EVALUATION OF IMAGING SERVICES INFORMATION QUALITY IN A COMPUTER-BASED PATIENT RECORD AND ASSOCIATED CLINICAL DATA MART

by

David Scott Memel

A thesis submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Medical Informatics

The University of Utah

December 2001

THE UNIVERSITY OF UTAH GRADUATE SCHOOL

SUPERVISORY COMMITTEE APPROVAL

of a thesis submitted by

David Scott Memel

This thesis has been read by each member of the following supervisory committee, and by majority vote has been found to be satisfactory.

Chair: Reed M. Gardner

Peter J. Haug

R. Scott Evans

THE UNIVERSITY OF UTAH GRADUATE SCHOOL

FINAL READING APPROVAL

To the Graduate Council of the University of Utah:

I have read the thesis of David Scott Memel in its final form and have found that (1) its format, citations, and bibliographic style are consistent and acceptable; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the supervisory committee and is ready for submission to The Graduate School.

Date

Reed M. Gardner
Chair, Supervisory Committee

Approved for the Major Department

Reed M. Gardner
Chair/Dean

Approved for the Graduate Council

David S. Chapman
Dean of The Graduate School

ABSTRACT

The past quarter century has seen increasing demands on the healthcare industry to manage and improve access to care, manage and improve processes and outcomes of care, and manage and decrease the costs of care. Forces driving the changes in healthcare also are driving changes in the information requirements of healthcare organizations and the way in which they must manage and use that information. One of the keys to more effectively managing information is measuring, improving, and maintaining the quality of information.

The broad objective of this thesis research was to evaluate the quality of imaging services data in the HELP System and the Imaging Services Data Warehouse at LDS Hospital. The dimensions of quality evaluated were accuracy, consistency, and completeness. The study design was a retrospective correlational evaluation study. Correlational evaluation studies explore the relationship between a set of variables that is measured but not manipulated in any way and is designed to facilitate the development and refinement of information resources. The independent variables of the research performed for this thesis were the processes for the production, capture, storage, and utilization of imaging services data. The dependent variable was the quality of those data. Multiple data element pairs, for each case in the inpatient and outpatient study population, were evaluated for quality problems. Quantitative analyses were used to determine the magnitude of the information quality problems. Qualitative analyses were used to identify the sources and potential impacts of the information quality problems.

For the range of data element pairs evaluated, in the inpatient population, inaccurate information was present in 22.5% to 63.0% of cases, inconsistent information was present in 16.0% to 40.7% of cases, and incomplete information was present in 1.2% to 40.7% of cases. In the outpatient population, inaccurate information was present in 36.5% to 46.2%, inconsistent information was present in 3.8% to 40.4%, and incomplete information was present in 0.0% to 23.1%. The quantitative results confirm the presence of information quality problems, and the qualitative results demonstrate the potential impacts of poor information quality on the delivery of healthcare services.

This thesis is dedicated to all those people who have supported and encouraged me to hang in there and bring it to completion. I would particularly like to thank Heintje, who has been there for 19 years through thick and thin.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ACRONYMS AND ABBREVIATIONS

BE	Barium Enema
CEO	Chief Executive Officer
CFO	Chief Financial Officer
CIO	Chief Information Officer
CMO	Chief Medical Officer
CPR	Computer-Based Patient Record
CPRS	Computer-Based Patient Record System
CPT-4	Current Procedural Terminology, 4th Edition
CQI	Continuous Quality Improvement
CT	Computer Tomography
CXR	Chest X-Ray
DC	Data Class
DQ	Data Quality
DRG	Diagnosis Related Group
ED	Emergency Department
EDW	Enterprise Data Warehouse
HCFA	Health Care Financing Administration
HEDIS	Health Plan Employer Data and Information Set
HELP	Health Evaluation through Logical Processing
HIPAA	Health Insurance Portability and Accountability Act
HIS	Health Information Systems
ICD-9-CM	International Classification of Disease, 9th Revision, Clinical Modification
IHC	Intermountain Health Care
IOM	Institute of Medicine
IQ	Information Quality
ISDW	Imaging Services Data Warehouse
ISS	Imaging Services Scheduler
JCAHO	Joint Commission on Accreditation of Healthcare Organizations
LAN	Local Area Network
MD	Medical Doctor
MIT	Massachusetts Institute of Technology
MRI	Magnetic Resonance Imaging
NCQA	National Committee for Quality Assurance
NM	Nuclear Medicine
OASIS	Outcome and Assessment Information Set
ODBC	Open Database Connectivity
ODS	Operational Data Store

OIG	Office of Inspector General
PAS	Patient Account Services
RIS	Radiology Information System
SNOMED	Systematized Nomenclature of Human Medicine
SQL	Structured Query Language
TDQM	Total Data Quality Management
TQdM	Total Quality data Management
TQM	Total Quality Management
US	Ultrasound
WAN	Wide Area Network

CHAPTER I

BACKGROUND AND PURPOSE

The Changing Healthcare Environment

Healthcare delivery in the United States has been undergoing continuous change since the beginning of this century. The changes have included growth and sophistication in the knowledge, tools, and technology supporting clinical care; evolution of healthcare delivery organizations into integrated delivery systems; changes in the methods of payment for healthcare services; and recognition of healthcare as an "industry."[1] The healthcare industry has been, and continues to be, impacted by a wide range of economic, social, technological, biological, political, and regulatory forces. These forces, acting alone and in concert, are continuing to drive and accelerate the changes that are occurring in healthcare delivery in the United States. In recent years, a number of major trends have emerged as key drivers in the accelerating pace of change. These trends include the following: market-driven healthcare, rising consumerism, mass customization of medicine, internal and external restructuring of healthcare delivery systems and payment methods, increasing sophistication in information technology, and the "digitization" of information.[2-9]

The factors that have led to the development of these trends are numerous. Among these factors are changes in the demographics of people seeking healthcare services and the epidemiology of the reasons they are seeking care. One of the major demographic changes is the general aging of the American population. Primary epidemiological changes include the dramatic rise of infectious diseases such as acquired immune deficiency syndrome. Factors related to the clinical aspects of care include the recognition of "preventable" medical errors,[10-13] studies documenting marked variation in care and "unnecessary" care,[14-17] an emphasis on wellness, a transition to evidence-based medicine, and trends toward self-care and shared decision-making.[18-23] Emerging regulations and quality standards (e.g., Joint Commission on Accreditation of Healthcare Organization's (JCAHO) Oryx, Health Care Financing Administration's (HCFA) Outcome Assessment Information Set (OASIS), National Committee for Quality Assurance's Health Plan Employer Data and Information Set (HEDIS), and provider profiling or "report cards") also play a role in the development of the new trends in healthcare delivery. Economic factors include changes in entitlements, changes in delivery models, and changes in payment methods.[1,24-26] It is this range of factors and the trends associated with them, that are causing the healthcare industry to experience increasing demands for managing and improving access to care, managing and improving processes and outcomes (clinical, functional, financial, and satisfaction) of care, and managing and decreasing the costs of care. In other words, there is an increasing demand for total accountability.

