEC	$0.00
													($5,333.49)

9.4 IDENTIFYING INDIVIDUAL AND FACILITY-RELATED FACTORS THAT COULD INFLUENCE CODING PRODUCTIVITY

9.4.1 Descriptive Analyses

Analysis of the productivity data begins with descriptive statistics which are critical to explore. Descriptive statistics include measures of central tendency (mean, trimmed mean, and median) and dispersion (standard deviation, minimum and maximum values to calculate range).

Coding Time (N= 323,112)

The average coding time is 39.46 minutes (95% CI= 39.33-39.60) with a standard deviation of ± 35.91 minutes. The coding time ranges from .10 to 593.80 minutes -minimum and maximum values respectively- which indicates that data is widely spread. Median coding time is 30.60 minutes which suggests that distribution is skewed to the right (median < mean). Furthermore, skewedness and kurtosis statistics are 3.355 and 22.979 respectively which suggests deviation from normality (skewed to the right with higher peak).

Length of Stay (LOS) (N= 323,112)

The average LOS is 4.86 days (95% CI= 4.82-4.87) with a standard deviation of ± 7.00 days. The minimum and maximum LOS are 1 and 357.00 days respectively (range= 356). Skewness (13.89) and kurtosis (405.64) suggest non-normality in LOS distribution; right-skewed & leptokurtic distribution.

Case Mix Index (CMI) (N= 323,112)

The average CMI is 1.57 (95% CI= 1.570-1.573) with a standard deviation of ± 0.36 points. The minimum and maximum CMI are .68 and 10.47 respectively (range= 9.79). Skewness (3.68) and kurtosis (35.21) suggest non-normality in LOS distribution; right-skewed & leptokurtic distribution.

Relative DRG Weight (*N*= 323,112)

The average relative DRG Weight is 1.59 ± 1.53 (95% CI= 1.58- 1.59). DRG relative weight ranges from .16 to 26.25- minimum and maximum values. The median and mode weight are 1.14 and .59 respectively. Skewness (.665) and kurtosis (56.72) suggest non-normality in LOS distribution; right-skewed & leptokurtic distribution.

Bed Size (*N*= 323,112)

The average bed size is 509 beds with a standard deviation of 296 beds (95% CI= 508-510). Hospital beds in the sample ranges from 25 to 1346 beds which represent minimum and maximum values, respectively. Skewness (5.61) and kurtosis (-.484) suggest distribution is slightly skewed to the right with a flat-topped curve (platykurtic distribution).

Inpatient Coding Productivity Observations

A total of 323,112 cases were analyzed based on data provided from 119 facilities. These facilities are in 25 different states across the US including: AR, CA, CT, DE, FL, GA, ID, IL, IN, KS, LA, MA, MI, NC, NJ, NM, NY, OR, PA, SC, TN, TX, VA, WA, and WI. Thirty-seven percent of participating facilities were non-trauma centers and 51 percent of these facilities are designated trauma centers (Level-1= 29%; Level-2= 18%, and Level-3= 4%). No information was provided regarding trauma status of the remaining 12% of the facilities. Sixty-three percent of the facilities hold teaching status while only 25 percent are non-teaching facilities. No information was provided regarding teaching status of the remaining 12% of the facilities.

In addition, participating facilities have a wide range of bed capacity ranges from 25 to 1,346 beds. The average bed count is 509 beds with standard deviation of ± 296 beds. Finally, average CMI is 1.57 with a standard deviation of .36 units and ranges between .68 and 10.47.

The mean coding time of all facilities is 39.46 minutes with standard deviation of ±35.91 minutes. When excluding the highest and lowest 5% of the data, mean coding time decreases by 4 minutes (trimmed mean). Average length of stay (LOS) is 5 days with standard deviation of 7 days. Lowest and highest length of stay are 1 and 357 days, respectively. Also, average DRG relative weight of all cases is 1.59 with standard deviation of ±1.53 and it ranges between .16 and 26.25. Table 31 presents distribution of coding time, LOS, CMI, DRG Weight, and Bed Size.

Table 31. Descriptive Statistics of Productivity Data

	Mean	Trimmed Mean	Standard Deviation	Minimum	Maximum	N
CODING TIME	39.46	35.53	35.91	0.10	593.80	323,112
LENGHT OF STAY (LOS)	4.86	4.04	7.00	1	357.00	323,112
CASE MIX INDEX (CMI)	1.57	1.55	0.36	0.68	10.47	323,112
DRG WEIGHT	1.59	1.39	1.53	0.16	26.25	323,112
BED SIZE	509.00	495.01	296.00	25.00	1346.00	323,112

Coding Productivity Over Time

Productivity continues to improve steadily over time. Coding productivity has improved consistently over the 10-month period (N= 323,112). The average coding time in October (2015) is 43.68 minutes compared to 37.45 minutes in July (2016). Figure 19 represents the average coding time for the 10-month period (October 2015-July 2016).

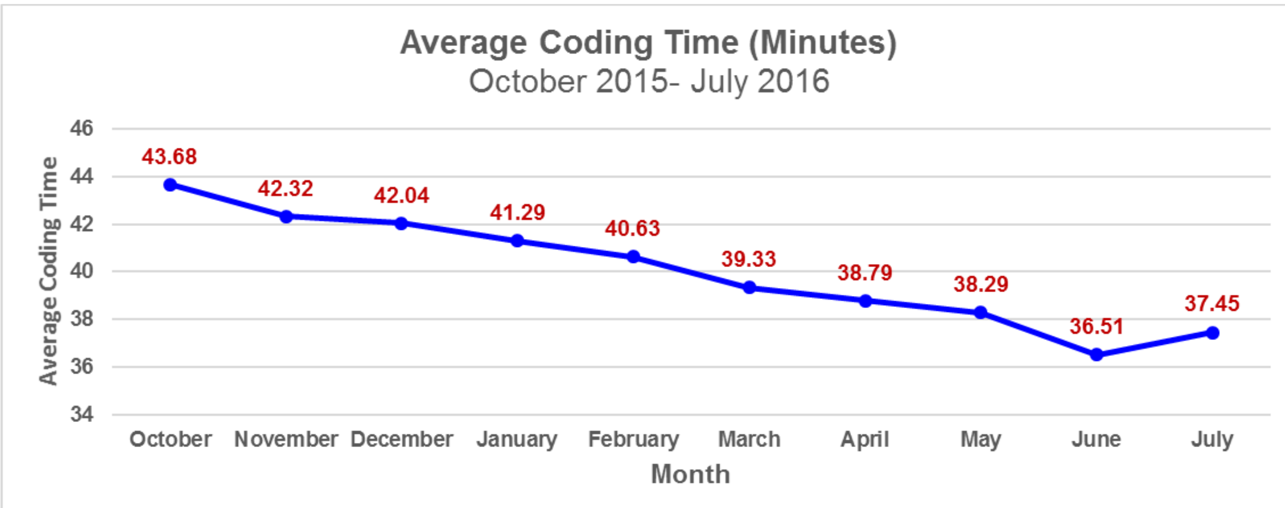


Figure 19: Average Coding Time for 10-month period (October 2015-July 2016)

Comparing to Standard Coding Times

Coding productivity with ICD-10 has decreased by nearly 22 percent when compared to ICD-9 productivity for the first five months of the new code set's use (October 2015 to February 2016). When compared to ICD-9, however, coding productivity has only decreased by 11 percent for the next five months (March 2016 to July 2016). In fact, ICD-10 coding productivity has consistently improved over time in terms of the number of coded records and average coding time. Further details are provided in figure 20.

Inpatient Coding	ICD-9 Sample Standard AHIMA	ICD-9 May 2015 – Sept. 2015 Baseline Study CIOX Health N = 84,627 cases	ICD-10 Oct. 2015 – Feb. 2016 Sample 1 CIOX Health N = 157,248 cases	ICD-10 March 2016 – July 2016 Sample 2 CIOX Health N = 165,864 cases
Records/day	24	14	11.5	12.6
Records/hour	3	1.8	1.4	1.6
Minutes/Record	20 minutes	34.2 minutes	41.9 minutes	38.1 minutes
Productivity Impact *(Based on average time per record)*			-22% (compared to CIOX Baseline ICD-9 Study	-11% (compared to the CIOX Baseline ICD-9 Study)

Figure 20: Comparing Standard Coding Productivity Times

Coding Time by LOS

Approximately 80 percent of the cases in the March 2016 to July 2016 data set had a LOS of six days or less. The lowest coding time was 27.8 minutes (LOS = one to two days) while the highest coding time was 79.5 minutes (LOS > 10 days). Distribution of coding time by six LOS categories is provided in figure 21. One can see that as LOS increases, coding productivity times increase as well, which is to be expected.

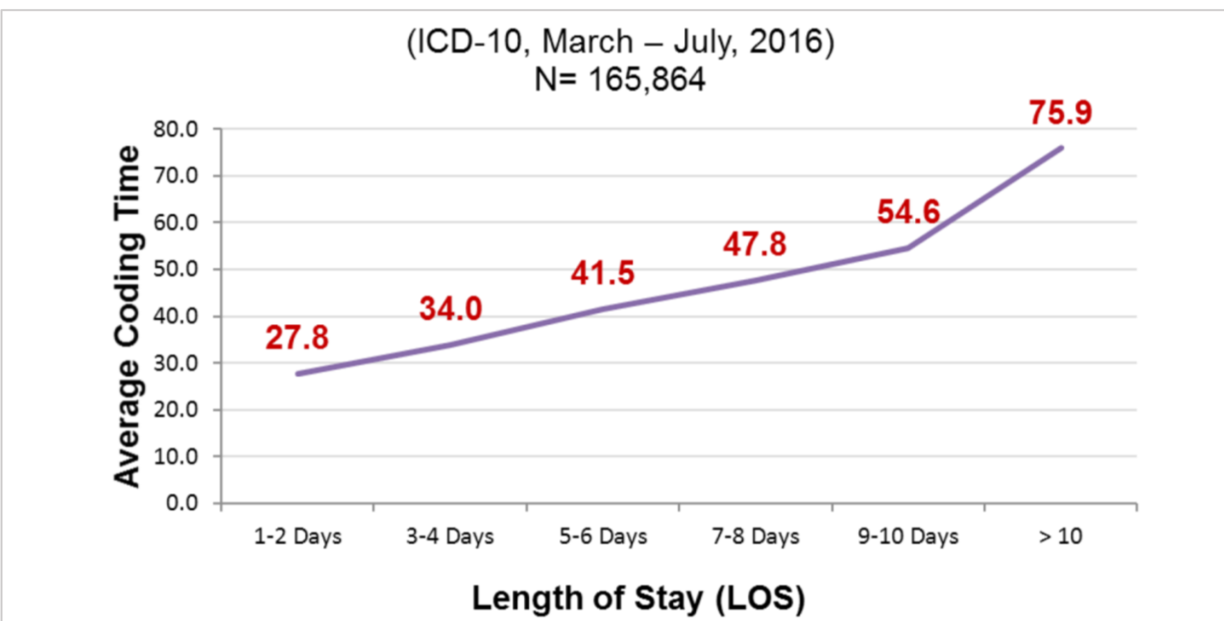


Figure 21: Average Coding Time by LOS

Coding Productivity by CMI

Also, as expected, the average coding time increased as the CMI increased. See figure 22 for a chart graphing this with cases studied between March 2016 and July 2016.

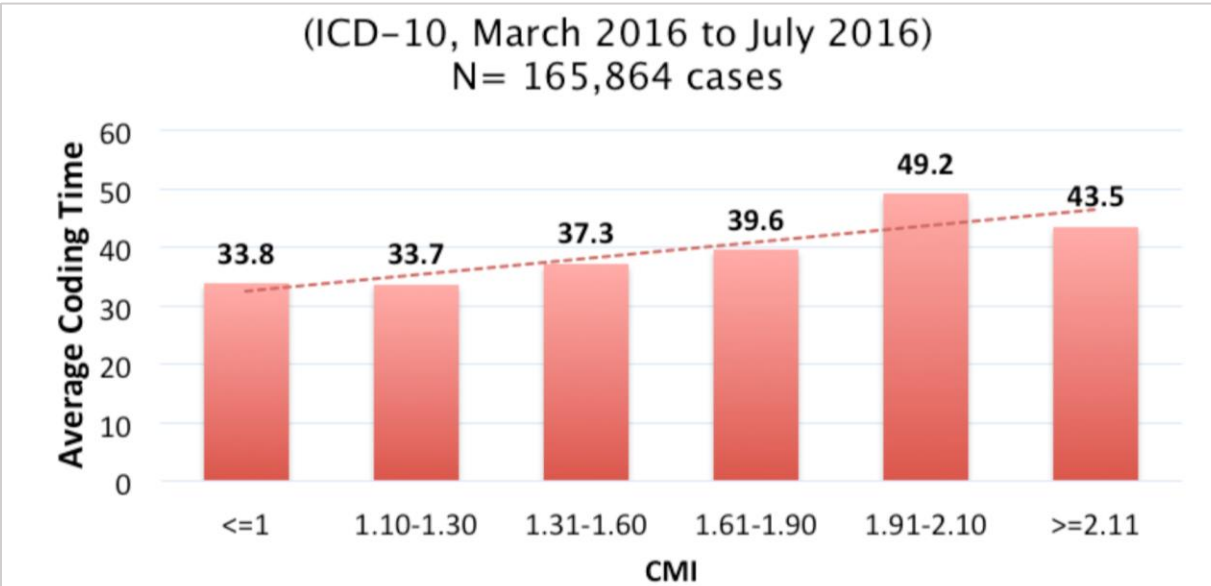


Figure 22: Average Coding Time by CMI

CMI and Coding Time by Months

It was observed that productivity gains did not come at the cost of CMI. This is a very important observation since CMI is a key indicator or metric in organizations. Otherwise stated, as productivity time continued to improve, CMI was observed to increase (see figure 23).

Top DRGs by Month

Normal newborn (DRG weight = 1.649) was the highest DRG for three consecutive months (April 2016 to June 2016) while septicemia (DRG weight = 1.7926) and vaginal delivery (DRG weight = 0.5865) were the top DRGs for the months of March 2016 and July 2016 respectively. The three highest DRGs for each month and their sample of cases are provided in the table in figure 24.

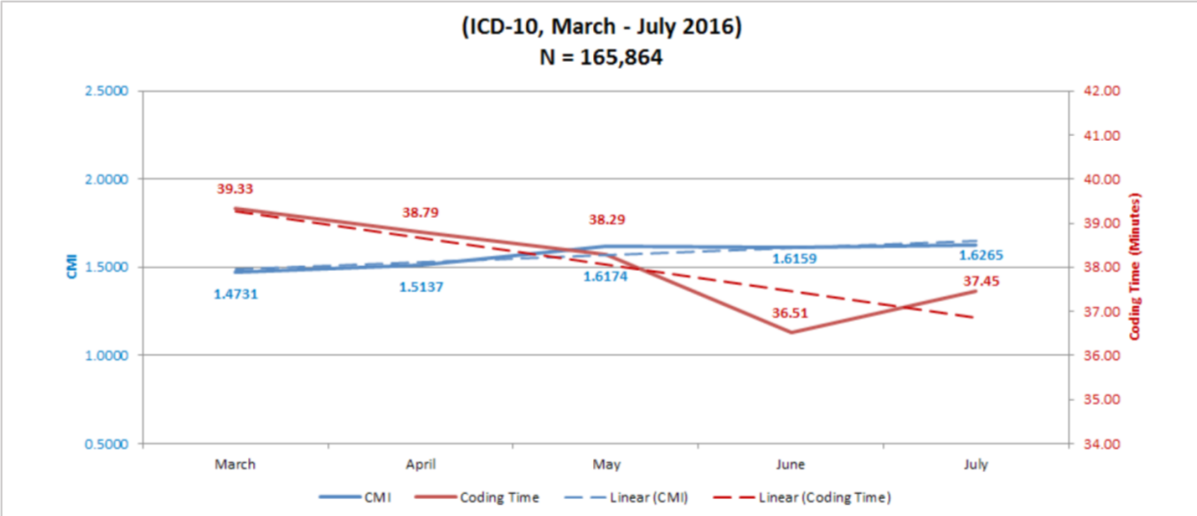


Figure 23: CMI and Coding Time by Month

2016 MONTH	DRG WEIGHT	DRG	N
March	1.7926	SEPTICEMIA OR SEVERE SEPSIS W/O MV 96+ HOURS W MCC	1453
	0.5865	VAGINAL DELIVERY W/O COMPLICATING DIAGONSES	1365
	0.1649	NORMAL NEWBORN	1223
April	0.1649	NORMAL NEWBORN	1443
	0.5865	VAGINAL DELIVERY W/O COMPLICATING DIAGONSES	1337
	1.7926	SEPTICEMIA OR SEVERE SEPSIS W/O MV 96+ HOURS W MCC	1257
May	0.1649	NORMAL NEWBORN	1416
	0.5865	VAGINAL DELIVERY W/O COMPLICATING DIAGONSES	1415
	1.7926	SEPTICEMIA OR SEVERE SEPSIS W/O MV 96+ HOURS W MCC	1294
June	0.1649	NORMAL NEWBORN	1314
	0.5865	VAGINAL DELIVERY W/O COMPLICATING DIAGONSES	1300
	1.7926	SEPTICEMIA OR SEVERE SEPSIS W/O MV 96+ HOURS W MCC	1195
July	0.5865	VAGINAL DELIVERY W/O COMPLICATING DIAGONSES	1315
	0.1649	NORMAL NEWBORN	1204
	1.7926	SEPTICEMIA OR SEVERE SEPSIS W/O MV 96+ HOURS W MCC	1061

Figure 24: Top Three DRGs by Month (March - July 2016)

Coding Time by Teaching and Trauma Status

The mean coding time for teaching facilities is 41. 85 minutes compared to non-teaching facilities in which average coding time is 34.49 minutes. Coding time also varies by Trauma status. Level-III trauma facilities have the lowest mean coding time (36 minutes) compared to level-I trauma facilities (45 minutes). In general, coding time decreases in facilities with higher trauma status. Table 32 demonstrates average coding time by teaching and trauma status.

Table 32. Coding time by teaching and trauma status.

		Mean	95.0% Lower CL for Mean	95.0% Upper CL for Mean	Standard Deviation	Minimum	Maximum
Teaching Status	Non-Teaching	34.49	34.24	34.75	32.820	0.1	538.21
	Teaching	41.85	41.67	42.03	38.20	0.6	593.82
Trauma Status	Non-Trauma	36.06	35.85	36.27	33.83	0.2	565.23
	Trauma Level-1	44.96	44.68	45.25	41.00	0.8	593.88
	Trauma Level-2	40.72	40.40	41.046	36.81	0.7	523.56
	Trauma Level-3	36.17	35.47	36.88	30.47	0.4	440.40

Coding Time by Bed Size

Facilities with a bed count of 900-999 beds have the highest mean coding time (58 ± 53 minutes) compared to facilities with a bed capacity between 100-199 in which coding time is lowest (33 ± 33 minutes). Table 33 and figure 25 demonstrate average coding time for facilities with different bed capacity (bed size).

Table 33. Coding time by bed size

Bed Size	Mean	95.0% Lower CL for Mean	95.0% Upper CL for Mean	Standard Deviation	Minimum	Maximum
<100	40.43367	39.29568	41.57166	36.19239	0.3	412.2
100-199	33.04041	32.68649	33.39433	33.17787	0.2	565.2
200-299	35.3099	35.04987	35.56994	31.91482	0.2	538.2
300-399	38.14078	37.7514	38.53016	31.48119	0.1	475.4
400-499	45.39972	44.95493	45.84451	39.60867	0.1	523.5
500-599	40.87163	40.56113	41.18213	33.79512	0.3	482.3
600-699	35.48513	34.93426	36.03601	32.84922	0.4	431
700-799	38.0379	37.53323	38.54257	34.13892	0.1	548.8
800-899	39.91138	39.34133	40.48144	39.09396	0.5	515
900-999	58.26093	57.41923	59.10264	53.00428	0.4	593.8
1000-1099	52.89986	50.02278	55.77693	38.63114	0.9	277.1
1100-1199	42.18533	41.67302	42.69764	34.31209	0.7	475.3
>= 1200	38.29066	36.83296	39.74837	41.79258	0.1	522.4

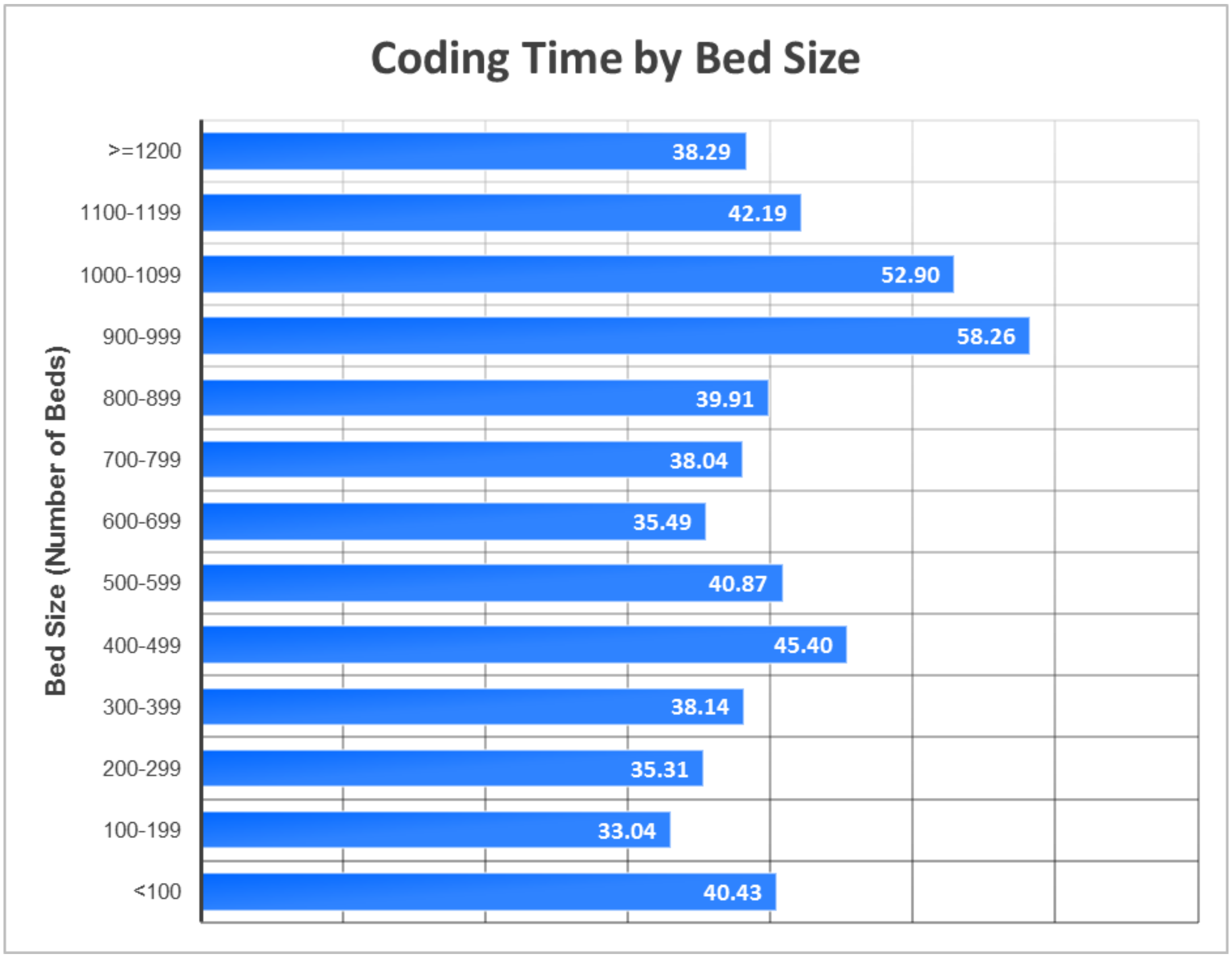


Figure 25: Coding time by bed size

Average Coding Time by MDCs

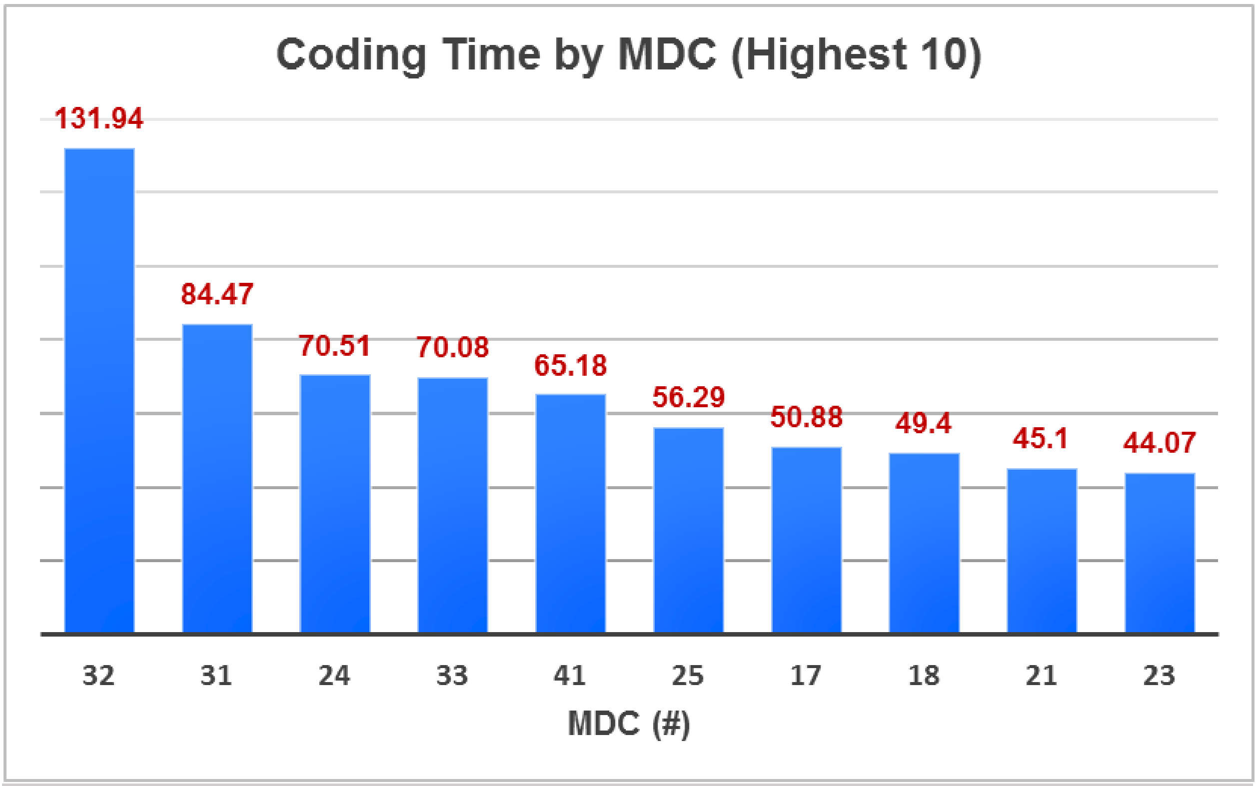
Mean coding time varies across specialties. Lowest and highest coding times belong to MDC #15 NEWBORNS & OTHER NEONATES and MDC#31 TRACHEOSTOMY with average coding times of 20.14± 24.98 minutes (95% CI= 19.80-20.48) and 131.94±93.91 minutes (95% CI= 126.41-137.48), respectively. Below is the average coding time for each MDC along with 95% CI intervals, standard deviation, as well as minimum and maximum values. Table 34 and figure 26 represent average coding time for each MDC.

Table 34. Average Coding Time by MDC

MDC		Coding Time	95.0% Lower CL for Mean	95.0% Upper CL for Mean	Standard Deviation	Max	Min
1	NERVOUS SYSTEM	42.37	41.89	42.86	34.74	468.60	0.10
2	EYE	41.57	37.80	45.34	42.13	415.10	0.30
3	EAR, NOSE, MOUTH & THROAT	38.68	37.36	40.00	34.37	593.80	0.40
4	RESPIRATORY SYSTEM	39.92	39.52	40.33	34.56	523.50	0.10
5	CIRCULATORY SYSTEM	43.64	43.27	44.00	36.52	516.80	0.10
6	DIGESTIVE SYSTEM	42.45	42.03	42.86	34.97	518.10	0.30
7	HEPATOBILIARY SYSTEM & PANCREAS	43.37	42.61	44.14	38.42	547.60	0.10
8	MUSCULOSKELETAL SYSTEM & CONN TISSUE	39.86	39.48	40.25	33.97	538.20	0.10
9	SKIN, SUBCUTANEOUS TISSUE & BREAST	38.39	37.64	39.13	31.80	435.50	0.30
10	ENDOCRINE, NUTRITIONAL & METABOLIC DISEASES & DISORDERS	36.81	36.18	37.44	32.65	506.80	0.30
11	KIDNEY & URINARY TRACT	41.77	41.23	42.30	33.91	520.10	0.10

Table 34 (continued)							
12	MALE REPRODUCTIVE SYSTEM	35.74	34.13	37.35	30.43	474.30	0.40
13	FEMALE REPRODUCTIVE SYSTEM	37.93	36.85	39.01	30.74	300.10	0.40
14	PREGNANCY, CHILDBIRTH & THE PUERPERIUM	25.83	25.56	26.10	21.46	419.50	0.10
15	NEWBORNS & OTHER NEONATES	20.14	19.80	20.48	24.98	496.50	0.10
16	BLOOD, BLOOD FORMING ORGANS, IMMUNOLOG DISORD	40.34	39.24	41.44	34.16	448.80	0.30
17	MYELOPROLIFERATIVE DISEASES & DISORDERS, POORLY DIFFERENTIATED NEOPLASM	50.88	49.22	52.54	42.25	394.90	0.10
18	INFECTIOUS & PARASITIC DISEASES, SYSTEMIC OR UNSPECIFIED SITES	49.40	48.86	49.95	40.68	515.00	0.10
19	MENTAL DISEASES & DISORDERS	25.30	24.78	25.82	23.22	394.00	0.20

Table 34 (continued)							
20	ALCOHOL/DRUG USE & ALCOHOL/DRUG INDUCED ORGANIC MENTAL DISORDERS	32.81	31.69	33.92	30.65	463.50	0.30
21	INJURIES, POISONINGS & TOXIC EFFECTS OF DRUGS	45.10	44.11	46.09	37.65	468.50	0.10
22	BURNS	42.59	35.53	49.66	33.55	181.20	1.00
23	FACTORS INFLUENCING HLTH STAT & OTHR CONTACTS WITH HLTH SERVCS	44.07	42.27	45.86	38.60	481.50	0.50
24	MULTIPLE SIGNIFICANT TRAUMA	70.51	67.20	73.83	50.89	391.30	0.40
25	HUMAN IMMUNODEFICIENCY VIRUS INFECTIONS	56.29	52.03	60.56	43.28	359.70	0.40
31	MAJOR TRANSPLANT	84.47	75.46	93.49	86.26	545.10	0.60
32	TRACHEOSTOMY	131.94	126.41	137.48	93.91	565.20	0.60
33	AUTOLOGOUS BONE MARROW TRANSPLANT	70.08	61.96	78.21	47.17	330.00	0.20
41	UNRELATED OPERATING ROOM PROCEDURE	65.18	63.23	67.13	50.10	547.10	0.30



32	TRACHEOSTOMY
31	MAJOR TRANSPLANT
24	MULTIPLE SIGNIFICANT TRAUMA
33	AUTOLOGOUS BONE MARROW TRANSPLANT
41	UNRELATED OPERATING ROOM PROCEDURE
25	HUMAN IMMUNODEFICIENCY VIRUS INFECTIONS
17	MYELOPROLIFERATIVE DISEASES & DISORDERS, POORLY DIFFERENTIATED NEOPLASM
18	INFECTIOUS & PARASITIC DISEASES, SYSTEMIC OR UNSPECIFIED SITES
21	INJURIES, POISONINGS & TOXIC EFFECTS OF DRUGS
23	FACTORS INFLUENCING HLTH STAT & OTHR CONTACTS WITH HLTH SERVCS

Figure 26: Coding Time by MDCs

Average Coding Time by DRGs

DRG# 3 (ECMO OR TRACH W MV 96+ HRS OR PDX EXC FACE, MOUTH & NECK W MAJ O.R) was associated with the highest coding time compared to all DRGs included in the data sets followed by DRG#7 (LUNG TRANSPLANT) with average coding times of 154.83 and 129.45 minutes, respectively. Table 35 represents the average coding time by DRGs (top 50).

Table 35. Average Coding Time by DRGs

DRG	DRG Description	Mean	95.0% Lower CL for Mean	95.0% Upper CL for Mean	Standard Deviation	Min	Max
3	ECMO OR TRACH W MV 96+ HRS OR PDX EXC FACE, MOUTH & NECK W MAJ O.R.	154.83	146.15	163.52	104.76	0.80	565.20
7	LUNG TRANSPLANT	129.45	68.93	189.97	90.08	17.20	330.70
4	TRACH W MV 96+ HRS OR PDX EXC FACE, MOUTH & NECK W/O MAJ O.R.	115.46	108.00	122.92	75.41	1.20	543.30
453	COMBINED ANTERIOR/POSTERIOR SPINAL FUSION W MCC	114.28	80.33	148.24	121.97	1.80	503.60
5	LIVER TRANSPLANT W MCC OR INTESTINAL TRANSPLANT	114.13	90.34	137.93	88.02	3.70	334.00

Table 35 (continued)							
11	TRACHEOSTOMY FOR FACE, MOUTH & NECK DIAGNSES W MCC	106.90	89.75	124.05	68.11	7.50	411.90
14	ALLOGENEIC BONE MARROW TRANSPLANT	106.17	90.43	121.91	76.84	2.10	489.60
456	SPINAL FUS EXC CERV W SPINAL CURV/MALIG/INFEC OR 9+ FUS W MCC	103.24	84.43	122.06	68.93	2.00	435.70
957	OTHER O.R. PROCEDURES FOR MULTIPLE SIGNIFICANT TRAUMA W MCC	102.39	91.84	112.94	64.73	1.00	391.30
984	PROSTATIC O.R. PROCEDURE UNRELATED TO PRINCIPAL DIAGNSIS W MCC	99.78	65.46	134.11	32.71	63.20	147.60
28	SPINAL PROCEDURES W MCC	99.76	74.85	124.67	86.71	3.80	446.30

Table 35 (continued)							
955	CRANIOTOMY FOR MULTIPLE SIGNIFICANT TRAUMA	99.09	80.89	117.29	42.10	28.90	195.60
20	INTRACRANIAL VASCULAR PROCEDURES W PDX HEMORRHAGE W MCC	98.57	82.98	114.16	69.14	1.00	385.70
228	OTHER CARDIOTHORACIC PROCEDURES W MCC	95.77	76.13	115.40	71.94	1.00	323.50
901	WOUND DEBRIDEMENTS FOR INJURIES W MCC	94.77	69.19	120.35	68.51	19.10	279.70
405	PANCREAS, LIVER & SHUNT PROCEDURES W MCC	94.07	83.01	105.13	67.84	2.00	507.30
826	MYELOPROLIF DISORD OR POORLY DIFF NEOPL W MAJ O.R. PROC W MCC	93.96	68.37	119.54	59.16	18.70	290.40
834	ACUTE LEUKEMIA W/O MAJOR O.R. PROCEDURE W MCC	92.63	81.03	104.24	66.88	0.60	394.90

Table 35 (continued)							
576	SKIN GRAFT &/OR DEBRID EXC FOR SKIN ULCER OR CELLULITIS W MCC	92.11	53.00	131.22	76.06	1.00	358.10
939	O.R. PROC W DIAGNSES OF OTHER CONTACT W HEALTH SERVICES W MCC	91.59	63.45	119.72	81.90	1.90	381.10
969	HIV W EXTENSIVE O.R. PROCEDURE W MCC	89.54	56.24	122.84	60.14	14.90	214.80
216	CARDIAC VALVE & OTH MAJ CARDIOTHORACIC PROC W CARD CATH W MCC	88.21	77.23	99.18	65.68	0.50	435.60
907	OTHER O.R. PROCEDURES FOR INJURIES W MCC	87.05	77.86	96.25	68.40	0.50	420.90
12	TRACHEOSTOMY FOR FACE, MOUTH & NECK DIAGNSES W CC	86.90	66.65	107.15	74.20	0.60	419.60

Table 35 (continued)							
616	AMPUTAT OF LOWER LIMB FOR ENDOCRINE, NUTRIT, & METABOL DIS W MCC	86.29	66.11	106.48	64.77	1.70	431.00
823	LYMPHOMA & NN-ACUTE LEUKEMIA W OTHER O.R. PROC W MCC	85.54	68.34	102.74	62.41	0.50	381.30
856	POSTOPERATIVE OR POST-TRAUMATIC INFECTIONS W O.R. PROC W MCC	84.53	76.68	92.39	59.27	0.60	405.10
239	POSTOPERATIVE OR POST-TRAUMATIC INFECTIONS W O.R. PROC W MCC	83.51	73.82	93.21	59.88	0.70	389.90
295	DEEP VEIN THROMBOPHLEBITIS W/O CC/MCC	83.50	-97.66	264.66	72.93	24.40	165.00
463	WND DEBRID & SKN GRFT EXC HAND, FOR MUSCULO-CONN TISS DIS W MCC	82.71	72.42	93.00	63.36	1.30	423.30