Meeting the demand for total accountability by the healthcare industry is a significant challenge and will require "cross-cultural" understanding and cooperation between healthcare consumers and clinical, administrative, and financial healthcare workers. Meeting the demand for total accountability also will require well-designed, efficient, and integrated clinical, administrative, and financial processes and the ability to make "informed" clinical, administrative, and financial decisions. The key to designing effective and efficient processes and to making sound decisions is the availability of high quality, integrated information delivered when and where it is needed, in a manner useful to knowledge workers, decision-makers, and healthcare consumers.[3,4,27,28]

Proactively managing the voluminous amount of clinical, administrative, and financial information produced and used in healthcare will enable more effective and efficient delivery of healthcare services. The greater effectiveness and efficiency will be experienced by healthcare consumers through the growing practice of evidence-based medicine, consumer empowerment through access to their own information, and improvement in the processes of healthcare delivery (e.g., scheduling, registration, patient care activities, managed care, billing). Other benefits will include improved support of health services research, consumer education and self-care, cost- and clinical-effectiveness assessments, and process and outcomes improvement. Risk management, regulatory and accreditation requirements, contracting needs, data-driven policy development and evaluation, and strategic planning also will be facilitated through improved information management.

To manage proactively, and more effectively, information in healthcare, the industry needs to evolve from the traditional focus on information systems management to a more comprehensive focus on the management of the processes, people, and systems that provide the framework for producing, capturing, storing, maintaining, integrating, and delivering information as a product throughout healthcare organizations.[3,4,27,29,30]

Data, Information, and Knowledge

Understanding the difference between data, information, and knowledge will help provide a framework for discussing the concept of the information value chain, the principles of information management, and the field of information quality (IQ).

The terms "data," "information," and "knowledge" often are used interchangeably. However, there is a conceptual difference between these three terms.[31,32] Data are a representation of facts about things or entities in the real world. Examples of data are "38, productive, and infiltrate." Without context, these individual pieces of data have no meaning. When data are put into context, they acquire meaning and become information. Using the three pieces of data above as an example and putting them into context by qualifying them further - "Mr. Jones has a temperature of 38° centigrade, a productive cough, and an infiltrate on chest X-ray (CXR)" - these data become clinical information about a particular patient. The application of information in a specific context becomes knowledge. Through experience, people learn to understand the significance of information and use it to make informed decisions. These decisions often lead to an action that is taken with the goal of having a positive impact on a person or situation. Continuing with the example above, a physician would use the information about Mr. Jones' temperature, cough, and chest x-ray and apply their experience to draw the conclusion that Mr. Jones has pneumonia. The physician then would use the knowledge that Mr. Jones has pneumonia to make a decision to treat him with antibiotics, the desired impact of the treatment being the resolution of Mr. Jones' illness. Figure 1 graphically depicts the steps in the transformation from data to information to knowledge.

Depending on the context in which the terms data, information, and knowledge are being used, the differences described above may or may not be pertinent. Some authors do not believe it is practical to differentiate between these three terms in daily usage.[32,33] I believe differentiating between them is beneficial in some situations, such as the discussion of the information value chain. However, for the sake of simplicity, the terms data and information will be used interchangeably in this thesis.

The Information Value Chain

Data, information, and knowledge have become increasingly well recognized as reusable, strategic resources for industries ranging from banking to healthcare. The information value chain as defined by English is:

> [T]he entire collection of processes and computer applications that create, update, extract, interface, transform data, and present information from its original inception or knowledge creation, whether in electronic or other form, to its final retrieval and information presentation to the knowledge workers.[31]

Stated more simply, the information value chain is the process of transforming data into information, and the delivery of that information when and where it is needed, for transformation into knowledge that can be used for making informed decisions. Supporting the information value chain (i.e., managing information as a product) requires four primary components: 1) Understanding the information consumers' needs; 2) a technical information systems infrastructure; 3) planning, developing, and

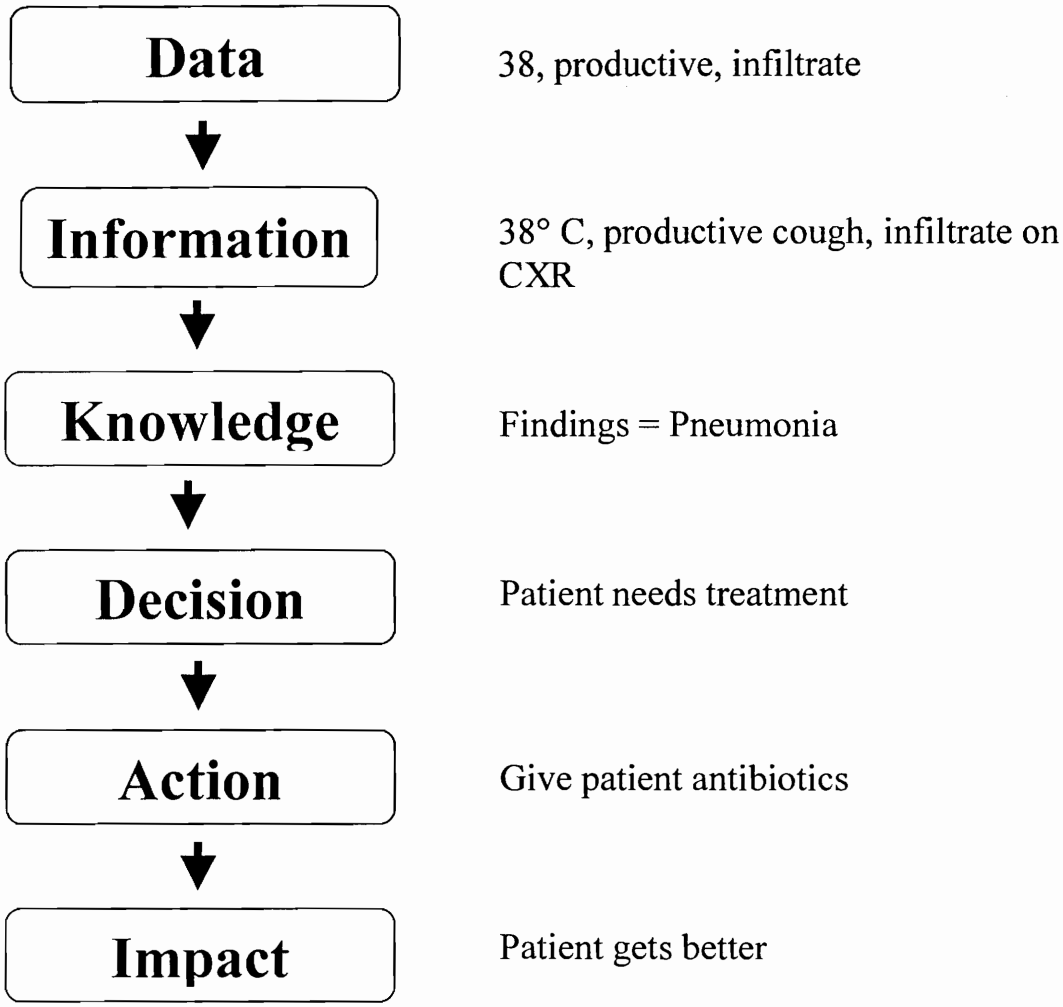


Figure 1. Transformation from data to information to knowledge.