Table 35 (continued)							
219	CARDIAC VALVE & OTH MAJ CARDIOTHORACIC PROC W/O CARD CATH W MCC	81.68	74.81	88.55	59.22	0.70	424.40
326	STOMACH, ESOPHAGEAL & DUODENAL PROC W MCC	81.64	75.78	87.51	62.49	0.30	518.10
23	CRANIO W MAJOR DEV IMPL/ACUTE COMPLEX CNS PDX W MCC OR CHEMO IMPLANT	81.48	74.88	88.07	51.99	0.70	355.20
981	EXTENSIVE O.R. PROCEDURE UNRELATED TO PRINCIPAL DIAGNSIS W MCC	81.17	77.44	84.91	58.68	0.60	547.10
423	OTHER HEPATOBILIARY OR PANCREAS O.R. PROCEDURES W MCC	80.71	61.53	99.88	39.78	2.30	148.90

Table 35 (continued)							
659	KIDNEY & URETER PROCEDURES FOR NN-NEOPLASM W MCC	80.27	71.80	88.74	58.84	1.60	384.30
215	OTHER HEART ASSIST SYSTEM IMPLANT	80.20	13.92	146.48	53.38	1.00	141.70
853	INFECTIOUS & PARASITIC DISEASES W O.R. PROCEDURE W MCC	79.82	77.35	82.29	57.09	0.40	514.00
231	CORONARY BYPASS W PTCA W MCC	79.65	63.95	95.35	44.28	3.50	166.50
163	MAJOR CHEST PROCEDURES W MCC	79.19	71.98	86.41	62.34	0.90	478.70
260	CARDIAC PACEMAKER REVISION EXCEPT DEVICE REPLACEMENT W MCC	78.72	64.36	93.09	47.26	2.80	219.90
870	SEPTICEMIA OR SEVERE SEPSIS W MV 96+ HOURS	78.53	74.75	82.30	51.13	0.50	410.20
503	FOOT PROCEDURES W MCC	78.39	65.37	91.41	32.91	17.40	147.60

Table 35 (continued)							
237	MAJOR CARDIOVASC PROCEDURES W MCC OR THORACIC AORTIC ANEURYSM REPAIR	77.85	45.14	110.56	59.07	1.20	209.10
408	BILIARY TRACT PROC EXCEPT ONLY CHOLECYST W OR W/O C.D.E. W MCC	77.53	68.17	86.90	60.73	0.00	547.60
800	SPLENECTOMY W CC	77.50	45.02	109.98	58.66	18.00	226.00
40	PERIPH/CRANIAL NERVE & OTHER NERV SYST PROC W MCC	77.43	68.10	86.75	48.87	1.10	240.90
21	INTRACRANIAL VASCULAR PROCEDURES W PDX HEMORRHAGE W CC	77.14	55.32	98.96	62.54	3.20	274.60
270	OTHER MAJOR CARDIOVASCULAR PROCEDURES W MCC	76.84	71.54	82.13	50.51	0.90	334.70

9.4.2 Bivariate Analyses

Correlation Coefficients

Pearson and Spearman correlation coefficients (*r*, *rho*) were calculated for the relationship between variables. Table 36 demonstrates zero-order correlations among the variables in the productivity data set.

Table 36. Zero-order correlations among the variables (productivity data)

		(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	Coding Time	-						
(2)	Length of Stay	0.330**	-					
(3)	Case Mix Index	.119**	.091**	-				
(4)	DRG Weight	.323**	.334**	.134**	-			
(5)	Bed Size	.096**	.028**	.080**	.016**	-		
(6)	Teaching Status	.088**	.035**	.202**	.048**	.420**	-	
(7)	Trauma Status	.050**	.015**	.114**	.013**	.241**	.260**	-
	N	323,112	323,112	323,112	323,112	323,112	323,112	323,112

****** Correlation is significant at the 0.01 level (2-tailed)

Table 36 presents correlations between the different variables. In this context, however, we are more interested in examining how different predictors affect coding productivity. First, the strongest correlation was found between bed size and teaching status (r (323,110) = 0.420, $p < .01$), indicating that a significant linear relationship between teaching status and bed size. Facilities holding teaching status tend to have greater bed capacity. Also, a moderate positive correlation was found between coding time and length of stay (r (323,110) = 0.330, $p < .01$), indicating a significant linear relationship between both variables. Coders need more time to code patients charts with longer hospital stays. Also, other moderate positive correlations were found between DRG relative weight and coding time (r (323,110) = 0.323, $p < .01$) as well as between DRG relative weight and CMI (r (323,110) = 0.334, $p < .01$). Cases with higher DRGs require additional coding time and are attributed with higher CMIs. Furthermore, many weak but significant correlations were found between different variables. These relationships include: teaching status and CMI (r (323,110) = .202, $p < .01$); trauma status and bed size (r (323,110) = .241, $p < .01$).

Although we did not find any strong correlations between the variables, they can still be included in our model as predictors of coding productivity since they are statistically significant.

Scatterplots

In addition to correlation coefficients, scatter diagrams were created to visually depict the correlations between coding time, LOS, DRG weight, bed size, and CMI (Figure 27- Figure 31).

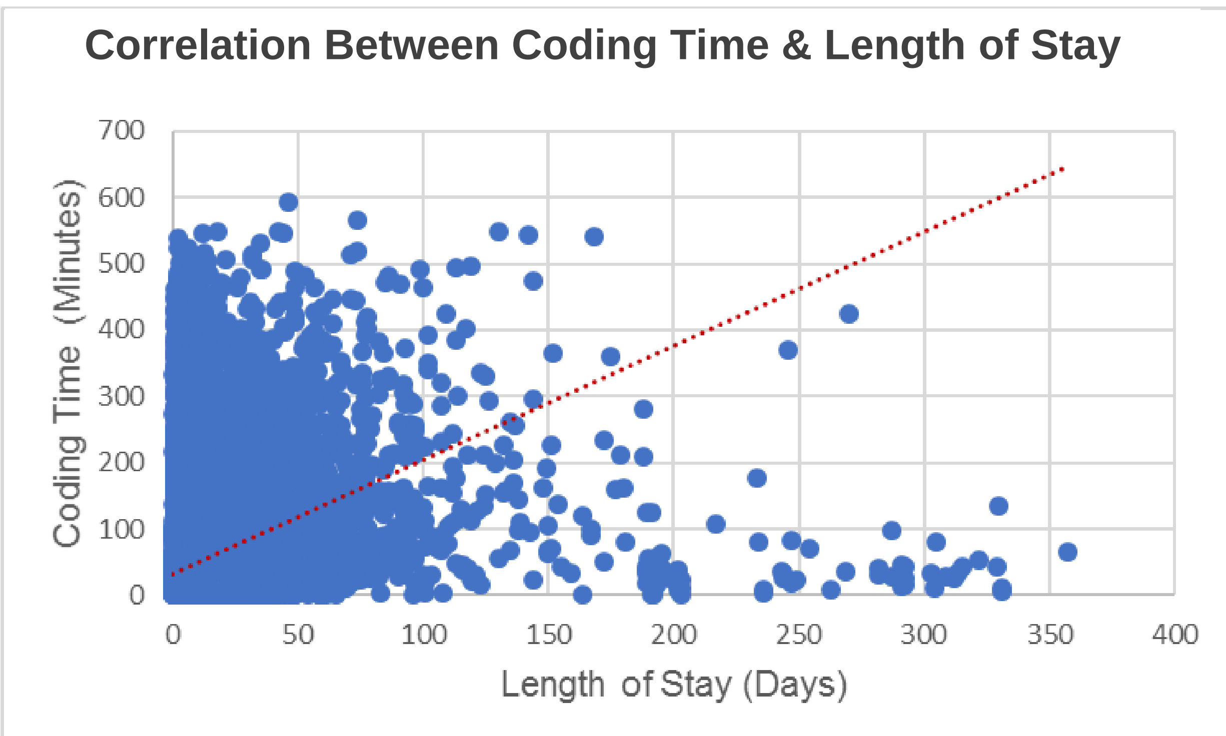


Figure 27: Coding Time by Length of Stay

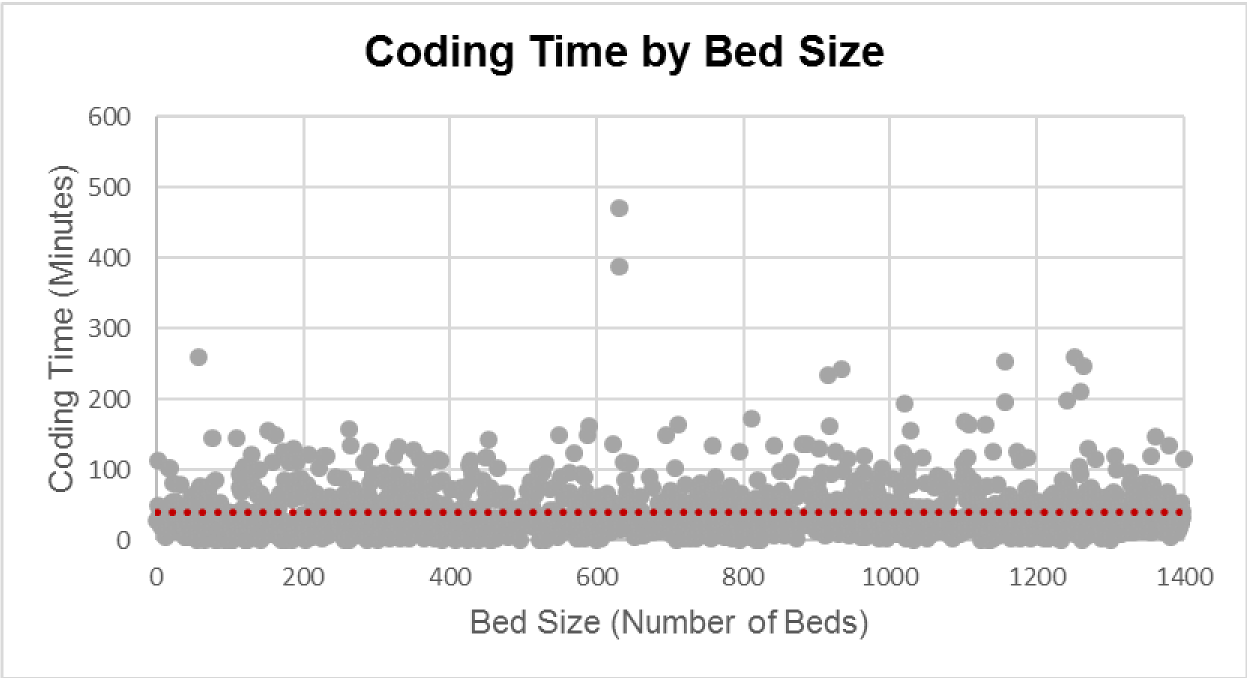


Figure 28: Coding Time by Bed Size

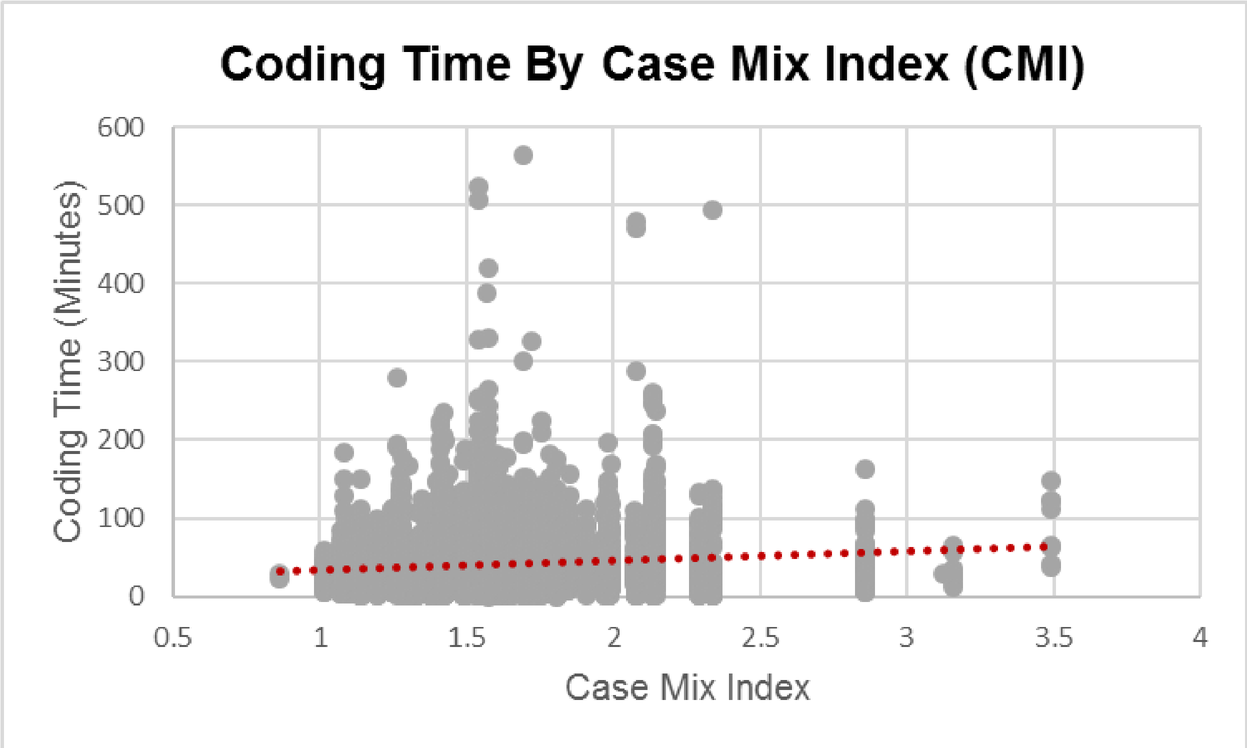


Figure 29: Coding Time by CMI

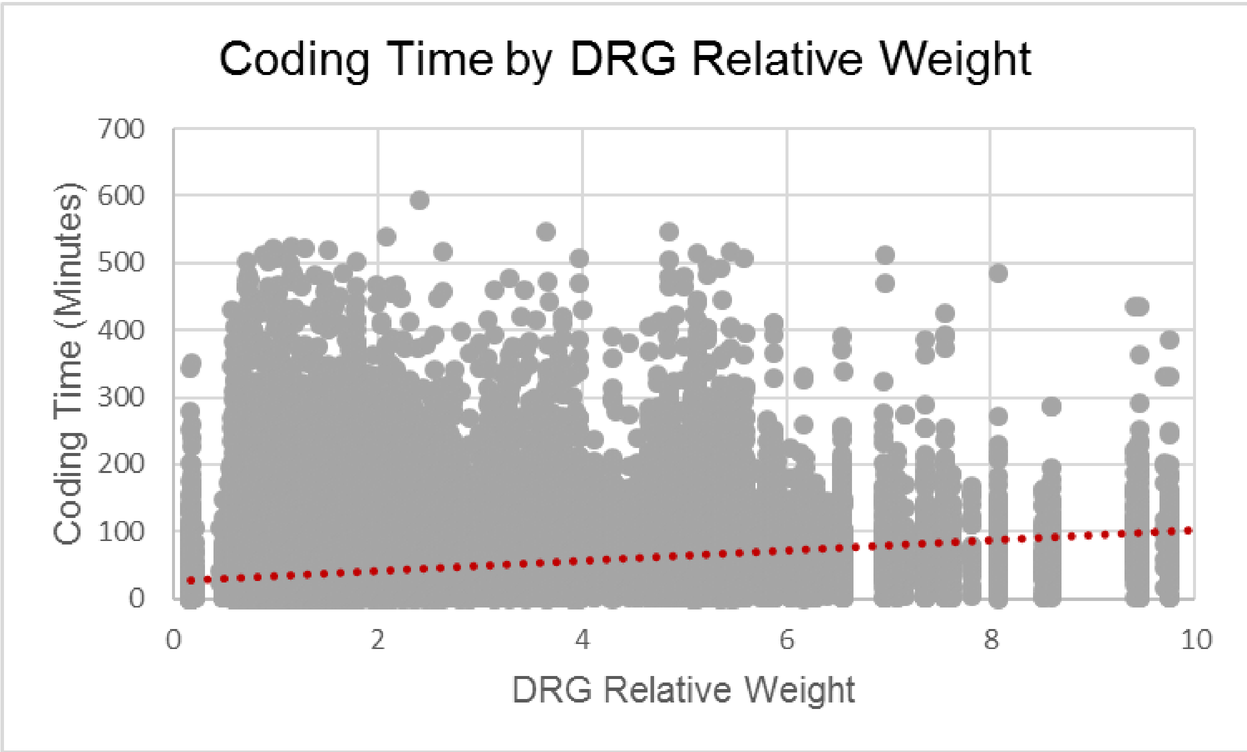


Figure 30: Coding Time by DRG Relative Weight

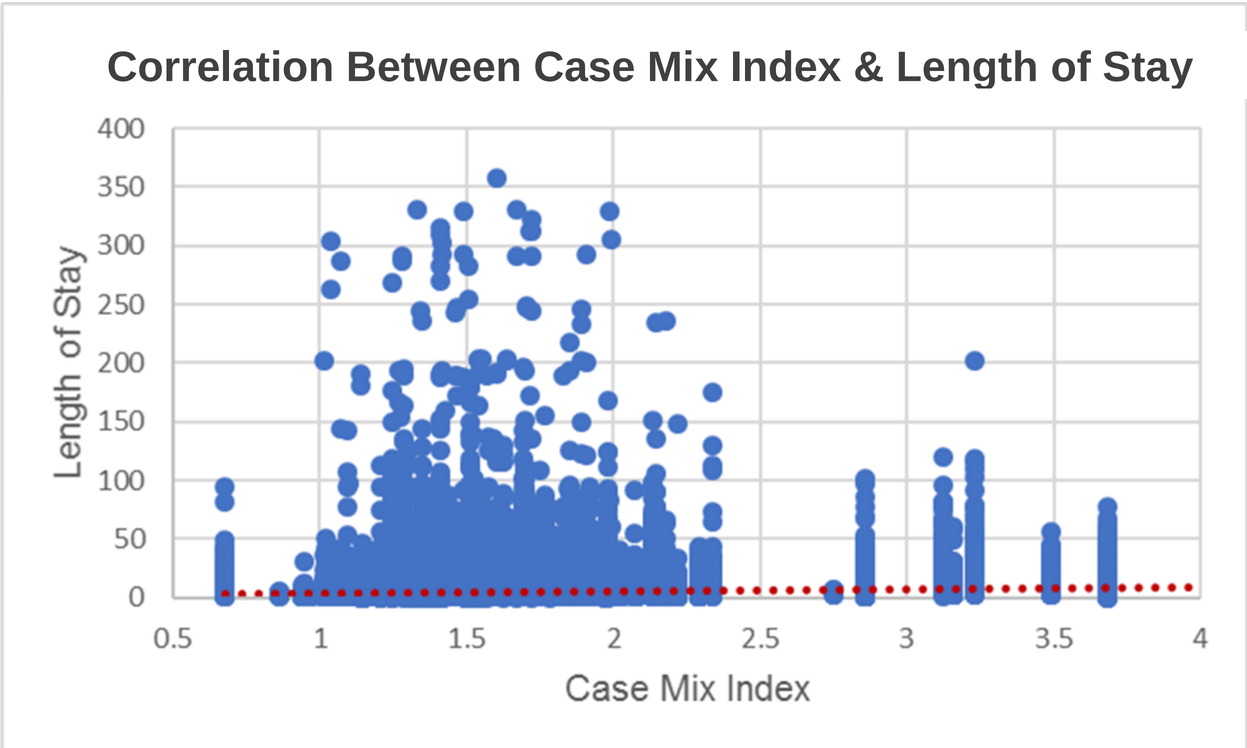


Figure 31: Correlation Between CMI and LOS

9.4.3 Multiple Linear Regression

Predicting Coding Productivity based on LOS and CMI

To predict ICD-10 coding time based on LOS and CMI, a model using multiple linear regression was developed. LOS and CMI were found to be significant predictors of ICD-10 coding productivity. Average coding time is 39.54 minutes for LOS of 4.86 and CMI of 1.57. Coding time increases by approximately two minutes for each additional day in LOS accounting for all variables in the model. Also, coding time increases by 9 minutes for each additional unit increase in CMI (.9 minute for each .10 increase in CMI) accounting for all other variables in the model. Furthermore, LOS and CMI combined account for 12 percent of ICD-10 coding productivity.