implementing processes (including education for producers, custodians, and consumers) for managing the flow and use of information; and 4) planning, developing, and implementing a program for monitoring and continuously improving the quality of information.[31,32,34]

The technical infrastructure required to support the information value chain has been described by Inmon, Imhoff, and Sousa as an "information factory."[35] The components of the factory include the following: applications, a data integration and transformation layer, a data warehouse, one or more data marts, an operational data store (ODS), metadata, and the communications infrastructure (local area network [LAN], wide area network [WAN], internet, intranet). The information factory also includes the integration of external data with an organization's internal data. The information factory uses the raw material, data; turns it into the finished product, information; and delivers it to the people who will use it in their daily work. Thus, the information factory creates a foundation for information delivery and decision-making activities. Building the information factory is an iterative process. It must be built in the context of the people and processes it will be used to support. Figure 2 shows the basic structure of a generic information factory.

Although it is true that information is a strategic resource for most industries and it is essential to have a robust technical infrastructure that facilitates the use of the information, without people to use it, information has no value. Therefore, knowledgeable people are the most important resource in any organization.[31] Providing these people with the knowledge and skills to perform their work in a way that will allow the organization to achieve its goals requires giving them the tools

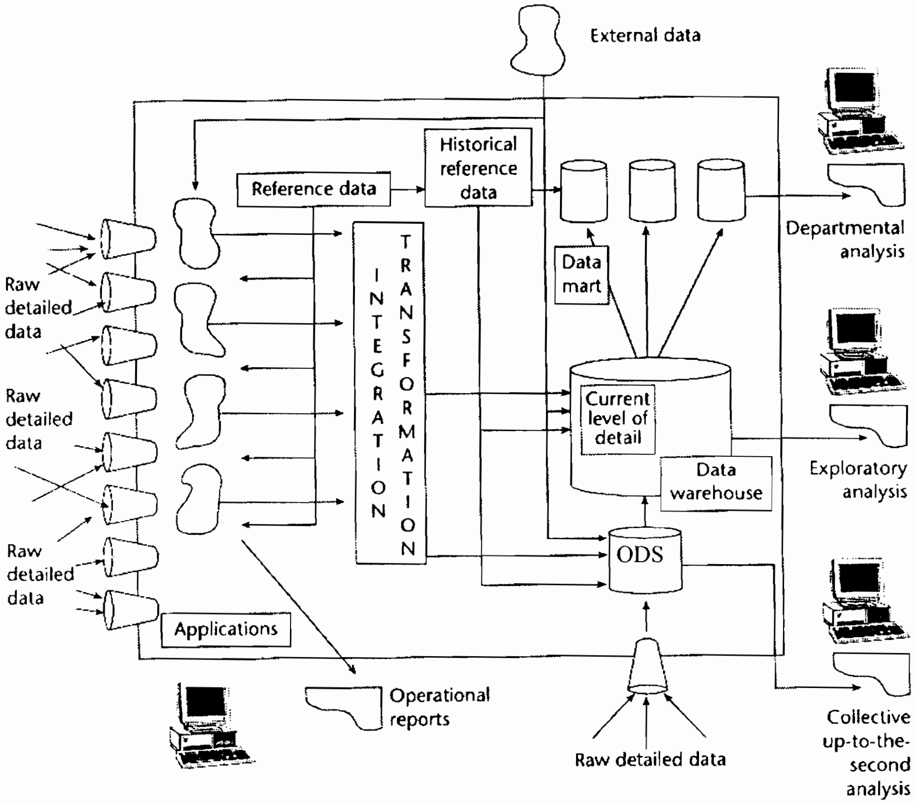


Figure 2. The basic structure of a generic information factory. (Adapted from *Corporate Information Factory* by Inmon WH, Imhoff C, Sousa R. Copyright © 1998 W.H. Inmon, Claudia Imhoff, and Ryan Sousa.[35] Permission granted by Wiley-Liss, Inc., a subsidiary of John Wiley & Sons, Inc.)

they need and educating and empowering them to use them. Information is one of those tools. In the case of information, people must be educated about determining the information they need to make decisions, from where and how to acquire that information, and how to interpret and use it once it has been acquired.

The forces driving the changes in the healthcare industry also are driving changes in the information requirements of healthcare organizations and the way in which they must manage and use that information.[4,36,37] Therefore, an evolution towards focusing on the comprehensive management of information as an essential resource is critical to the survival of the healthcare industry into the 21st century.

Information Management

The comprehensive approach to information management differs from a purely technology-centered (traditional) approach, in that it is "process-centered." The process-centered approach is broader than the technology-centered approach, because it expands the focus of information management to include the people, processes (including work flow), organizational structure, governance, and technology required for the production, capture, storage, integration, maintenance, quality, delivery, and utilization of information.[29] The process-centered approach is driven by the business goals and strategic initiatives of the organization. Thus, in the new paradigm, information management can be defined as the systematic approach to assessing and addressing the information needs (not limited to computerized information systems) required to enable an organization's business strategy and optimally support its core processes, by designing and implementing the roles, processes, and systems that facilitate the collection, flow, and use of information within the organization.[37-40] In the case of healthcare organizations, the way in which information is managed must support easy access to high quality care at "reasonable" costs, while maintaining and improving the operational performance and financial viability of the organization.

The importance of focusing on the people and processes, in addition to the technology, for the proactive management of information was recognized in healthcare, as far back as 1976.[39,41] However, it is only recently this topic has begun to receive the attention it needs in the healthcare industry to realize the full value of its potential. Using the new principles of information management, an organization can better determine how to build their information factory, how to identify the information they need to achieve their goals, and how to use the information most effectively.

An information management strategy includes a number of components that must work in concert for the strategy to be successful. The first component required for success is executive commitment to the following: creating standard operating processes across the organization, developing a comprehensive information management strategy, and recognizing the strategic importance of information systems. The second component is a well-defined information management governance structure. The third component for success is an information systems department organizational structure, which is aligned with organizational processes, supports information and workflow integration across clinical, financial, and administrative functions, and is independent of geographic location. Because most healthcare organizations already have made significant investments in information technology, information management plans must be designed to leverage existing investments in information systems where possible and incorporate an assessment of ongoing initiatives. The fourth, fifth, and sixth components are an enterprise data management strategy, data model, and metadata (information about the characteristics of the data in information systems); management of security and confidentiality; and an enterprise information quality program.[4,29,31-33,42,43]

Bringing these components together requires healthcare organizations to adopt a new approach to planning that is aligned with the comprehensive approach to information management and positions information systems as a tool to enable the business strategy of the organization. To be successful, the comprehensive approach to information management requires collaboration between the chief executive officer (CEO), chief financial officer (CFO), chief information officer (CIO), chief medical officer (CMO), healthcare informaticists, the medical staff, and the entire spectrum of clinical and nonclinical workers in the organization who will be using the information to support the organization's core processes.[44-47]

In their work on information quality, Strong, Lee, and Wang describe the concept of an information manufacturing process that is supported by the foundation of an information systems infrastructure.[33] Inmon's concept of an information factory[35] is consistent with the information systems infrastructure of Strong, Lee, and Wang. The information manufacturing model of Strong, Lee, and Wang provides a good framework from which to begin building an information management strategy. Particularly useful is their concept of information producers, information custodians, and information consumers. In healthcare, examples of information producers includes nurses, pharmacists, physicians, patients, admitting staff, accounting staff, and automated instruments in the laboratory or intensive care unit. The information custodians include members of the information technology staff, as well as those people responsible for data administration. Information consumers overlap with information producers and include patients, clinicians, managers, executives, accreditation and regulatory agencies, governing boards, and the community.