Table 37. Predicting Coding Productivity Based on LOS & CMI

Model	Predictors	Regression Equation	R2	SE
1	LOS CMI	Coding Time= 39.54+ 1.679(ALOS)+ 8.883 (CMI)	0.117	33.745

A multiple linear regression was calculated to predict coding time based on LOS and CMI (table 37). A significant regression equation was found (F(2, 323,111)= 19267.781, p < .001), with an R of .117. Coding time is equal to 39.459+ 1.679 (LOS)+ 8.883 (CMI), where LOS is measured in days.

Coding time increases by approximately two minutes for each additional day in LOS accounting for all variables in the model. Also, coding time increases by 9 minutes for each additional unit increase in CMI (0.9 minute for each 0.10 increase in CMI) accounting for all other variables in the model. Both LOS and CMI are significant predictors: combined account for approximately 12 percent of ICD-10 coding productivity. Thus, 12% of the variation in coding time can be explained by differences in LOS and CMI. It takes more time to code cases with longer hospital stays and as CMI for the facility increases.

The mean coding time for facilities with an average LOS of 5 days and a CMI of 1. 68 is 39.46. In this case, 95% of the time, the mean coding time would be within 67.5 minutes of being correct.

Adding Predictors to LOS-CMI Model

To improve predictive power of the LOS-CMI model, we added more predictors to help further explain ICD-10 coding productivity. LOS, CMI, Relative DRG weight, bed size, teaching status, and trauma status were all included in the second model (table 38).

Table 38. Predicting Coding Productivity Based on LOS, CMI, DRG Weight, Bed Size, Trauma Status, and Teaching Status

Model	Predictors	Regression Equation	R2	SE
2		Coding Time= 33.811+	0.175	33.743
	LOS	1.387(ALOS) +		
	CMI	5.100 (CMI) +		
	Bed size	.007 (Bed Size) +		
	DRG Weight	.5227 (DRG Weight) +		
	Teaching Status	1.507 (Teaching) +		
	Trauma Status	2.040 (Trauma)		

A multiple linear regression was calculated to predict coding time based on LOS and CMI bed size, DRG weight, teaching status, and trauma status. A significant regression equation was found ($F(6, 323{,}111) = 8371.984$, $p < .001$), with an R of .175. Coding time is equal to 33.811+ 1.387 (LOS)+ 5.1 (CMI)+ .007 (Bed Size) + .5227 (DRG weight) + 1.507 (Teaching) + 2.040

(Trauma) *where* LOS is measured in days; Bed Size is the number of beds; Teaching is coded as 1=Non-Teaching & 2= Teaching; and Trauma is coded as 1= Non-Trauma & 2=Trauma.

Coding time increased by 1.39 minutes for each additional day in LOS accounting for all variables in the model. Also, coding time increased by approximately 5 minutes for each additional unit increase in CMI (.5 minute for each .10 increase in CMI) accounting for all other variables in the model. For each 10 additional facility beds, coding time increases by approximately .1 minutes (.007 for each bed).

In addition, coding time increased by .5 minute for each additional unit increase in DRG weight accounting for all other variables in the model. Finally, it takes an additional 1.5 minutes on average to code patient charts in teaching facilities compared to non-teaching facilities and an additional 2 minutes on average in trauma centers compared to non-trauma facilities.

LOS, CMI, Bed Size, DRG Relative Weight, Teaching Status, and Trauma Status are significant predictors. Combined, they account for 17.5% of variability in coding time. In general, coders spend more time coding cases with longer LOS, and higher DRG relative weight. The coding time further increases in facilities with higher CMI and greater bed capacity. Teaching facilities as well as trauma centers tend to have increased coding time compared to their non-teaching and non-trauma counterparts.

The mean coding time for facilities with average length of stay of 5 days, CMI of 1. 68, DRG weight of 1.59, bed size of 509, designated as teaching hospitals and trauma centers is 33.81 minutes. In this case, 95% of the time coding time would fall between 0.11 and 67.55 minutes.

Adding Interaction Terms

Interaction between LOS and CMI was added to model 2 to see if there is a significant correlation between both variables. Interaction was significant but adding this interaction term to the model would only increase its predictive power by 0.2%. Therefore, interaction terms were not included in the predictive model. Also, keeping a simple model would increase its utilization by coders, coding managers, and other professionals with basic knowledge of statistics. Furthermore, there is no theoretical basis to support inclusion of other interaction terms: none of the variables were found to have mediating effects on coding time.

9.4.4 Hierarchical Linear Modeling

Hierarchical Linear Modeling is used to analyze nested data that are presented in multiple levels (Hox, 2002; Kreft & Leeuw, 2007; Raudenbush, 2014). In this case, data on two different levels was analyzed: level-1 is patient data and level-2 is facility data so patients are nested within facilities (clusters). There are many reasons to use HLM in analyzing coding productivity data. First, HLM has advantages over multiple linear regression for estimating the standard error for clustered data (Kreft & Leeuw, 2007). Second, HLM can reduce aggregation bias that arises when the results of aggregated data are different than results produced at the original level of observation (Kreft & Leeuw, 2007). Finally, multilevel analysis like HLM allows for a straightforward estimation of cross-level interactions (Hox, 2002; Kreft & Leeuw, 2007; Raudenbush, 2014).

Predicting Mean Coding Time Using HLM (Random-Slope Model)

This is the basic model in HLM where we allow variations in slopes. In other words, in this model we allow each facility to have a different mean coding time while holding other predictors constant across the facilities (Kreft & Leeuw, 2007).

A hierarchical linear model was developed to predict coding time based on LOS and CMI bed size, and DRG weight (teaching status and trauma status were not found to be significant predictors on coding productivity)- table 39.

Coding time increases by 1.36 minutes for each additional day in LOS accounting for all variables in the model. Also, coding time increases by approximately 8 minutes for each additional unit increase in CMI (.8 minute for each .10 increase in CMI) accounting for all other variables in the model. For each 10 additional facility beds, coding time increases by approximately .12 minutes (.012 for each bed). In addition, coding time increases by .5 minute for each additional unit increase in DRG Weight accounting for all other variables in the model.

Table 39. Hierarchical Linear Model to Predict Coding Time based on LOS and CMI bed size, and DRG weight

Model	Predictors	Regression Equation	SE
3		Coding Time= 40.02+	31.97
	LOS	1.36 (ALOS) +	
	CMI	8.44 (CMI) +	
	Bed size	.012 (Bed Size) +	
	DRG Weight	5.03 (DRG Weight)	

LOS, CMI, Bed Size, and DRG Relative Weight were found to be significant predictors of coding productivity. Combined, they account for 18.2% of variability in coding time (fixed effect).

After controlling for fixed effect, the proportion of variance in coding time due to facility effect (random effect) is 7%. In general, coders spend more time coding cases with longer LOS, and higher DRG relative weight. The coding time further increases in facilities with higher CMI and greater bed capacity. The mean coding time for facilities with average length of stay of 5 days, CMI of 1. 68, DRG weight of 1.59, bed size of 509 is 40.03 minutes.

9.5 IDENTIFYING IMPACT OF CODING TIME ON PRODUCTIVITY

A Pearson correlation coefficient was calculated for the relationship between coding time and coding accuracy. A moderate positive correlation was found (r (757)=0.316, p< .001), indicating significant linear relationship between the two variables. Increased coding time is associated with higher coding accuracy (figure 32).

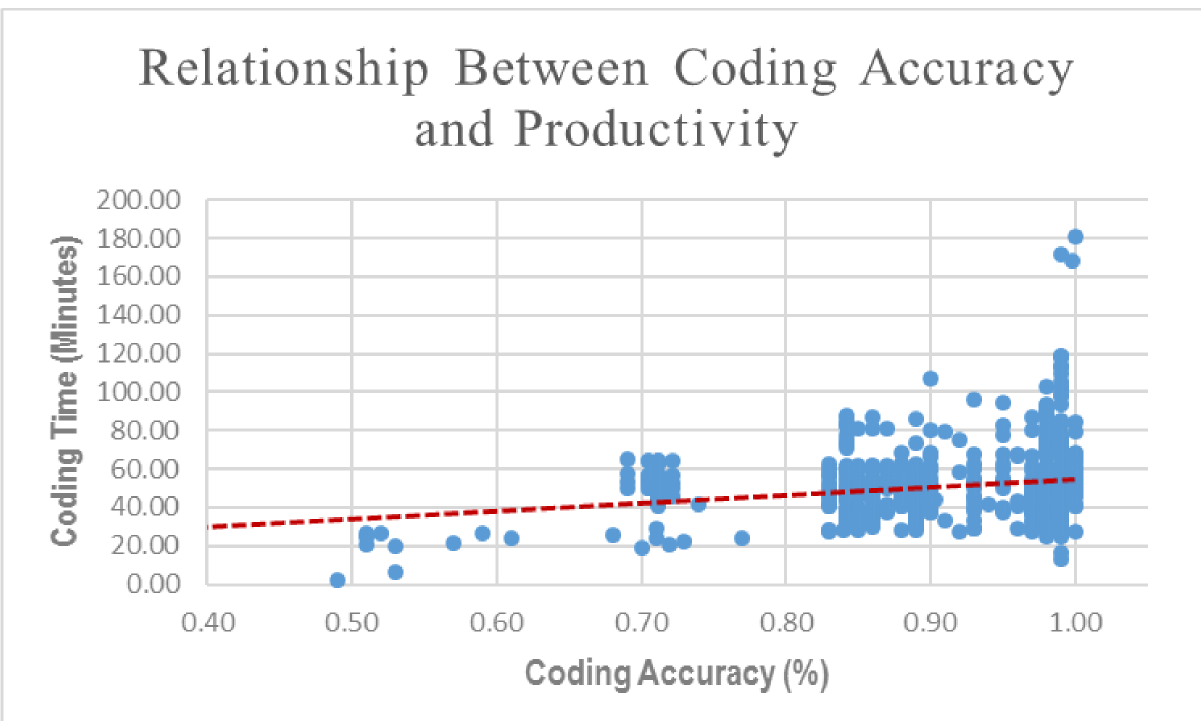


Figure 32: Relationship Between Coding Accuracy and Productivity (Coding Time)

9.6 DEVELOPING A PREDICTIVE MODEL TO PREDICT CODING QUALITY AND PRODUCTIVITY

9.6.1 Conceptual Framework:

Figure 33 represents a framework of factors influencing coding quality and productivity.

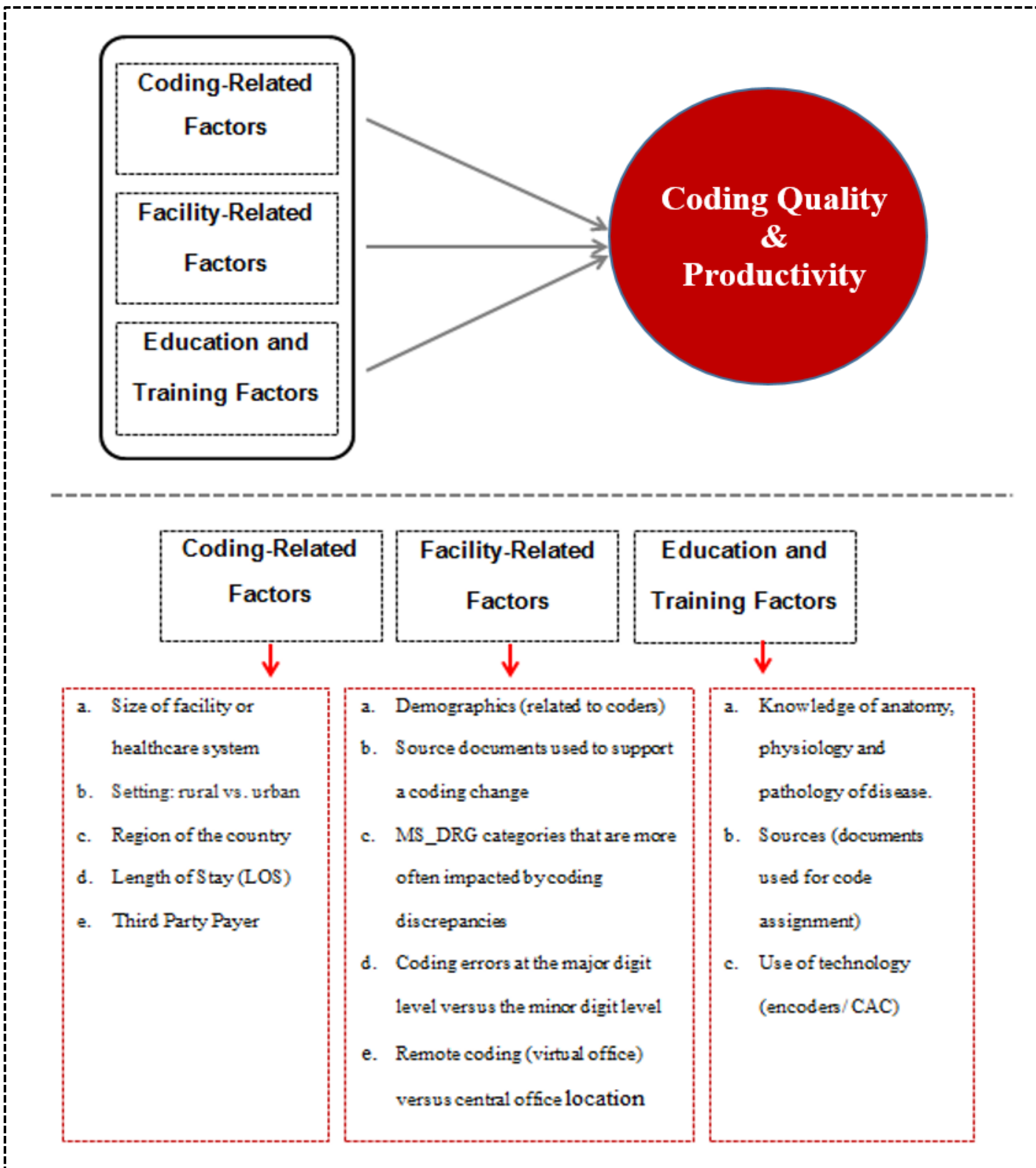


Figure 33: Conceptual Framework of Factors Influencing Coding Quality & Productivity

9.6.2 Prediction of Coding Productivity

The models shown in Table 40 can be used to predict coding productivity based on the chosen predictors. The first model includes LOS and CMI as predictors of coding productivity that was developed using multiple linear regression. The second model includes DRG relative weight, bed size, teaching status, and trauma status in addition to LOS and CMI. It was also developed using multiple linear regression. The third model which is an HLM model includes LOS, CMI, DRG relative weight, bed size as predictors of coding productivity. Teaching status and trauma status were excluded from Model 3 as they were not statistically significant. A statistically significant difference was found between Model1/Model2 and Model 3. However, all models practically generate comparable results when it comes to prediction of coding time.

9.6.3 Prediction of Coding Quality

There are many factors that can influence coding quality. Although LOS and CMI were not found to be significant predictors of coding quality, DRG weight was found to be a significant predictor. Furthermore, evidence based on qualitative analysis suggests that documentation plays critical role in predicting coding quality. Higher unspecified codes rate and physician query rate suggest issues related to documentation that could influence coding quality. Also, depth of coding can be an indicator of coding quality. Coding errors increases in cases where coders must assign multiple codes. A model to predict coding quality was not generated due to insignificant results. However, insignificant results can be attributed to the small sample size of the accuracy data (N=1010) compared to the productivity data (N= 323,112).

Table 40: Prediction of Coding Productivity (Final Models)

	Model 1		Model 2		Model 3	
	b	β	b	β	b	β
Independent Variable	(s.e)		(s.e)		(s.e)	
Length of Stay (LOS)	1.679	-	1.387	.255	1.36	-
	(.009)		(0.011)		(0.010)	
Case Mix Index (CMI)	8.883	-	6.100	.062	8.44	-
	(0.173)		(0.190)		(0.55)	
Relative DRG Weight	-	-	5.227	.221	5.03	-
			(0.047)		(0.46)	
Bed Size	-	-	.007	.056	0.012	-
			(0.000)		(0.004)	
Teaching Status	-	-	.007	.018	-	-
			(0.186)			
Trauma Status	-	-	2.040	.027	-	-
			(0.177)			
Constant	39.459		33.811		40.02	
R^2	0.117		0.175		.182	
Adjusted R^2	0.117		0.175			
N	323,112		323,112		323,112	

9.6.4 Case Studies

In this section, three case studies were selected to provide examples on prediction of coding time for three medical specialties. The case studies include the following: Oncology, Cardiovascular System, and Musculoskeletal System and Connective Tissues. Cardiovascular system and musculoskeletal system were the first two ICD-10-CM chapters with respect to coding error. Oncology was selected due to its increased coding time compared to other specialties.

9.6.4.1 Case Study 1: Circulatory System

To select the cases that only pertain to circulatory system, the data was filtered in SPSS by MDCs. The total number of cases for MDC #5 is equal to 38, 885 cases which represents 12.03% of the entire dataset. Below is the sampling distribution for the circulatory system represented by mean, standard deviation, minimum and maximum values (table 41).