When developing an information management plan, among the questions that should be asked are the following: Who will the information consumers be, what business functionality will be supported/required by the various information consumers, what information is necessary to support that functionality, what data sources will be used to produce the information, and how will the information be accessed and presented?[29,31] It is important to remember that the development of an information management plan is an iterative process. The plan should be customer focused and use the principles of Continuous Quality Improvement (CQI) by performing a continuous cycle of assessing information consumer needs, assessing current organizational strengths and weaknesses in meeting those needs, and creating, reviewing, prioritizing, and implementing strategies for information production, storage, maintenance, and use that will fill the gaps identified.[31]

In developing the plan, it also is important to be aware of the variety of information needs that exist in healthcare. Healthcare information needs can be grouped into eight categories.[48] The categories, a description of the type of information they contain, and potential uses of the information are outlined below:

1. Patient specific information includes clinical, demographic, and account information about a specific patient. In general, this information is used in the daily processes (clinical, administrative, and financial) of care delivery.
2. Clinician specific information represents activities or performance of a specific clinician. The information may be used for performance improvement, credentialing, or compensation.
3. Aggregate information can be clinical, administrative, financial, or some combination thereof that is combined and summarized. It often is used as an outcome measure in quality improvement, as well as for regulatory reporting, and various types of organizational decision making.
4. Organization specific information includes clinical, administrative, and financial information. The information may be about services or performance of a facility or organization (e.g., operating budget, average length of stay, etc.). It often is used for internal and external benchmarking and for making decisions about business and clinical operations.
5. Community specific information includes demographics and information about preferences, satisfaction with services, or health status of the community. Information of this type is becoming increasingly important as the focus on population health continues to grow. The information in this category may be used by a healthcare organization, community organization, or government agency for designing community health programs, disease management programs, tracking patients/health plan members with chronic diseases, or planning for future resource use.

6. Job-related information includes information that is needed for a specific class of job. For example, a person working in the business development area of a healthcare organization may need the zip code distribution of all patients who use their services, to determine where to target marketing campaigns.
7. External comparative information is needed to assist in healthcare research and for comparison of regional, national, or international health-related outcomes.
8. Knowledge-based information includes facts, models, best practices, guidelines, care paths (the latter three are tools used in the practice of evidence-based medicine), etc., that can be used for education, designing or redesigning processes, or clinical and administrative problem solving.

Unfortunately, the healthcare industry is facing a variety of challenges in moving from the traditional approach of technology-centered information management to the new paradigm of process-centered information management. The challenges fall into three main categories. The first category, organizational challenges, may vary by organization; however, a number of common themes can be identified. For example, many, if not most healthcare organizations, are still struggling with what it means to be integrated. The question may pertain to how to integrate hospital and clinic facilities that were formerly independent or how to integrate physician groups with a hospital or hospital system.[4] In addition to the issues of how to design the flow of information and the technical infrastructure to support it, the major issues this challenge raises are how to address the cultural and legal barriers to the sharing and integration of information. Another issue many healthcare organizations are facing that impacts the resources put into information management is how to balance the need to maintain a "bottom line" with the need to invest in people with new skill sets (epidemiology, statistics, health services research, informatics, etc.) and an information infrastructure that will facilitate the delivery and continuous improvement of their services. Other issues include the recognition and implementation of standardization in organizational processes (including information management), engaging physicians in critically looking at how they practice and the relationship to outcomes, implementing clinical practice guidelines, educating people about the common principles of CQI, and training people how to "ask questions," analyze data, and use information technology as a tool.[28,49-56]

Technological challenges, the second category, are numerous. One of the most significant ones is the integration of new information systems with legacy systems. Equally important are the challenges of integrating clinical, administrative, and financial information systems and integrating external data sources with internal data sources.[57] Technical challenges that have become more widespread in the last couple of years include implementing real-time clinical decision support (this also is an organizational challenge from a change management perspective),[58-60] building and implementing data warehouses,[61-63] implementing tools to provide push (information delivery with no action on the part of the consumer) and pull (information delivery in response to a request from the consumer) information delivery for information consumers,[64-67] and complying with the various regulations and requirements arising from the Health Insurance Portability and Accountability Act (HIPAA) of 1996.[68-73] Among the technical challenges created by HIPAA is the need for healthcare organizations to implement solutions for 1) improving the security of data housed within the organization and transmitted over intranets or the internet, 2) performing electronic data interchange using a standard format, and 3) performing electronic signatures.

The third category of challenges healthcare (and nonhealthcare) organizations are facing in moving towards process-centered information management is that related to the information itself. The primary problem resulting from these challenges can be summarized in the concept of "data toxicity." Data toxicity is an overload of redundant, inaccurate, uninformative, or confusing "facts" leading to incorrect conclusions. In other words, so much data are being generated and disseminated, that the ability to produce the information needed is being lost in the process, and the quality of the information that is being produced is not being optimized.[74] The major challenges leading to the problem of data toxicity are the lack of data standardization, including terminology and information representation; the absence of standard processes for capturing and storing data; and the poor quality of data in both paper-based and electronic information sources.[75-78] All three of these issues negatively impact the healthcare industry's ability to integrate and use information of different types or information of the same type from different sources.

Each group of challenges (organizational, technical, and information related) has its own set of solutions. To overcome the information-related challenges, the healthcare industry must work concurrently on the issues of standards (terminology and representation) and information quality (including the processes for capturing and storing data). Various groups currently are working on standardizing data definitions, vocabularies, and data representation.[76] These efforts will help improve the ability to integrate clinical, administrative, and financial data for supporting population-based and organizational decision making. These efforts also will aid in the ability to integrate external databases and knowledge sources with internal organizational information. However, no matter how standardized the data are, if the information is of poor quality, it still is not useful to healthcare organizations. Unfortunately, in contrast to the development of data standards, few organized efforts focusing on developing and implementing information quality programs exist in healthcare, either at a national or local level. Extensive literature searches identified ongoing efforts in the United States, specifically targeted at improving information quality, in only four healthcare provider organizations [79-82] and one payer organization.[83]

Continuously assessing and improving the quality of information are essential to its effective use. Developing and implementing an information quality program are important parts of an information management strategy.[1,4,29] The National Committee on Vital and Health Statistics recently released a report mandated by HIPAA, entitled "Uniform Data Standards for Patient Medical Record Information."[84] In addition to advocating data standards, the authors of this report clearly state that emphasis must be placed on improving and maintaining the quality of data in medical records. Brennan and Stead, in an editorial in the *Journal of the American Medical Informatics Association*,[85] lend further support to the need for a focus on the quality of information in computer-based patient records. The topic of information quality is the focus of this thesis. The remainder of the introduction will be devoted to a review of the literature on information quality and the specific purpose of this project.

Information Quality: A Review of the Literature

Definition of Information Quality

The basic principles of the field of information quality and approaches to information quality improvement have been adapted from those used in the practice of Total Quality Management (TQM) and Continuous Quality Improvement (CQI). The principles are derived from those of quality pioneers such as Deming, Juran, Crosby, Ishikawa, Shewhart, and others. According to the quality literature, it is the consumer who will judge whether or not a product is fit for use and thus of good quality.[86-89] The consumer-centric approach to quality is continuing to evolve as evidenced by the recognition, in service industries like healthcare, that the consumer-base is heterogeneous and that there is a need to continuously adjust to their dynamically shifting expectations.[90] Therefore, most experts in the field of information quality agree that information quality cannot be assessed independent of the consumers who use it. These experts define quality information as "information that is fit for use by information consumers." The experts also agree that to be of high quality, information should be intrinsically good, contextually appropriate for the task, accessible to the information consumer, and clearly represented.[31,91,92]

English further refines the definition of IQ by identifying two primary types, inherent IQ and pragmatic IQ.[31] Inherent IQ is the degree to which data accurately reflect the real-world objects they represent. In other words, the degree to which the data are "correct." Pragmatic IQ is the degree of usefulness and the value the data have in supporting the enterprise processes that enable the achievement of enterprise objectives.[31] Pragmatic IQ is reflected in the degree of customer satisfaction derived by knowledge-workers who use the information to do their jobs.

The process of providing quality information, therefore, requires consistently meeting knowledge-worker and end-customer expectations through information and information services and enabling them to perform their jobs efficiently and effectively. The process of providing quality information applies to all purposes for which the information is used, including both present and likely future uses.[31]

Research in Information Quality

Major research efforts focused on information quality are actively being pursued. However, as noted previously, the programs organized specifically to focus on the field of IQ exist primarily outside of healthcare. A well-known example of the work being done external to the healthcare field is the "Total Data Quality Management Program" (TDQM) at the Sloan School of Management, Massachusetts Institute of Technology (MIT).[33,42,93-97] In addition to programs at academic centers such as MIT, individual consultants, such as Larry English of Information Impact International, Inc. and Thomas C. Redman, not only conduct research in the field of information quality but also help organizations apply the results of that research in a practical fashion.[31,34,43] One of the more successful examples of information quality work being done within healthcare is the IQ program at Cedars-Sinai Health System in Los Angeles, California.[79]

According to Strong, Lee, and Wang, before they began the TDQM program at MIT, data quality research focused on the intrinsic quality of data in databases, rather than the quality of data and information in the context of how it is produced and used.[91] In contrast, the overall objective of the TDQM research program is to establish a theoretical foundation for the field of information quality and to devise a practical method for business and industry to improve all aspects of the quality of information. The short-term goals of the MIT program are to "create a center of excellence among practitioners of data quality techniques, and act as a clearinghouse for effective methods and project experiences." The long-term goals include "creating a theory of data quality based on reference disciplines such as computer science, the study of organizational behavior, statistics, accounting, and the TQM field."[97] These reference disciplines are similar to those on which the field of healthcare informatics is based.

The three major components of the MIT TDQM research program are data quality definition, data quality analysis, and data quality improvement. The data quality definition research includes developing a definition of data quality and a method for data quality measurement. The data quality analysis research uses the products of the data quality definition research and focuses on the relationship between and the identification and calculation of the impacts of poor quality data and the benefits of high quality data on an organization's effectiveness.

The goal of the data quality improvement component of the TDQM research is to identify and address the areas for improvement opportunities that will result in significantly improving the overall quality of corporate information. The data quality improvement research has four subcomponents: business redesign, data quality motivation, use of new technologies, and data interpretation. Business redesign includes simplifying and streamlining operations in a way that will minimize the opportunity for data errors to occur. Data quality motivation evaluates the impact of employee rewards and benefits and ways to improve employee perceptions of the value of their work. The goal is to determine how to encourage employees to pay more careful attention to the way in which data are handled and thus improve the quality of information. The research on the use of new technologies focuses on new technologies for capturing data and the way in which these technologies can contribute to improving information quality. The data interpretation research is targeted at developing ways to assist information consumers in understanding the meaning of data so that it is not used incorrectly.

Wang and Strong identified three approaches used in the literature to study data quality: intuitive, theoretical, and empirical.[98] The intuitive method looks at accuracy, timeliness, precision, reliability, currency, completeness, relevance, accessibility, and interpretability of information. The theoretical method focuses on how data may become deficient during the "manufacturing process." The empirical method focuses on two primary issues: capturing quality attributes important to information consumers and analyzing data collected from consumers to determine characteristics they use to assess whether data are fit for use in their daily tasks. It is the empirical approach that primarily is used by Wang and Strong in their work.[95,96,98] The empirical approach is the best way to "capture the voice of the consumer" in the evaluation of information quality.

The basic goals and principles used by English in his research and consulting on information quality do not differ significantly from those of Wang, Strong, and

Lee.[33,42,92,95,97,98] Even the name of English's approach to information quality is similar to that of Wang, Strong, and Lee. It is called Total Quality data Management (TQdM).[31] The tools used in theTQdM approach to information quality are similar to those used in general applications of CQI. The TQdM methodology uses five processes to address the measurement and improvement of information quality and a sixth process that functions as "an umbrella for bringing about cultural and environmental changes to sustain information quality improvement as a management tool and a habit."[31] The five processes used to address measurement and improvement of information quality are as follows: 1) assessment of data definition and information architecture quality, 2) assessment (measurement) of information quality, 3) measurement of nonquality information costs, 4) reengineering and cleansing of data, and 5) improving information process quality. In the latter process, information quality problems discovered through the four initial processes are addressed through the analysis of root causes and the planning and implementation of process improvements that will prevent further data defects. The sixth process, the establishment of the "Information Quality Environment," is actually a combination of several processes. The establishment of an IQ environment requires systemic, management, and cultural changes.