Table 41. Descriptive Statistics (Circulatory System)

	Mean	95.0% Lower CL for Mean	95.0% Upper CL for Mean	Standard Deviation	Maximum	Minimum
Coding Time	43.64	43.27	44.00	36.52	516.80	0.7
Length of Stay (LOS)	5.00	5.00	5.00	9.00	315.00	1.00
Case Mix Index (CMI)	1.62	1.62	1.63	0.41	10.47	0.68
DRG Relative Weight	1.93	1.91	1.95	1.58	15.87	0.45
Bed Size (Count)	461.00	459.00	464.00	263.00	1346.00	25.00

Also, below is a breakdown of cases based on the DRG type (table 42). Medical cases approximately represent 67% of the cases (N=25,900) while surgical cases represent the remaining 33% of the cases (N= 12,985).

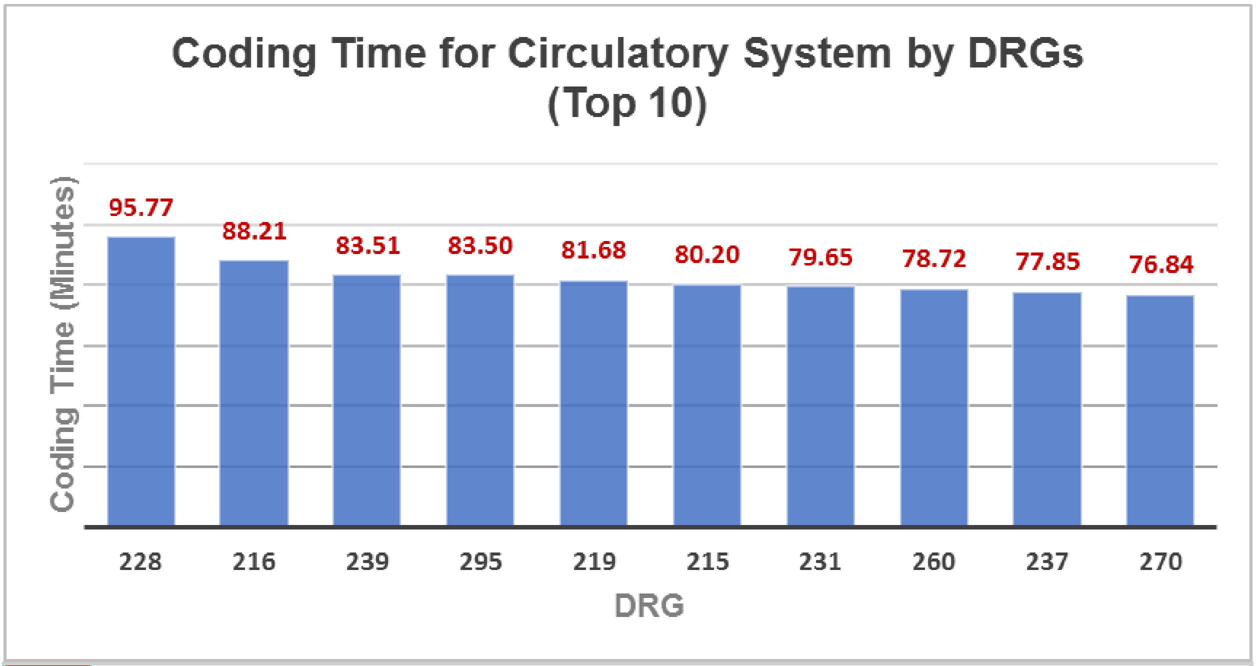
Table 42. Average Coding Time by DRG Type (Circulatory System)

	Variable	Mean	95.0% Lower CL for Mean	95.0% Upper CL for Mean	Standard Deviation	Min	Max
Medical N=25,900 (66.61%)	Coding Time	39.46	39.06	39.85	32.39	0.50	475.40
	DRG Relative Weight	1.11	1.10	1.11	0.42	0.45	2.79
	Length of Stay	4.48	4.38	4.59	8.31	1.00	315.00
Surgical N=12,985 (33.39%)	Coding Time	51.98	51.25	52.71	42.38	0.80	516.80
	DRG Relative Weight	3.57	3.54	3.60	1.76	1.08	15.87
	Length of Stay	5.87	5.72	6.02	8.90	1.00	269.00

Based on this analysis, the following represent the top 10 DRGs with highest mean coding time (table 43 and figure 34).

Table 43. Average Coding Time by DRG (Circulatory System)

DRG	DRG Relative Weight	ALOS	Coding Time	95.0% Lower CL for Mean	95.0% Upper CL for Mean	Standard Deviation	Min	Max
228	6.95	11.91	95.77	76.13	115.40	71.94	1	323.5
216	9.46	14.02	88.21	77.23	99.18	65.68	0.5	435.6
239	4.84	13.81	83.51	73.82	93.21	59.88	0.7	389.9
295	0.74	1.33	83.50	-97.66	264.66	72.93	24.4	165
219	7.56	10.76	81.68	74.81	88.55	59.22	0.7	424.4
215	15.87	7.2	80.20	13.92	146.48	53.38	1	141.7
231	7.81	12.33	79.65	63.95	95.35	44.28	3.5	166.5
260	3.73	10.73	78.72	64.36	93.09	47.26	2.8	219.9
237	5.08	6.47	77.85	45.14	110.56	59.07	1.2	209.1
270	4.73	8.87	76.84	71.54	82.13	50.51	0.9	334.7



228	OTHER CARDIOTHORACIC PROCEDURES W MCC
216	CARDIAC VALVE & OTH MAJ CARDIOTHORACIC PROC W CARD CATH W MCC
239	AMPUTATION FOR CIRC SYS DISORDERS EXC UPPER LIMB & TOE W MCC
295	DEEP VEIN THROMBOPHLEBITIS W/O CC/MCC
219	CARDIAC VALVE & OTH MAJ CARDIOTHORACIC PROC W/O CARD CATH W MCC
215	OTHER HEART ASSIST SYSTEM IMPLANT
231	CORONARY BYPASS W PTCA W MCC
260	CARDIAC PACEMAKER REVISION EXCEPT DEVICE REPLACEMENT W MCC
237	MAJOR CARDIOVASC PROCEDURES W MCC OR THORACIC AORTIC ANEURYSM REPAIR
270	OTHER MAJOR CARDIOVASCULAR PROCEDURES W MCC

Figure 34: Coding Time by DRGs (Circulatory System)

Predicting Mean Coding Time for Circulatory System

A multiple linear regression was calculated to predict coding time based on LOS, CMI and DRG relative weight (table 44). A significant regression equation was found ($F(3, 38, 885) = 1320.461$, $p < .001$), with an R of .175. Coding time is equal to 25.15+ 1.692 (LOS)+ 3.136 (CMI)+ 5.170 (DRG Weight), where LOS is measured in days.

Table 44: Predicting Mean Coding Time for Circulatory System

Model	Predictors	Regression Equation	R2	SE
Circulatory System	LOS CMI DRG Weight	Coding Time= 25.15+ 1.692(ALOS) + 3.136 (CMI) + 5.170 (DRG Weight)	0.113	34.79

The mean coding time for circulatory system cases with average LOS of 5 days, CMI of 1.62, DRG weight of 1.93 is 25.15 minutes.

9.6.4.2 Case Study 2: Musculoskeletal System and Connective Tissues

To select the cases that only pertain to musculoskeletal system and connective tissues, the data was filtered in SPSS by MDCs. The total number of cases for MDC #8 (Musculoskeletal System and Connective Tissues) is equal to 29, 630 cases which represents 9.17% of the entire dataset. Table 45 demonstrates the sampling distribution for the circulatory system represented by mean, standard deviation, minimum and maximum values.

Also, below is a breakdown of cases based on the DRG type (table 46). Medical cases approximately represent 22.37% of the cases (N=6,628) while surgical cases represent the remaining 77.63% of the cases (N= 23,002).

Table 45. Descriptive Statistics (Musculoskeletal System and Connective Tissues)

	Mean	95.0% Lower CL for Mean	95.0% Upper CL for Mean	Standard Deviation	Min	Max
Coding Time	39.86	39.32	39.91	33.97	0.34	538.23
Length of Stay (LOS)	3.96	3.00	4.00	4.32	1.00	322
Case Mix Index (CMI)	1.619	1.5	1.64	0.36534	0.6766	3.68
DRG Relative Weight	2.23	2.08	2.20	1.29	0.63	11.43
Bed Size (Count)	494	480	482	280	25	1346

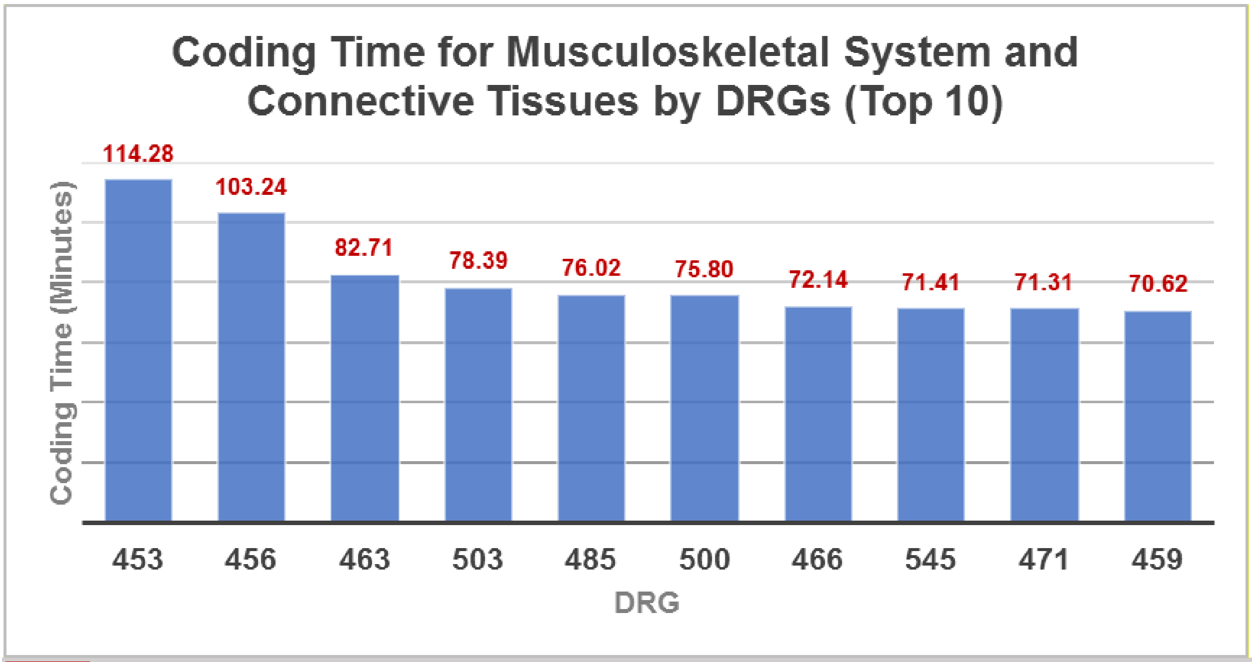
Table 46. Average Coding Time by DRG Type (Musculoskeletal System and Connective Tissues)

	Variable	Mean	95.0% Lower CL for Mean	95.0% Upper CL for Mean	Standard Deviation	Min	Max
Medical N=6,628 (22.37%)	Coding Time	39.83	32.00	33.40	32.10	0.30	475.7
	DRG Relative Weight	1.03	0.86	0.91	0.35	0.63	2.4409
	Length of Stay	4.60	3.00	4.00	6.15	1.00	322
Surgical N=23,002 (77.63%)	Coding Time	39.87	31.10	31.80	34.49	0.70	538.2
	DRG Relative Weight	2.58	2.08	2.20	1.25	0.91	11.4304
	Length of Stay	3.77	3.00	4.00	3.61	1.00	71

Based on this analysis, the following represent the top 10 DRGs with highest mean coding time (table 47 and figure 35).

Table 47. Average Coding Time by DRGs (Musculoskeletal System and Connective Tissues)

DRG	DRG Relative Weight	ALOS	Coding Time	95.0% Lower CL for Mean	95.0% Upper CL for Mean	Standard Deviation	Min	Max
453	11.43	10.00	114.28	61.00	90.90	121.97	1.80	503.60
456	9.41	11.00	103.24	74.90	111.90	68.93	2.00	435.70
463	5.10	11.00	82.71	59.10	84.40	63.36	1.30	423.30
503	2.27	8.00	78.39	61.70	92.90	32.91	17.40	147.60
485	3.21	13.00	76.02	48.30	81.10	56.09	1.20	233.00
500	3.20	10.00	75.80	60.30	75.30	48.55	0.80	287.20
466	5.04	8.00	72.14	52.30	75.70	45.73	1.90	279.50
545	2.44	10.00	71.41	48.90	70.40	54.59	1.90	314.70
471	4.90	8.00	71.31	48.60	73.80	54.48	1.10	324.40
459	6.55	8.00	70.62	53.70	70.20	41.23	1.70	200.20



453	COMBINED ANTERIOR/POSTERIOR SPINAL FUSION W MCC
456	SPINAL FUS EXC CERV W SPINAL CURV/MALIG/INFEC OR 9+ FUS W MCC
463	WND DEBRID & SKN GRFT EXC HAND, FOR MUSCULO-CONN TISS DIS W MCC
503	FOOT PROCEDURES W MCC
485	KNEE PROCEDURES W PDX OF INFECTION W MCC
500	SOFT TISSUE PROCEDURES W MCC
466	REVISION OF HIP OR KNEE REPLACEMENT W MCC
545	CONNECTIVE TISSUE DISORDERS W MCC
471	CERVICAL SPINAL FUSION W MCC
459	SPINAL FUSION EXCEPT CERVICAL W MCC

Figure 35: Coding Time by DRGs (Musculoskeletal System & Connective Tissues)

Predicting Mean Coding Time for Musculoskeletal System and Connective Tissues

A multiple linear regression was calculated to predict coding time based on LOS, CMI and DRG relative weight (table 48).

A significant regression equation was found (F(3, 29, 630)= 1418.465, $p < .001$), with an R of 0.126. Coding time is equal to 39.33 + 2.36 (LOS)+ 4.12 (CMI)+ 3.76 (DRG Weight), where LOS is measured in days.

Table 48. Predicting Mean Coding Time for Musculoskeletal System and Connective Tissues

Model	Predictors	Regression Equation	R2	SE
Circulatory System	LOS CMI DRG Weight	Coding Time= 39.33+ 2.36 (ALOS) + 4.12 (CMI) + 3.76 (DRG Weight)	0.126	31.76

The mean coding time for circulatory system cases with average LOS of 3.96 days, CMI of 1. 62, DRG weight of 2.23 is 39.33 minutes.

9.6.4.3 Case Study 3: Oncology

To select the cases that only pertain to oncology, the data was filtered to include DRGs related to oncology. The total number of cases of oncology in this sample is equal to 10,206 cases which represents 3.16% of the entire dataset. Table 49 demonstrates the sampling distribution Oncology m represented by mean, standard deviation, minimum and maximum values.

Also, below is a breakdown of cases based on the DRG type (table 50). Medical cases approximately represent 52.04% of the cases (N=5,311) while surgical cases represent the remaining 47.96% of the cases (N= 4,895).

Table 49. Descriptive Statistics (Oncology)

	Mean	95.0% Lower CL for Median	95.0% Upper CL for Median	Standard Deviation	Min	Max
Coding Time	48.15	47.40	48.91	38.89	0.9	474.30
Length of Stay (LOS)	5.37	5.24	5.50	6.49	1.00	188.00
Case Mix Index (CMI)	1.62	1.61	1.63	0.34	0.68	10.47
DRG Relative Weight	1.91	1.89	1.94	1.36	0.59	9.41
Bed Size (Count)	560	554	566	298	48	1346

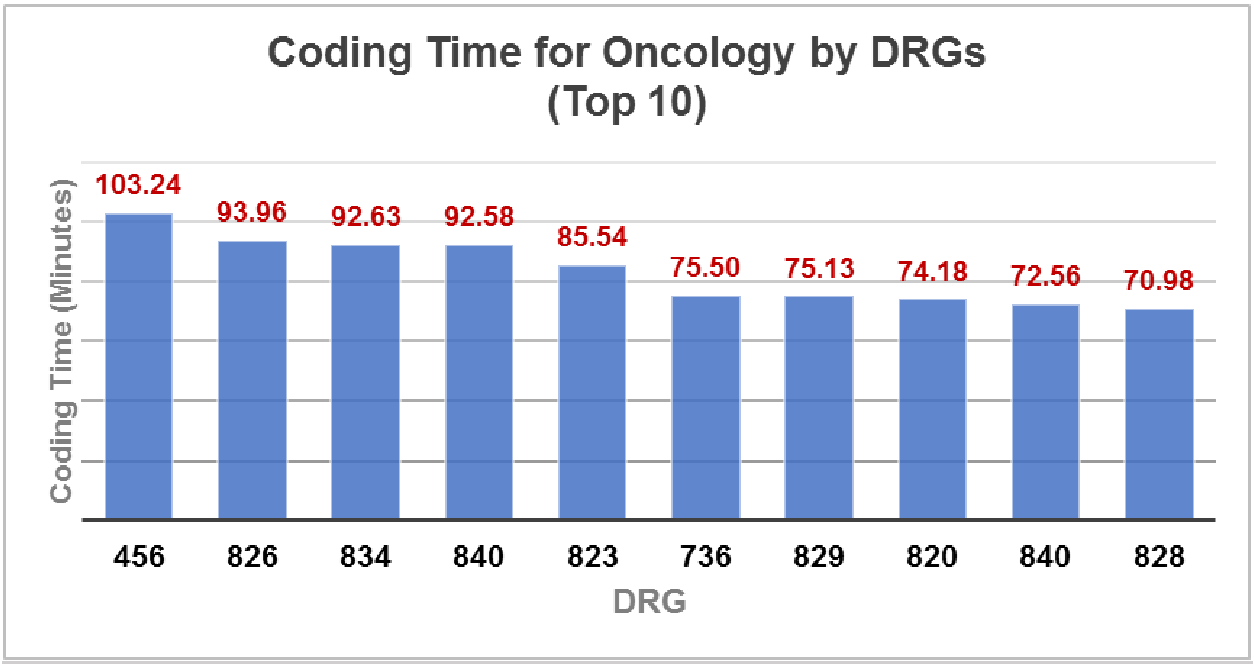
Table 50. Average Coding Time by DRG Type (Oncology)

	Variable	Mean	95.0% Lower CL for Mean	95.0% Upper CL for Mean	Standard Deviation	Min	Max
Medical N=5,311 (52.04%)	Coding Time	47.97	46.95	48.99	37.81	0.9	474.30
	DRG Relative Weight	1.65	1.62	1.68	0.98	0.59	6.13
	Length of Stay	6.06	5.88	6.25	6.93	1.00	181.00
Surgical N=4,895 (47.96%)	Coding Time	48.35	47.23	49.48	40.03	0.8	435.70
	DRG Relative Weight	2.20	2.16	2.25	1.63	0.85	9.41
	Length of Stay	4.62	4.45	4.78	5.88	1.00	188.00

Based on this analysis, the following represent the top 10 DRGs with highest mean coding time (table 51 and figure 36).