There are two key paradigm shifts underlying the sixth process in the TQdM approach to information quality improvement.[31] The first paradigm shift is changing the organizational perception of information as a byproduct of work processes to a perception of information as a direct product and essential resource of the organization. The second shift is changing the acceptance of the costs of nonquality information as a normal cost of doing business to the belief that nonquality information costs are unacceptable and may threaten the viability of the organization. The key to changing these perceptions in any organization is to make everyone accountable for their processes and the quality of the information they produce. The way to achieve the goal of making everyone accountable is by helping information producers understand who their customers are and what expectations these customers have, developing performance measures that include customer satisfaction with information products, and providing education about information quality for the entire organization.[31]

Although as far back as 1974 the Institute of Medicine (IOM) was performing studies on the quality of data contained in hospital discharge abstracts,[99] within healthcare active discussions about information quality only recently have begun to receive significant attention.[100] Research in the area of information quality in healthcare is relatively sparse and has focused primarily on paper-based medical records,[101-105] disease registries,[106-108] clinical trial databases,[109-112] and administrative databases.[113-116] Little research has been devoted to evaluating information quality in computer-based patient record systems (CPRS). Moreover, rather than addressing the general process issues or data issues that could improve overall information quality in CPRS, with few exceptions,[77,117-120] most of the research on CPRS information quality has focused on selected components of a general computer-based patient record or on the quality of information in a specialty specific computer-based patient record.[121-129] In their study of the accuracy of data in CPRs, Hogan and Wagner found only 26 studies for the years 1977-1995, in which the object of study was a CPR, a gold standard to which the computer-based records were being compared was identified, and in which a defined measure of data quality was evaluated.[117] Moreover, using a scoring system (scale of 0 to 18 points) based on three main factors: 1) CPR description, 2) methodology, and 3) study objective, Hogan and Wagner concluded the quality of these studies was variable and generally not high.

Even less attention has been focused on the quality of data in analytical databases and data warehouses used for healthcare decision making and health services research.[107,108,114,130-132] The absence of attention to data quality in analytical systems is unfortunate, in that of the many issues challenging the ability of organizations to develop successful data warehouses, data quality generally is regarded as among the most significant.[43] In addition to the relative paucity of research on the cause of and solutions to information quality problems in CPRS and analytical databases, the impact of information quality problems also has received little attention.[75,78,117,122,123,133-138]

Categories and Dimensions of Information Quality

As stated earlier, information quality generally is defined as "information that is fit for use by information consumers."[31,33,34,98] However, this definition does not convey the complexity behind the concept of information quality. Just as the quality of any manufactured product has several dimensions, so does the quality of information. Unfortunately, there is no agreement in the literature about the dimensions of information quality and their definitions. Thus, it has been difficult to compare or aggregate the various research efforts focused on measuring and improving information quality. In an attempt to address this issue and to identify the dimensions of IQ that are important to the information consumer, Wang and Strong developed a hierarchical conceptual framework in which they identified 15 dimensions of IQ and grouped those dimensions into four categories.[33,91,98] Wang and Strong began their work by proposing a preliminary framework based on the assumptions that data must be accessible to the consumer, the consumer must be able to interpret the data, the data must be relevant to the consumer, and the consumer must find the data to be accurate. Using marketing research survey techniques, the end result of their work was the development of the following categories and associated dimensions:

- Intrinsic Data Quality – Believability, accuracy, objectivity, reputation
- Contextual Data Quality – Value-added, relevancy, timeliness, completeness, appropriate amount of data
- Representational Data Quality – Interpretability, ease of understanding, representational consistency, concise representation
- Accessibility Data Quality – Accessibility, security

This framework captures the majority of dimensions described in the work of other authors and provides a relatively comprehensive model on which to base information quality research.[31,117,119,121,139,140] The emphasis Wang and Strong placed on identifying dimensions that are important to the consumer is exemplified by the dimensions in the Intrinsic Data Quality category. In their research, Wang and Strong determined that accuracy and objectivity of information are the two dimensions of IQ to which information technology workers generally pay attention.

Wang and Strong noted little attention is paid to the believability and reputation of the information. Although little attention is paid to these latter two dimensions, they are important barriers to overcome in getting information consumers to use the information.[33]

Although the framework of Wang and Strong is a step in the right direction, researchers in the field of information quality still face the challenge of standardizing on the definitions and measurement methods for the various dimensions. Depending on how a dimension is defined and measured and the way in which a study is designed, the results of an IQ study may be incongruous with the perception of the information consumers or not provide information useful in addressing the information consumers' needs. Using the accessibility dimension of information quality as an example, the potential for different interpretations of the results of an IQ study, in the absence of a clear definition, measurement method, and understanding of the consumers' perspective, can be demonstrated. When evaluating information accessibility in an enterprise-wide CPRS, information combined across autonomous systems may be technically accessible, but information consumers may view it as inaccessible because similar data items are defined, measured, or represented differently. Similarly, coded medical data may be technically accessible as text, but information consumers view it as inaccessible because it is the codes that are provided to them, and they cannot translate the codes into text. In the case of large volumes of data, the data may be technically accessible, but information consumers may view them as inaccessible because of excessive access time. If, in performing a study on information accessibility, the measurement was based on technical accessibility, the results could be interpreted by the researcher as confirming that a system provides easily accessible data and thus meets the information consumers' needs. However, even though the data may be technically accessible, for the reasons stated above, the system would not be meeting the consumers' needs from their perspective, and thus the researcher's conclusions would be erroneous. Therefore, when performing an evaluation study of the quality of information, it is important to identify the dimension(s) of information quality being evaluated, state the working definition of the dimension(s), consider the information consumer's perception of the dimension(s), and use a measurement method that will yield results that can be used to guide improvements in IQ from the consumers' perspective.

Categories and Causes of Information Quality Problems

General information quality problems. Quality problems may arise anywhere in the information value chain. An information quality problem can be defined as "any difficulty encountered along one or more quality dimensions that renders data completely or largely unfit for use."[91] Information quality problems are more than incorrect information values. They also include production problems and errors, technical problems with storage and access, and problems caused by changing information needs of consumers.[33]

Using their framework for information quality research, Strong, Lee, and Wang divided information quality problems into three categories: intrinsic quality problems, contextual quality problems, and accessibility quality problems.[91] Although representational data quality is a separate category in the information quality framework, problems with representational information quality were incorporated into the three categories listed above. Although not addressed as a separate category by Strong, Lee, and Wang, representational information quality problems clearly are a major barrier to the production and use of high quality information, especially in healthcare.[76]

Intrinsic data quality problems are manifested in a variety of ways; however, all problems in this category result in data not being used because of little added-value or questionable believability and objectivity (i.e., poor reputation). Two examples of intrinsic data quality problems are mismatches between different sources of the same data and the use of judgment or subjectivity in the data production process. In the case of mismatches between data sources, the problems appear as believability issues, which in turn lead to poor reputation of the data source and decreased use of the data. An example would be the storage of a single patient's medication data in separate inpatient and outpatient CPR databases. If these systems are not synchronized, through an interface, for example, changes in patient information may be made in one system and not the other. If the patient with differing medication information in the two systems comes to the clinic for an acute medical problem and gets admitted to the hospital, the medication information the clinic physician provides the admitting physician may not match the information in the inpatient CPR. The differing information may lead to confusion about the true medications the patient is taking and result in medication errors. The question in the mind of the information consumer (admitting physician in this case) becomes which source is accurate (i.e., the gold standard). Examples in which judgment or subjectivity is used in the production of data are the cases of coded data (e.g., ICD-9-CM or CPT codes) and subjective rating systems (e.g., severity of illness measurement systems). In the case of coded data, medical record coders use, in part, subjective judgment to decide on the "correct" procedure and diagnosis codes. In the case of severity of illness systems, nurses and physicians often use subjectivity in "rating" the patient's severity of illness. In both cases the problems appear as concerns about data objectivity and, thus, lead to lack of trust in the data and decreased use.

Contextual data quality problems directly lead to an inability to integrate or aggregate data and, thus, difficulties with data utilization. The underlying causes of contextual data quality problems are incomplete data, inadequately defined or measured data, inconsistent representation of data, poor relevancy of data, and little added-value to the information consumers' tasks. For example, DRG codes stored with decimal points in one facility of a healthcare delivery system and without decimal points in another facility preclude aggregating this information (without data conversion processes) and evaluating services across the organization. Another example is the case in which basic utilization measures are defined differently across divisions within a hospital. One division may define utilization as hospital days per thousand patients, whereas another defines it as hospital days per hundred patients. A final example is the case in which an outpatient facility in an integrated delivery system stores laboratory results using the units of grams per deciliter, whereas the hospital in the same organization stores the results for the same test as milligrams per liter. The difference in representation of these data would preclude integrating inpatient and outpatient lab results in a common database, unless a data conversion takes place.

Accessibility data quality problems arise from a variety of barriers consumers encounter in trying to access information. These barriers may be related to communications technology, security and confidentiality issues, representation and interpretability of data (and thus an overlap with contextual problems), and volume of data. Technical barriers to accessing data often are related to a poor communications infrastructure. Poor communications infrastructure includes limited or poorly designed networks and poor technical quality of the communications lines. Security and confidentiality barriers may be related to technical data security safeguards, organizational policies, and/or the processes of data custodians which may add time and effort to the process of data access and retrieval. Representation of data presents a barrier to access when multiple specialists are needed to interpret data across multiple databases and representations (e.g., ICD-9-CM or CPT codes) or when information consumers use different systems and do not understand the way data are represented in the systems. Representation barriers and interpretability problems are interrelated. For example, if imaging studies are stored only as the image data, as opposed to the image and the textual interpretation, when retrieving the imaging studies without the associated interpretation, the information consumer may not be able to understand the meaning of the images. Lastly is the issue of data volume. Large amounts of data can lead to timeliness issues in data access. As healthcare organizations are beginning to capture increasing amounts of clinical data and are using the data in process and outcomes improvement, the size of the analytical databases in which the information is being stored is growing to the gigabyte range. The queries run against these databases may require processing through millions of rows of data. Obtaining the results from these queries may take hours. The timeliness issues may be a result of several factors such as slow processing due to inadequate computing power, poorly designed databases, or poorly designed queries.

After defining the categories and dimensions of information quality problems, in a subsequent work, Strong, Lee, and Wang describe which of these problems are the key information quality problems in organizations they have observed.[33] They identify 10 key problems and categorize them in accordance with the concept that information is a product that is manufactured, stored, and utilized. The categories used are information production, information storage, and information utilization.

The problems in the information production category are 1) multiple sources of/processes for producing the same information yielding different values, 2) production of information using subjective judgments leading to bias, and 3) systemic errors in information production leading to lost information (e.g., data entry errors due to technical issues that prevent accurate data entry). The information storage category problems are a result of the difficulty in storing large amounts of information across different computer systems. They include 4) large volumes of stored information making it difficult to access in a reasonable time; 5) distributed heterogeneous systems leading to inconsistent data definitions, formats, and values; and 6) nonnumeric information being difficult to index. The information utilization quality problems are 7) the absence of the ability to perform automated content analysis across information collections (due to inconsistent definitions, names, formats, etc.); 8) the changing information consumers' tasks and organizational environment resulting in the information that is relevant and useful also changing; 9) easy access to information conflicting with requirements for security, privacy, and confidentiality; and 10) lack of sufficient computing resources limiting access.

Healthcare specific information quality problems. Problems with the quality of information in paper- and computer-based patient records, associated analytical databases (e.g., data marts and data warehouses), and claims-based administrative databases present a barrier to using the information for daily processes of care (including clinical decision support), health services research, regulatory reporting, and clinical quality improvement.[78,121,141-143] A prime example of the barrier caused by poor information quality in healthcare is the problem with risk adjustment strategies. The proliferation of proposals for risk adjustment strategies is one of the key components in the growth of outcomes research. Ultimately, the success or failure of the exploration for causal links in outcomes will depend to some degree on scientifically credible risk adjustment strategies.[144] Unfortunately, many of the risk adjustment strategies in use today are inherently imperfect. Iezzoni points out that problems with the quality and availability of data are one of the major problems in performing valid risk adjustment. She postulates the most immediate gains in risk adjustment strategies will be achieved by improvements not only in experimental design but also in information quality.[142] Another example of the barriers poor quality data pose to the use of information in healthcare is clearly articulated by Overhage, Tierney, and McDonald[59] when they state "probably the most challenging aspect of implementing a clinical decision support system is ensuring that the necessary data are available and valid." In their experience, lack of good data is one of the major barriers to the use of decision support systems. This experience is supported by that of Hogan and Wagner.[117,141]

In paper- or computer-based medical records, high quality patient data are data that represent the true state of the patient. Deviation of recorded medical information from the actual or "true" state of the patient always has been present in medical records.[121,141] Information flows into the medical record through a complex series of processes.[101,117] Information may come directly from the patient; from a physician, nurse, or other clinician; from transcriptionists; from data entry personnel; or from other information systems such as a laboratory system. As a result of the multiple processes and sources that bring information into medical records, the quality of information is influenced by a number of factors, and data may become inaccurate in a variety of ways. One factor influencing the quality of data in medical records is inconsistent or incomplete record keeping on the part of clinicians when compared to 1) what the patient told them, 2) what they told the patient, 3) suspected or confirmed diagnoses, and 4) actual treatment rendered. The information documented has been found to be influenced by the clinical relevance of the information as perceived by the clinicians.[125] The accuracy of information provided by patients and family members is another factor influencing the quality of medical record data. Operational process issues such as 1) the use of multiple mechanisms for data entry, 2) the type of user interacting with the system, 3) the kind of function the user performs, and 4) the kind of coding systems used also may impact data quality. Representational limitations for some patient characteristics, such as satisfaction, and the quality of human-machineinterfaces are two additional factors influencing medical record data quality.[105,115,116,118,119,125,127,128,145-147]