Table 51. Average Coding Time by DRGs (Oncology)

DRG	DRG Relative Weight	ALOS	Coding Time	95.0% Lower CL for Mean	95.0% Upper CL for Mean	Standard Deviation	Min	Max
456	9.41	11.00	103.24	84.43	122.06	68.93	2.00	435.70
826	2.30	7.00	93.96	68.37	119.54	59.16	18.70	290.40
834	3.14	11.00	92.63	81.03	104.24	66.88	0.60	394.90
840	1.19	4.00	92.58	-14.69	199.85	86.39	32.20	244.50
823	1.51	4.00	85.54	68.34	102.74	62.41	0.50	381.30
736	4.33	11.00	75.50	50.61	100.38	58.93	1.00	279.80
829	2.77	8.00	75.13	56.84	93.42	67.01	2.30	375.30
820	1.40	4.00	74.18	57.15	91.21	51.08	12.30	203.90
840	2.46	8.00	72.56	65.14	79.98	53.33	0.00	392.30
828	6.13	16.00	70.98	55.41	86.55	44.63	10.30	180.70



456	SPINAL FUS EXC CERV W SPINAL CURV/MALIG/INFEC OR 9+ FUS W MCC
826	MYELOPROLIF DISORD OR POORLY DIFF NEOPL W MAJ O.R. PROC W MCC
834	ACUTE LEUKEMIA W/O MAJOR O.R. PROCEDURE W MCC
840	LYMPHOMA & NON-ACUTE LEUKEMIA W MCC
823	LYMPHOMA & NN-ACUTE LEUKEMIA W OTHER O.R. PROC W MCC
736	UTERINE & ADNEXA PROC FOR OVARIAN OR ADNEXAL MALIGNANCY W MCC
829	MYELOPROLIF DISORD OR POORLY DIFF NEOPL W OTHER O.R. PROC W CC/MCC
820	LYMPHOMA & LEUKEMIA W MAJOR O.R. PROCEDURE W MCC
840	LYMPHOMA & NN-ACUTE LEUKEMIA W MCC
828	MYELOPROLIF DISORD OR POORLY DIFF NEOPL W MAJ O.R. PROC W/O CC/MCC

Figure 36: Coding Time by DRG (Oncology)

Predicting Mean Coding Time for Oncology

A multiple linear regression was calculated to predict coding time based on LOS, CMI and DRG relative weight. A significant regression equation was found ($F(3, 10{,}206) = 900.645$, $p < .001$), with an R of 0.126. Coding time is equal to 46.36+ 2.17 (LOS)+ 6.01 (CMI)+ 4.57 (DRG Weight), where LOS is measured in days. The mean coding time for circulatory system cases with average LOS of 5.37 days, CMI of 1. 62, DRG weight of 1.91 is 46.36 minutes.

Table 52. Predicting Mean Coding Time for Oncology

Model	Predictors	Regression Equation	R2	SE
Circulatory System	LOS CMI DRG Weight	Coding Time= 46.36+ 2.17 (ALOS) + 6.01 (CMI) + 4.57 (DRG Weight)	0.209	34.58

10.0 DISCUSSION

Coding constitutes one of the fundamental functions of HIM. Clinical coding is widely utilized in the health care system across the nation. Clinical terminology and classification systems have been developed to meet the increasing demand for data-driven decision making in health care, especially with rapid adoption of health information technology. Inpatient coding, however, represents the major focus of this dissertation research. ICD-10-CM is the system used for inpatient coding in the U.S. and it is the U.S clinical modification of the WHO's ICD-10.

ICD-10-CM represents the foundation of reimbursement in the U.S health care system. Also, coded clinical data is utilized to compile wide range of statistics and quality of care indicators. It can be used to evaluate clinical outcomes for individual patients, compare performance between health care organizations, or to compile the major causes of mortality and mobility at the public health level. In addition, this data can be utilized for education, research, and healthcare services utilization.

Therefore, it is important to address two major aspects related to quality of clinical coded data including quality and productivity. Ensuring quality of coded data can significantly contribute to reliable data-driven decision-making. However, accuracy can be useless if data is not processed in a timely and efficient manner. This dissertation research aims at identifying current coding trends, and factors that could influence coding quality and productivity. The significance of this study lies in three premises: (1) coding is not considered a revenue-generating activity and thus is underutilized in health care research; (2) this study tries a new approach to coding using quantitative and qualitative methods along with statistics and data analytics; and (3) it tries to establish a connection between coding quality and productivity- a topic that has never been addressed based on real data analysis.

This dissertation research utilized two different data sets for this purpose: (1) accuracy data set (N=1,010) and productivity data (323,112). All cases were provided by Ciox Health. The first data set includes audited coded data in ICD-9-CM while the second data set includes ICD-10-CM productivity data. SAS, SPSS, and Nvivo were used for data analysis. Data analysis includes univariate (descriptive), and bivariate analyses. Also, linear, and multiple regressions were performed in addition to t-tests. A hierarchical linear model was further developed to account for the nested productivity data.

Many factors were found to have a significant impact of coding quality and productivity. Although LOS and CMI were not found to be significant predictors of coding quality, DRG weight was found to be a significant predictor. Qualitative evidence suggests that documentation plays critical role in predicting coding quality. Furthermore, higher unspecified codes rate and physician query rate suggest issues related to documentation that could influence coding quality. In this sample, the unspecified codes rate and physician query rate were approximately 15% and 2%, respectively, which suggests relatively high documentation standards in the participated facilities.

History and physical examination, discharge summary, and progress notes were identified as the most frequent documents cited for coding change. Also, this study found that there are some issues related to the coding guidelines that could influence coding quality including: symptoms & signs, principle diagnosis, secondary and additional diagnoses, and combination coding. In general, the accuracy rate of this sample was around 94%. Furthermore, the accuracy rate increased to approximately 95% when accounting for depth of coding. The second method of measuring accuracy has been developed in this study to meet the demand for advanced coding metrics that account for complex variables including depth of coding when measuring coding accuracy (Stanfill, 2016). Depth of coding can be an indicator of coding quality. Coding errors increase in cases where coders must assign multiple codes (Stanfill, 2016). A model to predict coding quality was not generated due to insignificant results. However, insignificant results can be attributed to the small sample size of the accuracy data (N=1010) compared to the productivity data (N= 323,112).

Regarding coding productivity, many factors were found to be significant predictors of coding productivity (coding time in minutes). Specifically, LOS, CMI, DRG weight, bed size, trauma status, and teaching status were found to have statistically significant effects on coding time when performing multiple linear regression. However, trauma status and teaching status were not statistically significant when using HLM that was required to account for the nested design. In addition, this study found that there is a significant positive (moderate) correlation between coding time and coding quality. This leads us to the conclusion that if coding time increases, coding quality (accuracy) increases. Coders do not have to sacrifice quality for quantity. Although the relationship is locally linear, a possibility of existence of non-linear form should be further investigated.

It should be noted that data on coders' demographics could not be secured for this study. Linking attributes such as coder's education, years of experience, and credentials represents a major opportunity for future research. Based on secondary analysis performed by the researcher, coder's education and credentials accounts for approximately 17% of coding variability accounting for all variables in the model. Furthermore, clinical coding represents a promising area for qualitative research in HIM. Qualitative analyses could be exceptionally beneficial if applied in coding complex training on coding. Data-driven decision making can be made more effective, reliable and of less risk if mixed research methods, applied statistics, and data analytics techniques are utilized for health care quality improvement.

11.0 LIMITATIONS

The following represent major limitations and challenges to this study: (1) Sampling bias; (2) external validity; (3) non-blind review; (4) obtaining accuracy data on ICD-10; and (4) obtaining coders' demographics.

I. Sampling Bias: Sampling bias that is a result of using non-probability sampling. This is especially true in case of accuracy (N=1,010). However, in case of productivity sampling bias is not an issue due to very large sample (N= 323,112).

II. External Validity: The ability to generalize study's results (accuracy) could be limited since the sample used for this study in not representative of the entire population.

III. Non-blind Review: Auditors were not blinded in chart review process which could represent another source of bias with respect to the review process.

IV. Obtaining ICD-10 data: ICD-10 accuracy data could not be obtained from Ciox Health.

V. Correlation between accuracy and productivity: to establish a connection between coding productivity and accuracy, the two data sets were linked using DRGs. However, this violates the assumption of Pearson's correlation as the pairs of observations used for this purpose were not related (coming from two different populations).

VI. Obtaining Coders' demographics: Larger variance in coding productivity is yet to be explained. Specific information such as education, training, and years of the coders' experience would have significantly contributed to more accurate prediction of coding time. Coders' information could not be obtained in this study and therefore, a vital piece of information was missing from this analysis.

12.0 FUTURE WORK

This study aimed at identifying factors influencing current coding trends, coding quality, and productivity. Length of stay, CMI, DRG weight, bed size, teaching status, and trauma status were found to have significant impact on coding productivity.

However, variability in coding productivity due to coder-related factors represents a promising area for future research. Although length of stay, case mix index, and DRG weight were not found to be significant determinants of coding quality, the results are inconclusive due to sample size. Therefore, identifying significant determinants of coding quality is still an open area that needs further investigation if a larger and more representative sample could be granted for this purpose.

Furthermore, clinical coding and classification as a sub-specialty of HIM has great potential for qualitative research. Specifically, qualitative research should be further utilized in identifying potential coding scenarios based on the coding audit for education and training purpose. Clinical coding and classification represents one area where research is underutilized and there is more truth to be revealed using various methods of data analytics, statistics, and research.

13.0 CONCLUSION

Coding is one of the most critical functions of HIM and has different applications in health care. This study identified current coding trends and factors that might influence coding quality and productivity. It found that coding productivity in ICD-10 improved over time. Length of stay, case mix index, DRG weight, bed size, and teaching as well as trauma status were found to be significant factors that influence coding productivity. However, length of stay, case mix index, and DRG weight were not found to have significant influences on coding quality. Based on the qualitative analysis, H&P, discharge summary, and progress notes were identified as the three most common resources to guide coders through the coding audit process. Coders' demographics could not be granted for this study. However, factors related to coders such as education, credentials, and years of experience are believed to have significant impact on coding quality as well as productivity, which are to be further explored in future opportunities.

APPENDIX A

GLOSSARY

Case Mix Index: The average diagnosis-related group (DRG) relative weight for that hospital. It is calculated by summing the DRG weights for all Medicare discharges and dividing by the number of discharges.

Classification: A system that arranges or organizes likes or related entities; also, a system for assigning numeric or alphanumeric code to represent specific diseases and/or procedures.

Clinical terminology: A set of standardized terms and their synonyms that record patient findings, circumstances, events, and interventions with sufficient detail to support clinical care, decision support, outcomes research, and quality improvement.

Code Set: under HIPAA, any set of codes used to encode data elements, such as tables of terms, medical concepts, medical diagnostic codes, or medical procedure codes; includes both the codes and their descriptions.

Coding: The process of translating descriptions of diseases, injuries, and procedures into numeric or alphanumeric designations

Controlled medical terminology: A coded vocabulary of medical concepts and expressions used in healthcare

Controlled vocabulary: A restricted set of phrases, generally enumerated in a list and perhaps arranged into a hierarchy

Current Procedural Terminology (CPT): A comprehensive list of descriptive terms and codes published by the American Medical Association and used for reporting diagnostic and therapeutic procedures and other medical services performed by physicians

Derived Classification: One based on a reference classification such as ICD or ICF by adopting the reference classification structures and categories and providing additional detail or through rearrangements and aggregation of items from one or more reference classifications

Healthcare Common Procedure Coding System (HCPCS): A two-level classification system introduced in 1983 to standardize the coding systems used to process Medicare and Medicaid claims

> **HCPCS level I:** Current Procedural Terminology (CPT), developed by the American Medical Association
>
> **HCPCS level II:** Codes not covered by CPT and modifiers that can be used with all levels of codes, developed by the Centers for Medicare and Medicaid Services
>
> **HCPCS level III:** Codes, often called local codes, developed by local Medicare and/or Medicaid carriers for use in their geographic locations; eliminated on December 31, 2003

International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM): A classification system used in the United States to report morbidity information

International Classification of Diseases, Tenth Revision (ICD-10): The most recent revision of the disease classification system developed and used by the World Health Organization to track morbidity and mortality worldwide

National Center for Health Statistics (NCHS): The federal agency responsible for collecting and disseminating information on health services utilization and the health status of the population in the United States; developed the clinical modification to the International Classification of Diseases, Ninth Revision (ICD-9) and is responsible for updating the diagnosis portion of the ICD-9-CM

Prospective payment system (PPS): A type of reimbursement system based on preset payment levels rather than actual charges billed after a service has been provided; specifically, one of several Medicare reimbursement systems based on predetermined payment rates or periods and linked to the anticipated intensity of services delivered as well as the beneficiary' condition

Reference terminology: A set of concepts and relationships that provide a common reference point for comparisons and the aggregation of data about the entire healthcare process, recorded by multiple different individuals, systems, or institutions

Related Classification: Partially refers to a reference classification or is associated with the reference classification at specific level of structure only and describes important aspects of health or the health system not covered by reference or derived classifications

Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT): A systematized, multi-axial, and hierarchically organized controlled terminology developed by the College of American Pathologists and currently owned by the International Health Terminology Standards Development Organization

Terminology: A set of terms representing the system of concepts of a subject field

Vocabulary: A list or collection of clinical words or phrases with their meanings; also, the set of words used by an individual or group within a subject field

EXAMPLES I: *SOURCE DOCUMENTS USED TO IDENTIFY CODING ERRORS*

(Figures 37-43)

Cases in Which H&P Was Used for Coding Change
Adding personal and family history codes, "Add 412 history of heart attack per consultation and H/P"; "Add 305.1 tobacco use per H/P"
"Change 414.01 CAD of native vessel to 414.00 CAD of unspecified vessel, native or graft per H/P"
"Recommend adding V49.86 (Do not resuscitate status) per the last paragraph in the H&P and the Swing-bed Admission Orders form".
"Recommend adding 276.7 (hyperkalemia) as documented in the H&P and DS. Treatment began in the ER but monitoring continued during the in-house stay"
"Recommend adding 244.9 (hypothyroidism) as documented in the H&P. Patient under treatment with Levothyroxine"
"Recommend adding 428.0 (CHF) as documented in the H&P and as the component code to 428.22"
"Recommend adding V58.66 (long term aspirin) and V49.86 (do not resuscitate) per H&P documentation"
"From the H&P for the acute care stay, recommend adding 414.01 (CAD), 412 (old MI), 401.9 (hypertension), 250.00 (diabetes mellitus), V58.61 (long term use of coumadin), and V58.66 (long term use of aspirin)"
"Recommend adding 799.02 (hypoxemia) as documented in the H&P. This documents the effects on breathing by the pneumonia"
"Change 246.9 unspecified disorder of thyroid to 244.9 hypothyroidism per H/P"
"Add 272.4 hyperlipidemia per H/P"
"Change 369.9 unspecified visual loss to 368.8 other specified visual disturbances per H/P blurred vision"
"Report 2724. (dyslipidemia) on Plavix, V15.82 (h/o tobacco use) per H&P"

Figure 37. Cases in Which H&P Was Used for Coding Change

Cases in Which Discharge Summary Was Used for Coding Change
“Change POA from (Y) to (N) for E960.0 unarmed fight or brawl per discharge summary on second day father struck patient.”
Identifying organisms in discharge summary, “Add 041.49 E. coli per discharge summary”
“Recommend adding 285.9 (anemia) per the DS statement of anemia and the prescription for iron supplement. The patient's Hgb remained in the 8 gram range for the stay”
“Recommend adding 599.60 for obstructive uropathy as documented in the discharge summary. It was noted that a bladder catheter was inserted and the patient went home with it for removal after seeing the urologist in his office”
“Delete 288.60 leukocytosis, unspecified as this condition is inherent in diverticulitis. Per discharge summary patient's symptoms resolved on IV flagyl and cipro; discharged home on po flagyl and cipro x 10 days”
“Add 300.00 anxiety per discharge summary. Patient given Xanax for recurrent stressors”

Figure 38. Cases in Which Discharge Summary Was Used for Coding Change

Cases in Which Progress Notes Were Used for Coding Change
“Add 599.0 UTI per discharge summary and progress notes 6/22 patient was treated for same”
“Add 780.2 syncope and collapse (fainting) per progress notes 6/24”
“Recommend adding 599.0 (UTI) as documented in progress note 6/12. There are multiple entries of E coli positive urine culture”
“Recommend adding 03.31 (spinal tap) as documented in the progress notes on 6/12”
Discharge disposition in progress notes, “Change to (01) home per progress notes”
“Add 584.9 AKI (acute kidney injury) per progress notes 10/02”
“Add 724.5 back pain per progress note 2/17; patient given morphine sulfate extended release”
“Add 799.02 hypoxemia per progress note 2/20”
“Change 496 COPD to 491.21 COPD with acute exacerbation per pn 2/20”
No Error Assessed as electronically signed after coding. Add 250.00 diabetes per progress note 3/02”

Figure 39. Cases in Which Progress Notes Were Used for Coding Change

Cases in Which Anesthesia and Operative Report Were Used for Coding Change
BMI assigned per per-anesthesia record, “Add 278.00 obesity per pre-anesthesia record”
“Add 530.81 GERD (reflux) per pre-anesthesia record”
“Delete 87.53 interoperative cholangiogram per the operative report the procedure could not be performed due to a stone lodged in the neck of the gallbladder. The DRG does not change”
“Add 571.8 fatty liver per the body of the operative report. There was difficulty removing the gallbladder due to the patient's body habitus along with the gallbladder being intrahepatic”
“Recommend deleting 04.81 (injection of anesthetic into peripheral nerve for anesthesia). The documentation by both the surgeon and the anesthesiologist list the block along with the general. If the block is done as part of the operative anesthesia, it is not separately coded. If it is done for postoperative pain control regardless of the time of the performed, then it would be separately coded. There is no documentation that is was done for anything other than interoperative anesthesia. CPT Assistant December 2012, pg.12”

Figure 40. Cases in Which Anesthesia and Operative Report Were Used for Coding Change

Cases in Which Consultation Notes Were Used for Coding Change
“Change 414.01 CAD of native coronary artery to 414.00 CAD of unspecified, native or graft vessel per consultation note 11/02 patient had previous CABG. Add V45.81 status post CABG”
“Change 593.9 unspecified disorder of kidney and ureter (renal insufficiency) to 585.9 CKD (chronic kidney disease) per consultation patient assessment 'chronic kidney disease with increased creatinine”
“Add 783.7 adult failure to thrive per consultation note 2/15”
“Add 401.9 HTN per consultation”
“Add V45.82 (s/p ptca) per consult”

Figure 41. Cases in Which Consultation Notes Were Used for Coding Change

Cases in Which Psychiatric Evaluation Was Used for Coding Change
“Add V15.81 noncompliance with medical treatment per psychiatric evaluation”
“Add V15.81 noncompliance with medical treatment per psychiatric evaluation/H/P 'she has not been taking medication correctly.' Also noted in the consultation note 2/15.

Figure 42. Cases in Which Psychiatric Evaluation Was Used for Coding Change

Cases in Which Different Sources Were Used for Coding Change

"Recommend adding 311 (depression) per the H&P and DS"

"Recommend adding 571.5 (liver cirrhosis) as documented in the H&P and DS"

"Recommend adding 272.4 (hyperlipidemia on Zocor) and 244.9 (hypothyroid on Levothyroxine) per the H&P and DS."

"Recommend adding 571.5 (liver cirrhosis) as documented in the H&P and DS"

"Recommend adding 250.00 (diabetes mellitus) and V58.67 (long term use of insulin) as documented in the H&P and DS"

"Recommend adding the following secondary codes which are documented in the H&P and DS and under current treatment: 401.9 (hypertension), 493.90 (asthma), 424.0 (mitral valve disorder), and 276.51 (dehydration)"

"Add E950.4 self-inflicting poisoning; other specified drugs and medicinal substances per ED and progress notes"

"Add 244.9 hypothyroidism, 272.4 hyperlipidemia and 285.9 anemia per ED,H/P, consultation, progress notes"

"Add 272.4 hyperlipidemia per progress notes 11/20 and operative report"

"Add E888.8 other fall and E849.0 place of occurrence, home per ED, consultation and discharge summary"

"Add 272.4 hyperlipidemia per consult, progress notes and D/S"

"Change principal diagnosis from 780.60 fever, unspecified to 079.99 unspecified viral illness per discharge summary and consultation"

"Change principal diagnosis from 296.20 major depressive disorder, single episode, unspecified to 296.24 major depressive disorder, single episode, severe, specified as w/ psychotic behavior per

discharge summary and psychiatric evaluation/H/P"

Figure 43. Cases in Which Different Sources Were Used for Coding Change

EXAMPLES II: *ERRORS RELATED TO CODING GUIDELINES*

(Figures 44-50)

Changes Related to Principle Diagnosis
"Change principal diagnosis from 574.10 cholelithiasis with other cholecystitis to 574.71 cholelithiasis w/o choledocholithiasis with other cholecystitis; with obstruction per the body of the operative report the cholangiogram could not be performed due to the stone lodged into the neck of the gallbladder. The DRG does not change" "Recommend changing the principal diagnosis from 715.91 (OA unspecified whether generalized or localized; shoulder region) to 715.31 (OA, localized, not specified whether primary or 2nd; shoulder region) The documentation in the medical record does not document OA anywhere but in the shoulders. The below referenced coding clinic guides coding into the 715.3x category. Coding Clinic 4Q 2003, pg.118" "Recommend changing 038.42 (septicemia due to E. coli) to 038.9 (unspecified septicemia). There is no physician documentation of the sepsis being due to an organism and the blood cultures are negative" "Recommend changing the principal diagnosis from 715.96 (osteoarthrosis, unspecialized whether general or localized) to 715.36 (osteoarthrosis, localized, not specified whether primary or secondary). The OA is specified only for the knee and the RA is systemic" "Recommend changing the principal diagnosis from 715.96 (osteoarthrosis, unspecified whether generalized or localized) to 715.36 (osteoarthrosis, localized) as the documentation of OA includes only the knee" "Recommend changing the principal from 715.96 (osteoarthrosis, unspecified whether generalized or localized) to 715.36 (osteoarthritis, localized) per documentation in the record. The OA is confined to the knee. "Recommend changing the principal diagnosis from 038.42 (e coli septicemia) to 038.9 (unspecified septicemia). There is no linkage of organism to the septicemia and the blood culture is negative" "Change principal diagnosis from 414.01 CAD of native vessel to 414.00 CAD; unspecified vessel, native or graft in a patient with previous CABG and PTCA. DRG does not change"

“Change 786.50 chest pain, unspecified to 786.59 atypical chest pain per discharge summary. The DRG does not change”

“Would recommend using R04.0 for the epistaxis.

Figure 44: Changes Related to Principle Diagnosis

Changes Related to Secondary Diagnoses

“Recommend adding the following secondary codes which are documented in the H&P and DS and under current treatment: 401.9 (hypertension), 493.90 (asthma), 424.0 (mitral valve disorder), and 276.51 (dehydration)”

Secondary and principle: “Would recommend using G47.33 for the OSA. Would recommend using D63.8 for the anemia in other chronic diseases”

“Delete 285.9 (unspecified anemia) and report 280.9 (iron deficiency anemia) which is more specific”

“Report 2724. (dyslipidemia) on Plavix, V15.82 (h/o tobacco use) per H&P”

“Add V58.67 Type 2 Diabetes Mellitus patient on insulin – Coding Clinic directives is to report V58.67 to show the use of insulin for these patients”

“Delete 584.9 and report 584.5 per consult as this is the more definitive code”

Figure 45: Changes Related to Secondary Diagnoses

Changes Related to Combination Codes

“Recommend deleting 997.49 (Other digestive system complications) as this code is included in 536.49”

“Recommend changing 788.42 (polyuria) to 791.9 (pyuria) per documentation. Polyuria is a symptom of the newly diagnosed diabetes mellitus and would not be separately coded.

Figure 46: Changes Related to Combination Codes

Changes Related to Symptoms & Signs

“Unless needed for medical necessity, cough and wheezing are considered symptoms of the definitive diagnosis and as such are not separately coded”

“Delete 780.97 (altered mental status) this is a symptom code and would not be reported separately. Documentation states " altered mental status is secondary to infection”. (pneumonia)

Figure 47: Changes Related to Symptoms & Signs

Changes Related to V Codes
“Recommend adding V58.66 (long term aspirin) and V58.67 (long term use of insulin). Official Coding Guidelines”
Recommend adding V58.66 (long term aspirin) and V49.86 (do not resuscitate) per H&P documentation
“Recommend adding V58.66 (long term aspirin) and V58.67 (long term insulin) per the chart documentation. Long term aspirin poses a bleeding risk and long term insulin documents progression of the diabetes”

Figure 48: Changes Related to V Codes

Changes Related to CC/MCC/POA
“Add 296.20 major depressive disorder per discharge summary. This is a CC, however, does not affect DRG”
“Add 486 pneumonia per D/S. This is a MCC; does not change DRG in this case”
“Add 290.3 senile dementia with delirium per D/S and consultation 02/15. This is a CC; does not change DRG”
“Change 401.1 hypertension to 403.90 hypertensive kidney disease per coding guidelines”
“Add V58.67 Type 2 Diabetes Mellitus patient on insulin – Coding Clinic directives is to reportV58.67 to show the use of insulin for these patients”

Figure 49: Changes Related to CC/MCC/POA

Changes Related to Place of Occurrence
Place of occurrence- Heart Attack
“Delete E849.9 place of occurrence, unspecified per Coding Guidelines below: Place of Occurrence Guideline- *Use an additional code from category E849 to indicate the Place of Occurrence for injuries and poisonings. The Place of Occurrence describes the place where the event occurred and not the patient’s activity at the time of the event.* Do not use E849.9 if the place of occurrence is not stated”.
“Change E849.0 place of occurrence, home to E849.7 place of occurrence, hospital per discharge summary on second inpatient day father struck patient and police were called”

Figure 50: Changes Related to Place of Occurrence

EXAMPLES III: *RECOMMENDED CHANGES PER DOCUMENTATION*

(Figures 51-53)

Lack of Supporting Documentation
"Recommend deleting 296.90 (episodic mood disorder) due to a lack of supporting documentation in the record" "Delete 99.77 (adhesion barrier) as this was not documented in the record as being used during this procedure" "Recommend changing 286.9 (other and unspecified coagulation defects) to 287.5 (thrombocytopenia) per documentation in the record. There was no documentation to support a coagulopathy" "Recommend changing 327.23 (obstructive sleep apnea) to 780.57 (unspecified sleep apnea) due to a lack of documentation that the sleep apnea was obstructive in nature" "It is noted that the record contained documentation from both the IP admission and the Swing Bed. None of the chronic health conditions (ex. htn) are coded. This leaves gaps in the capture of the general health of the patient and the comorbidities that may have an impact on this patient's therapy" "Recommend changing 337.21 (RSD upper limb) to 337.20 (RSD unspecified site) as the record does not document what area(s) are involved"

Figure 51: Lack of Supporting Documentation

Addition and Deletion Per Documentation
"Recommend adding 427.31 (atrial fib) and V58.61 (long term use of coumadin) per documentation in the record" "Recommend adding 428.0 (CHF) per documentation. The treatment of this patient including holding the digoxin that was prescribed for treating the CHF" "Recommend adding 585.9 (chronic renal insufficiency). There is a code also underlying renal disease instruction under code 403.90 in the ICD-9 Code Book" "Recommend adding 274.9 (gout). The patient is treated with allopurinol which can be renal toxic and was held for a bit to help facilitate resolution of the kidney failure"

“Recommend adding 600.00 (BPH) as this condition was a possible contributing source of the documented hematuria”

“Recommend adding 57.94 for the urinary catheter inserted at discharge”.

“Recommend adding V58.66 (long term use of aspirin) and V58.64 (long term use of NSAID) as these meds can help encourage bleeding. In this patient, she was encouraged to avoid these meds after discharge”

“Recommend adding 584.9 (acute renal failure) and E930.6 (adverse effect antimycobacterial antibiotic) as the patient went into acute renal failure with vancomycin charted as the cause”

“Delete 285.9 (unspecified anemia) and report 280.9 (iron deficiency anemia) which is more specific.”

Figure 52: Addition and Deletion Per Documentation

Recommended Query

“Recommend clarification regarding patient's final diagnosis as documentation is conflicting. Discharge summary does not state patient with Alzheimer's and consultation states dementia and Alzheimer's Disease”

“On admission, the patient's hemoglobin was 10.6 (5/9) and fell to 7.9 on 5/10. The patient received two units of packed cells. The patient returned to surgery for postoperative oozing from the liver bed. Would query for acute postoperative blood loss anemia”

Query for conflicting documentation: “Recommend clarification regarding patient's final diagnosis as documentation is conflicting. Discharge summary does not state patient with Alzheimer's and consultation states dementia and Alzheimer's Disease.

Figure 53: Recommended Query

BIBLIOGRAPHY

AHIMA. "Taking Coding to the Next Level through Clinical Validation." *Journal of AHIMA* 88, no.1 (January 2014): expanded web version.

AHIMA & AMIA Terminology and Classification Task Force. (2007). Healthcare terminologies and classification: An action agenda for the United States.

AHIMA “Working Smarter with Computer-Assisted Coding." AHIMA Today, the Convention Daily Newsletter, October 5, 2009. Brouch K. AHIMA Project offers insights into SNOMED, ICD-9-CM mapping process. J AHIMA 2003:74:52-5.

AHIMA e-HIM^{TM} Work Group on Computer-Assisted Coding. "Delving into Computer-assisted Coding" (AHIMA Practice Brief). *Journal of AHIMA* 75, no.10 (Nov-Dec 2004): 48A-H (with web extras).

AHIMA. "Automated Coding Workflow and CAC Practice Guidance." *Journal of AHIMA* 81, no.7 (July 2010): 51-56.

AHIMA. & AMIA, (2007). Healthcare terminologies and classification: Essential keys to interoperability.

AHIMA. “Automated Coding Workflow and CAC Practice Guidance (Update).” *Journal of AHIMA*, 84, no.11 (November-December 2013): expanded web version.

AHIMA Practice Brief. “Automated Coding Workflow and CAC Practice Guidance (2013 Update).” Journal of AHIMA 84, no. 11 (2013). Available at http://library.ahima.org/xpedio/groups/public/documents/ahima/bok1_050535.hcsp?dDocName=bok1_050535.

AHIMA. “Data Quality Management Model (2012 Update).” Journal of AHIMA83, no. 7 (July 2012): 62–67. Available at http://library.ahima.org/xpedio/groups/public/documents/ahima/bok1_049664.hcsp?dDocName=bok1_049664.

American Health Information Management Association. A Coder’s Guide to Natural Language Processing and Coded Data. Chicago: AHIMA, 2011.

Alakrawi, Zahraa M; Watzlaf, Valerie J.M.; Nemchik, Scot; Sheridan, Patty Thierry. "New Study Illuminates the Ongoing Road to ICD-10 Productivity and Optimization" Journal of AHIMA 88, no.3 (March 2017): 40-45.

American Health Information Management Association. CAC 2010: Industry Outlook and Resources Report. Chicago: AHIMA, 2011.

Arner, Kathy. "When Guidelines Depend on the Setting: Comparing, Contrasting Facility Reporting and Professional Fee Coding." *Journal of AHIMA* 78, no.2 (February 2007): 70-71.

Averill, R., Bowman, S., & Bowman, S. (2012). There are critical reasons for not further delaying the implementation of the new icd-10 coding system. *Journal of AHIMA*, *83*(7).

Beinborn, Julie. "Automated Coding: The Next Step." Journal of AHIMA 70, no. 7 (1999): 38–43.

Benjamin P. Rosenbaum, MD; Robert R. Lorenz, MD, MBA; Ralph B. Luther, MBA; Lisa Knowles-Ward, RHIT, CCS; Dianne L. Kelly, RN; and Robert J. Weil, MD. "Improving and Measuring Inpatient Documentation of Medical Care within the MS-DRG System: Education, Monitoring, and Normalized Case Mix Index." *Perspectives in Health Information Management* (Summer 2014): 1-11.

Bhargava, B., Temkin, T., Fireman, B., Eaton, A., McCaw, B., Kotz, K., & Amaral, D. (2011). A predictive model to help identify intimate partner violence based on diagnoses and phone calls. *Am J Prev Med*, *41*(2), 129-135.

Bower-Jernigan, Patricia; Chenoweth, Ann; James, Jaime. "Realigning HIM to the New Healthcare Environment: Case Studies in HIM Transformation Due to Accountable Care and Pay-For-Outcomes Initiatives." Journal of AHIMA 85, no.8 (August 2014): 34-38.

Brownfield, Cathy; Didier, Donna M. "Making the Most of External Coding Audits: From Preparation to Recommendations " Journal of AHIMA 80, no.7 (July 2009): 34-38.

Boyd, A., John, J., Burton, M., Jonen, M., Gardeux, V., Achour, I., Q Luo, R., Zenku, I., Bahroos, N., Brown, S., Vanden Hoek, T., Lussier, Y. The discriminatory cost of ICD-10-CM transition between clinical specialties: metrics, case study, and mitigating tools *Journal of the American Medical Informatics Association*: amiajnl-2013-001358.

Broberg, C., McLarry, J., Mitchell, J., Winter, C., Doberne, J., Woods, P., & Burchill, L. (2014). Accuracy of administrative data for detection and categorization of adult congenital heart disease patients from an electronic medical record. *Pediatr Cardiol*, 1-7.

Bronnert, J., Bonnie , C., Eichenwald, S., Eminger, H., & Flanagan, J. (2011). Cac 2010-11 industry outlook and resources report. *AHIMA*.

Brouch, Kathy. "AHIMA Project Offers Insights into SNOMED, ICD-9-CM Mapping Process." Journal of AHIMA 74, no.7 (July/August 2003): 52-55.

Burton, Ben. "How to Lead Your Organization in Compliance, Ethics, and Customer Service." Journal of AHIMA 85, no.8 (August 2014): 22-25.

Butler, Mary. "Not So Fast! Congress Delays ICD-10-CM/PCS: Examining How the Delay Happened, Its Industry Impact, and How Best to Proceed." Journal of AHIMA 85, no.6 (June 2014): 24-28.

Butler, Mary. "Survey: Coding Productivity Dipped After ICD-10 Implementation" (Journal of AHIMA website), June 13, 2016.

Chapman, W., Nadkarni, P., Hirschman, L., D'Avolio, L., Savova, G., & Uzuner, O. (2011). Overcoming barriers to nlp for clinical text: the role of shared tasks and the need for additional creative solutions. *J Am Med Inform Assoc.*, *18*(5), 540-543.

Chavis, S. "Two Systems, One Direction." For the Record 25, no. 14 (2013): 10. Available at http://www.fortherecordmag.com/archives/1013p10.shtml.

Combs, Tammy. "Understanding CDI Metrics" (Journal of AHIMA website), February 24, 2016.

Cornet, R., and N. Keizer. "Forty Years of SNOMED: A Literature Review."BMC Medical Informatics and Decision Making 8 (2008): 1–6.

Chute, C. (2000). J am med inform assoc. *Clinical Classification and Terminology: Some History and Current Observations*, *7*, 298-303.

Chute, C. (2013). State of the art and science: Coding patient reimbursement for care, and the icd transition. *Virtual Mentor; American Medical Association Journal of Ethics*, *15*, 596-599.

Chute, C., Huff, S., Ferguson, J., Walker, J., & Halamka, J. There Are Important Reasons For Delaying Implementation Of The New ICD-10 Coding System. *Health Affairs, 31*, (2012): 836-842.

CMS. (n.d.). Icd-9 code lookup. Retrieved from https://www.cms.gov/

Cockbain, A., Carolan, M., Berridge, D., & Toogood, G. (2012). Performance and quality indicators: the importance of accurate coding. Ann R Coll Surg Engl, 94, 46-50.

Cornet, R., & Keizer, N. (2008). Forty years of SNOMED: a literature review. *BMC Medical Informatics and Decision Making*, (8), 1-6.

DeAlmeida, D., Watzlaf, V., Anania-Firouzan, P., Salguero, O., Rubinstein, E., Abdelhak, M., Parmanto, B., Evaluation of Inpatient Clinical Documentation Readiness for ICD-10-CM, *Perspectives in Health Information Management*. 2014 11(Winter): 1h. Published online Jan 1, 2014.

DeVault, Kathryn. “Best Practices for Coding Productivity: Assessing Productivity in ICD-9 to Prepare for ICD-10.” Journal of AHIMA 83, no. 7 (July 2012): 72-74.

Desai, Anna. "Focus on the Future: Environmental Scan Illuminates the Path Ahead for HIM."*Journal of AHIMA* 84, no.8 (August 2013): 48-52.

Dimick, Chris. "Achieving Coding Consistency" Journal of AHIMA 81, no.7 (July 2010): 24-28.

Dimick, Chris. "Treating Healthcare with Health "I"T." *Journal of AHIMA* 83, no.9 (September 2012): 30-37.

Drake, Suzanne P. "Coding and Quality Measures Brief: Pressure Ulcer Rate" CodeWrite (December 2016).

Drake, Suzanne P. "Coding and Quality Measures Brief. Patient Safety Indicator 11: Postoperative Respiratory Failure Rate" CodeWrite (January 2017).

Dunn, Rose T. and Christina Mainord. "The Latest Look at Coding Trends." Journal of AHIMA 72, no.7 (2001): 94-96.

Dunn, Rose. "Developing Facility-specific Productivity Measures." Journal of AHIMA 72, no.4 (2001): 73-74.

Duarte, J., S. Castro, M. Santos, A. Abelha, and J. Machado. “Improving Quality of Electronic Health Records with SNOMED.” Procedia Technology 16 (2014): 1342–50.

e-HIM Work Group on Benchmark Standards for Clinical Coding Performance Measurement. Quality Subgroup. "Collecting Root Cause to Improve Coding Quality Measurement." Journal of AHIMA 79, no.3 (March 2008): 71-75.

Endicott, Melanie. "Maintaining Productivity During the Transition to ICD-10" (Journal of AHIMA website), July 08, 2015.

Endicott, Melanie. "Effects of ICD-10 on Coding Productivity" (Journal of AHIMA website), March 10, 2016.

Eramo, Lisa A. "Healthcare’s Data Revolution: How Data is Changing the Industry and Reshaping HIM’s Roles." *Journal of AHIMA* 84, no.9 (Sept 2013): 26-32.

Farishta, Mehnaz. "More than a Database: Mining Your Data for Decision-Making Success." *Journal of AHIMA* 72, no.10 (2001): 28-32.

Farley, D. (2012). Code theory: the impact of icd-10 on predictive modeling. *Predictive Modeling News*, *5*(8).

Flanagan, J., & Santos, M. (2009). Defining the standards for automated e&m coding through coding consistency methodology. *Perspectives in Health Information Management*.

Forman, S. (2013). Compliance program leveraging of audit resources. Journal of Health Care Compliance, 55-70.

Friedman, C., Rindflesch, T., & Corn, M. (2013). Natural language processing: State of the art and prospects for significant progress, a workshop sponsored by the national library of medicine. *Journal of Biomedical Informatics*, (46), 765–773.

Friesner, D., Roseman, R., & McPherson, M. (2007). Does a single case mix index fill all hospitals? empirical evidence from washington state .*Research in Healthcare Financial Managment*,*11*(1), 35-55.

Giannangelo, K. *Healthcare code sets, clinical terminologies, and classification systems*. (2nd ed.). Chicago: American Health Information Management Association (AHIMA).

Gjertsen, F., Bruzzone, S., Vollrath, M., Pace, M., & Ekeberg, O. (2013). Comparing icd-9 and icd-10: The impact on intentional and unintentional injury mortality statistics in Italy and Norway. 44, 132-138. doi: Injury. Int. J. Care Injured.

Gleberman, Amy L. "Incentive Programs Offer Aid to Increase Coding Productivity." Journal of AHIMA 71, no. 1 (2000): 75-77.

Glenn, M. "Synergizing ICD-10: Integrating an ICD-10 Implementation into Other Compliance Programs Will Reduce Costs, Maximize Investments." Journal of AHIMA84, no. 2 (March 2013): 34–38.

Gøeg, K. R., R. Chen, A. R. Højen, and P. Elberg. "Content Analysis of Physical Examination Templates in Electronic Health Records Using SNOMED CT." International Journal of Medical Informatics 83, no. 10 (2014): 736–49.

Godbey-Miller, Barbara. "ICD-10: How to Move Your Coding Professionals to the Head of the Class [Sponsored Article]" (Journal of AHIMA website), January 2016.

Grzybowiski, D. (2012). Computer-assisted coding and electronic health records EHR. *Electronic Health Records Briefing*, 7-10.

Gurrieri, Joseph J; Karban, Karen M. "The Good, Bad, and Reality of Offshore Coding: Some Turn to Distant Shores to Fill US Coding Demands ." *Journal of AHIMA* 84, no.9 (Sept 2013): 44-48.

Hayrinen, K., Saranto, K., & Nykanen, P. (2008). Definitions, structure, content, use and impacts of electronic health records: A review of the research literature. *International Journal of Medical Informatics* , (77), 291-304.

Hevnegaard, A., Street, A., Sorensen, T., & Gyrd-Hansen, D. (2009). Comparing hospital costs: What is gained by accounting for more than a case-mix index?. *Social Science and Medicine*, (69), 640-647.

He, D., Mathews, S., Kallo, A., & Hutfless, S. (2013). Mining high-dimensional administrative claims data to predict early hospital readmissions. *J Am Med Inform Assoc*, (0), 1-8.

Hill, C., Hunter, A., Jobnson, L., & Coustasse, A. (2014). Medicare fraud in the united states: can it ever be stopped?. The Health Care Manager, 33(3), 254-260.

Hohl, Corinne M., et al. ICD-10 codes used to identify adverse drug events in administrative data: a systematic review. *Journal of the American Medical Informatics Association* (2013): amiajnl-2013.

Houser, Shannon H.; Meadows, Jean M. "Transitioning to Electronic Clinical Quality Measures in the Informatics Era" Journal of AHIMA 88, no.2 (February 2017): 28-31.

Humphreys, B., Lindburg, D., Schoolman, H., & Barnett, O. (1998). The unified medical language system: An informatics research collaboration. *Journal of the American Medical Informatics Association*, *5*(1), 1-11.

International Health Standards Development Organisation. SNOMED CT Starter Guide. 2014. Available at http://ihtsdo.org/fileadmin/user_upload/doc/download/doc_StarterGuide_Current-en-US_INT_20140222.pdf.

International Health Standards Development Organisation. SNOMED CT® Frequently Asked Questions. Available at http://ihtsdo.org/fileadmin/user_upload/doc/download/doc_FAQ_Current-en-US_INT_20130731.pdf.

Inza, L., Calvo, B., Armananzas, R., Bengoetxea, E., Larranaga, P., & Lozano, J. (2010). Machine learning: An indispensable tool in bioinformatics. In L. Inza (Ed.), *Bioinformatics Methods in Clinical Research* (pp. 25-48).

Jiang, G., Jiang, G., Pathak, J., & Chute, C. (2009). Formalizing ICD coding rules using formal concept analysis. *Journal of Biomedical Informatics*, *42*, 504-517.

Jensen, K., Cooke, C., & Davis, M. (2014). Fidelity of administrative data when researching down syndrome. *Medical Care*, *52*(8), 52-57.

Kate, R. (2013). Towards converting clinical phrases into SNOMED CT expression. *Biomedical Informatics Insights*, *6*(1), 29-37.

Koutroumanos, J. (2010). Comparing coding between interventional radiologists and hospital coding departments. Clinical Audit, 2, 33-36.

Kusserow, R. (2014). Dissecting the internal audit function in health care. Journal of Health Care Compliance, 47-50.

Land, Daniel. "Auditing ICD-10 Through the Lens of Education" Journal of AHIMA 87, no.10 (October 2016): 82-83.

Lasserre, J., Arnold, S., Vingron, M., Reinke, P., & Hinrichs, C. (2012). Predicting the outcome of renal transplantation. *J Am Med Inform Assoc.*, *19*, 255-262.

Lau, F., Kuziemsky, C., Price, M., & Gardner, J. (2010). A review on systematic reviews of health information system studies. *J Am Med Inform Assoc.*, (17), 637-645.

Leon-Chisen, Nelly. "Coding and Quality Reporting: Resolving the Discrepancies, Finding Opportunities" Journal of AHIMA 78, no.7 (July 2007): 26-30.

Linder, Karen M. "Finding the New Normal with ICD-10-CM/PCS" Journal of AHIMA 87, no.4 (April 2016): 56-59.

Lorence, Daniel P. "Productivity: How Do You Measure Up?" Journal of AHIMA 70, no. 5 (1999): 35-39.

Madden, R., Sykes, C., & Ustun, T. (2008.). World health organization family of international classifications: Definition, scope and purpose.

Mahjan, R., Moorman, A., Liu, S., & Rupp, L. (2013). Use of the international classification of diseases, 9th revision, coding in identifying chronic hepatitis b virus infection in health system data: implications for national surveillance. *J Am Med Inform Assoc.*, (20), 441-445.

Martin, Ginny. "Using Additional Diagnoses to Improve Productivity" CodeWrite (February 2016).

Mathias, J., Agrawal, A., Feinglass, J., Cooper, A., Baker, D., & Choudhary, A. (2013). Development of a 5 year life expectancy index in older adults using predictive mining of electronic health record data. *J Am Med Inform Assoc.*, *20*, e118–e124.

McBride, Susan, et al.. "Data Mapping." *Journal of AHIMA* 77, no.2 (February 2006): 44-48.

Mea, D., et al. “A Web-based Tool for Development of a Common Ontology between ICD11 and SNOMED-CT.” 2014 IEEE International Conference on Healthcare Informatics (2014): 144–48.

Melton, G., & Hripcsak, G. (2005). Automated detection of adverse events using natural language processing of discharge summaries. *J Am Med Inform Assoc.*,*12*, 448–457.

Mendez, C., Harrington, D., Christenson, P., & Spellberg, B. (2014). Impact of hospital variables on case mix index as a marker of disease severity. *Population Health Management*, *17*(1), 28-34.

Mikroyannidi, E., Stevens , R., Lannone, L., & Rector, A. (2012). Analyzing syntactic regularities and irregularities in SNOMED-CT. *Journal of Biomedical Semantics*, *3*(8).

Miller, P.J., and F.L. Waterstraat. "Apples to Apples: Using Autobenchmarking to Measure Productivity." Journal of AHIMA 75, no.1 (January 2004): 44-49.

Mitchell, Glenn. "Synergizing ICD-10: Integrating an ICD-10 implementation into other compliance programs will reduce costs, maximize investments." *Journal of AHIMA* 84, no.2 (March 2013): 34-38.

Mitra, I., Malik, T., Homer, J., & Loughran, S. (2009). Audit of clinical coding of major head and neck operations. Ann R Coll Surg Engl, 91, 245-248.

Moar, K., & Rogers, S. (2012). Impact of coding errors on departmental income: an audit of coding of microvascular free tissue transfer cases using opcs-4 in uk . British Journal of Oral and Maxillofacial Surgery, 50, 85-87.

Morsch, M. (2010). Computer-assisted coding: The secret weapon. *Health Management Technology*, 24-26. Chute, C. (2000). J am med inform assoc. *Clinical Classification and Terminology: Some History and Current Observations*, *7*, 298-303.

Morsch, Mark L., David S. Byrd, and Daniel T. Heinze. “Factors in Deploying Automated Tools for Clinical Abstraction and Coding.” Unpublished work. 2007.

Nadkarni, P., & Darer, J. (2010). Migrating existing clinical content from icd-9 to snomed. *J Am Med Inform Assoc.*, (17), 602-607.

Nadkarni, P., Ohno-Machado, L., & Chapman, W. (2011). Natural language processing: an introduction. *J Am Med Inform Assoc.*, *18*, 544-551.

Naran, S., Hudovsky, A., Antscherl, J., Howells, S., & Nouraei, S. (2014). Audit of accuracy of clinical coding in oral surgery. British Journal of Oral and Maxillofacial Surgery, 52, 735-739.

Nicole Blair Johnson, Locola D. Hayes, Kathryn Brown, Elizabeth C. Hoo. & Kathleen A. Ethier. (2014). CDC National Health Report: Leading Causes of Morbidity and Mortality and Associated Behavioral Risk and Protective Factors—United States, 2005–2013. CDC. Retrieved form: http://www.cdc.gov/mmwr/preview/mmwrhtml/su6304a2.htm

Nichols, Cindy. "Measuring True Productivity." 2011 AHIMA Convention Proceedings, October 2011.

Nouraei, S., Hudovsky, A., Virk, J., Chatrath, P., & Sandhu, G. (2013). An audit of the nature and impact of clinical coding subjectivity variability and error in otolaryngology. Clinical Otolaryngology, 38, 512-524.

Osborn, Carol E. "Practices and Productivity in Acute Care Facilities." Journal of AHIMA 71, no. 2 (2000): 61-66.

Pitsikoulis, John; Doty, Laura. "Developing Productivity Standards through Analytics [sponsored article]" (Journal of AHIMA website), July 15, 2016.

Price, Kurt, and Dean Farley. "How Does Your Coding Measure Up?: Analyzing Performance Data Gives HIM a Boost in Managing Revenue." *Journal of AHIMA* 76, no.7 (July-August 2005): 26-31.

Rea, S., Pathak, J., Savova, G., Oniki, T., Westberg, L., Beebe, C., Tao, C., & Parker, C. (2012). Building a robust, scalable and standards-driven infrastructure for secondary use of ehr data: The sharpn project. *Journal of Biomedical Informatics*, (45), 763-771.

Reed, G. (2010). Toward icd-11: Improving the clinical utility of WHO's international classification of mental disorders. Professional Psychology: Research and Practice, 41(6), 457-464.

Rode, D. Embracing ICD-10 Advocacy: HIM Professionals Must Offer Implementation Help to Physicians. *Journal of AHIMA* 84, no.2; (2013, Feb):16-18.

Rode, D. Why ICD-10 Can't Wait. *Journal of AHIMA*, 79, 10 (2008): (18 - 20).

Rode, Dan. "Taking the Next Step Forward for ICD-10." *Journal of AHIMA* 75, no.1 (January 2004): 14-15.

Rode, Dan. "Why ICD-10 Can't Wait" *Journal of AHIMA* 79, no.10 (October 2008): 18-20. Romano, PS. ICD-9-CM to ICD-10-CM/PCS Conversion of AHRQ Quality Indicators.

Rybnicek, D., Hanthorn, K., Pfaff, E., Bulsiewicz, W., Shaheen, N., & Dellon, E. (2014). Administrative coding is specific, but not sensitive, for identifying eosinophilic esophagitis. Diseases of the Esophagus, (27), 703-708.

Salmasian, H., Freedberg, D., & Friedman, C. (2013). Driving comorbidities form medical records using natural language processing. (0), 1-4.

Sacks, G., Dawes, J., Russell, M., Lin, A., Maggard-Gibbons, M., Winogard, D., Chung, H., & Tomlinson, J. (2014). Evaluation of hospital readmissions in surgical patients: Do administrative data tell the real story?. *JAMA Surg*, *149*(12), 759-764.

Sett, Ajit; Hickman, George T; Karban, Karen. "Trust But Verify: Safeguards in Contracting for Outsourced Coding Services." Journal of AHIMA 85, no.6 (June 2014): 40-44.

Sion, D., Dooling, J., Glondys, B., Jones, T,.Kadlec, L., Overgaard, S., Ruben, K., & Wendicke, A. “Data Quality Management Model (2015 Update)” Journal of AHIMA 86, no.10 (October 2015): expanded web version.

Spencer, Carol. "Two Key Metrics in Evaluating the Effectiveness of Hospital Coding Services: Paid Claims Error Rate and Coding Accuracy." 2010 AHIMA Convention Proceedings, September 2010.

Stanfill, M., Williams, M., Fenton, S., Jenders, R., & Hersh, W. (2010). A systematic literature review of automated clinical coding and classification systems. *J Am Med Inform Assoc.*, *17*, 646-651.

Stanfill, Mary H; Hsieh, Kang Lin; Beal, Kathleen; Fenton, Susan H. "Preparing for ICD-10-CM/PCS Implementation: Impact on Productivity and Quality" Perspectives in Health Information Management (Summer, July 2014).

Stanfill, Mary H. "Call for Additional Coding Metrics" Journal of AHIMA 86, no.4 (April 2015): 56-57.

Stausberg, J., Lehmann, N., Kaczmarek, D., & Stein, M. (2008). Reliability of diagnoses coding with icd-10. *International Journal of Medical Informatics* ,*77*, 50-57.

Steinbusch, P., Joost, O., Zuurbier, J., & Schaepkens, F. (2007). The risk of upcoding in casemix systems: A comparative study. *Health Policy*, (81), 289-299.

Steindel, S. International classification of diseases, 10th edition, clinical modification and procedure coding system: descriptive overview of the next generation HIPAA code sets

Journal of the American Medical Informatics Association. May-Jun; 17(3): (2010):274–282.

Stoner, Jean, Michael Nossal, Philip Resnik, Andrew Kapit, and Richard Toren. "Assessing Coder Change Rates as an Evaluation Metric." Unpublished work. 2007.

Sun, J., McNaughton, C., Sun, J., Zhang, P., Perer, A., Gkoulalas-Divanis, A., Denny, J., & Kirby, J. (2013). Predicting changes in hypertension control using electronic health records from a chronic disease management program. *J Am Med Inform Assoc.*, (0), 1-8.

Tully , M., & Charmichael, A. (2012). Computer-assisted coding and clinical documentation. *Healthcare Financial Management* , 46-49.

Walker, R., Hennessy, D., Johansen, H., Sambell, C., Lix, L., & Quan, H. (2012). Implementing of ICD-10 in Canada: how has it impacted coded hospital discharge data. BMC Health Services Research, 12(149).

Watzlaf, V., Garvin, J. H., Moeini, S., & Anania-Firouzan, P. The Effectiveness of ICD-10-CM/PCS in Capturing Public Health Diseases. *Perspectives in Health Information Management*, 4, (2007, Jun 12): 6.

Watzlaf, Valerie J.M.; Nemchik, Scot; Hoerner, Mike; Sheridan, Patty Thierry. "ICD-10 Coding Productivity Study Highlights Emerging Standards" Journal of AHIMA 87, no.8 (August 2016): 44-47.

W. C. Morris et al. "Assessing the Accuracy of an Automated Coding System in Emergency Medicine." Proceedings of the AMIA Annual Symposium (2000): 595–599.

Weinstein, L., Jack, T., Radano, T., Kalina , P., & Eberhardt , J. (2009). Application of multivariate probabilistic (bayesian) networks to substance use disorder risk stratification and cost estimation. *Perspectives in Health Information Management*.

White, Susan E. "Predictive Modeling 101: How CMS's Newest Fraud Prevention Tool Works and What It Means for Providers." *Journal of AHIMA* 82, no.9 (September 2011): 46-47.

Wilson, Donna D; Hampton-Bagshaw, Kim; Jorwic, Therese M; Bishop, Jean; Giustina, Elizabeth. "New Focus on Process and Measure: Raising Data Quality with a Standard Coding Workflow and Benchmarks" Journal of AHIMA 79, no.3 (March 2008): 54-58.

Wilson, Donna, and Rose Dunn, technical editor. *Benchmarking to Improve Coding Accuracy and Productivity*. Chicago, IL: AHIMA, 2009.

Wilson, Donna D. "Five RAC Coding Targets: Demonstration Program Identified Key Areas of Improper Payment" *Journal of AHIMA* 80, no.5 (May 2009): 64-66.

WHO. (2012). ICD-10 Browser. Retrieved from http://apps.who.int/classifications/icd10/browse/2010/en

WHO. (2012). ICD-11 Browser . Retrieved from http://apps.who.int/classifications/icd11/browse/f/en

Yang, C., & Reinke, W. (2006). Feasibility and validity of international classification of diseases based case mix indices. *BMC Health Services Researh*,*6*(125), 1-10.

Ying, T., Sun, X., Nunez, C., & Johannes, R. (2013). Using electronic health record data to develop inpatient mortality predictive model: Acute laboratory risk of mortality score (alarms). *J Am Med Inform Assoc.*, (0), 1-9.