The ways in which these factors may lead to inaccurate information include patient's giving poor information, patients being misidentified, clinical findings being misinterpreted, laboratory studies being misanalyzed, clinical findings being misrecorded, laboratory reports being misfiled, transcription or coding errors, or queries being constructed incorrectly resulting in misselected data.[119] In addition, if the true state of a patient changes with time due to effects of disease or treatments, data errors may accumulate from a lack of recent observations by the patient's clinician. Alternatively, the patient may not tell their physician about a change, such as the fact they have changed their medications or stopped taking them. The physician's lack of knowledge of the change or discontinuation of medications would result in inaccurate information in the record. If patients receive treatment outside of their "home" institution and do not tell their physician, this will result in an incomplete record. Many coding problems stem from incomplete or conflicting medical record documentation.[102,127] Errors in diagnosis and procedure codes also may occur due to ambiguity in coding classification systems, limitations on the number of diagnoses and procedures recorded on a face sheet or in a computer-based record, or financial reasons such as the increasing need to justify utilization of healthcare resources such as hospital days, medications, and diagnostic procedures. The need to justify utilization of resources may lead to bias in reporting to justify clinical care and thus insure reimbursement, rather than focusing on truly reflecting the state of the patient.[146]

While large claims databases have some advantages over clinical databases derived from CPRs, with respect to population-based health services analyses, the claims databases are not free from information quality problems. The advantages of administrative databases when compared to a single organization's clinical data repository or data warehouse are the large numbers of individuals included, the broad cross-section of patients and clinical practice, the length of time over which patients have been followed, the high proportion of patients for whom follow-up data are available, and the comparatively low cost. The problems include variations in coding, errors in coding, incompleteness in coding, and errors in clinical diagnoses in the primary data sources resulting in errors in diagnoses in the administrative database.[133]

Impact of Poor Information Quality

Information quality is not only about the accuracy of data but also about the accuracy of the inferences made about that data by information consumers.[148] Errors in databases outside of the healthcare industry have been measured in the 10% range and higher for a variety of applications.[95] Studies that have examined information quality in healthcare (CPRS, research, administrative and epidemiological databases) have found error rates that range from 5% to 51%.[77,108,116,123,128,135-138,147,149-151]

Poor quality information causes process failures and information scrap and rework that waste money, materials, people's time, and facility resources.[31] Inaccurate, out-of-date, or incomplete data can have significant impacts both socially and economically.[95] The problems in the FBI identification records and the resulting impact on the due legal process are examples of the social impact of poor information quality outside of healthcare. In a study performed by Laudon,[152] 74% of FBI identification records exhibited significant quality problems. These errors resulted in some people being detained when they were not guilty and some guilty individuals being erroneously released. Social and economic impacts of poor quality data cost billions of dollars each year.[91] A recent report of 500 medium-size corporations with annual sales of more than $20 million demonstrated that more than 60% of the corporations have problems with data quality.[98] According to Larry English and Thomas Redman,[31,34] 10%-15% of operating budgets for many organizations go into scrap and rework due to poor information quality. Every hour spent hunting for missing data, correcting inaccurate data, or working around data problems is an hour of cost only.

One of the most significant problems caused by poor-quality information, however, is that it frustrates the most important organizational resource - people. Poor quality information keeps workers from effectively performing their jobs and alienates customers by providing wrong information to and about them.[31]

Data in computer-based patient records, associated analytical databases, and administrative databases are used for a variety of purposes. The range of purposes includes daily patient care activities, automated clinical decision support, process and outcomes measurement and management, health services research, health system planning and management, regulatory and accreditation reporting, and reimbursement.[78,99,119,123,127,136,153-155] In healthcare, poor quality information has social, economic, and clinical consequences. For example, alerting systems sending false alerts or failing to send alerts to physicians and other care givers may result in treatment errors. Poor quality information can lead to erroneous research conclusions such as researchers underestimating compliance with standards of care. Erroneous results such as these may lead to poor health system planning. Without good quality information, clinical quality improvement efforts may be impeded, and healthcare organizations may end up in noncompliance with government reporting regulations, commit inadvertent fraudulent billing, or make decisions that could negatively impact their viability.[13,75,117,122-124,133,134,137,138,145]

Wilton and Pennisi performed a study on the quality of information in their pediatric immunization tracking system that demonstrates the potential impact of poor quality information in healthcare program planning.[128] Information on immunizations in their institution is first recorded on paper, then transcribed into a computerized tracking system. Based on the information in the tracking system, they concluded that 22.5% of children who received primary care at their institution were under immunized. However, they also found that 50% of children who were under immunized according to the immunization database had inaccurate records in the computerized tracking system when compared to written records. A subsequent study performed by them, based on a comparison of the computer and written records, revealed the source of the information quality problems to be 10.2% transcription errors and 38.4% information incompleteness errors. The latter group of errors were a result of a combination of the processes for recording information on children being immunized outside of Wilton and Pennisi's institution and the process for transcribing the written record into the computerized tracking system. In the case of a child immunized at an outside institution, the immunization information was recorded in a portion of the written record that was not transcribed into the electronic record. When under immunization rates were reevaluated based on the written records, the rate dropped from 22.5% (based on the computer records) to 10.9%.

Although not specifically addressing information quality issues, the work performed on adverse drug events has demonstrated the importance of having accurate, high quality, easily accessible information, and the clinical and economic impacts of poor quality clinical information. Adverse drug events have been estimated to increase patient costs by an average of $2000 per event. Without accurate information on medication allergies or current drug therapies and in the presence of transcription errors for dose and frequency of medication administration, the likelihood of adverse drug events is higher, and the potential clinical and economic impact is greater than if accurate information was available and transcription errors did not occur.[156-160]

Studies performed by Lloyd and Rissing;[149] Doremus and Michenzi;[137] Johnson and Appel;[135] Aas;[138] Reidel, Brown, and Charles et al.;[115] and Hsia, Krushat, and Fagan et al.[151] demonstrate significant information quality problems in administrative databases. The results of their studies show the negative impacts of information quality problems on reimbursement and healthcare organization revenue. The impacts of the poor information quality identified in these studies resulted in losses of as much as $3 million in revenues for the healthcare organizations studied. These same quality problems also resulted in impacts on calculation of hospital case mix rates. In some cases, the case mix rates calculated with the poor quality information led to higher levels of reimbursement from Medicare than was supportable by the source medical records.

Purpose of Study

In his keynote address to the 1992 conference on "Medical Effectiveness Research Data Methods," J. Michael Fitzmaurice, Ph.D., Director of the Office of Science and Data Development of the Agency for Healthcare Research and Quality (formerly Agency for Health Care Policy and Research) stated: "As our nation faces the challenge of attempting to leverage the enormous power of its healthcare system by improving its effectiveness, we are drawn together at this conference by the potential that better methods and better data will lead to better patient outcomes."[143] Fitzmaurice went on to say that the "ideal data" are rarely available and that there are a number of needs and challenges that must be met to acquire better data and increase the effectiveness of our healthcare system's ability to address society's problems of access, quality, and the cost of healthcare. These needs and challenges include 1) understanding the data, the purpose for which they are collected, how they are collected, how they are checked for errors and corrected, how the elements are coded, and what the codes are intended to mean; 2) determining how the data failings may influence research results; 3) developing methods to reduce bias and large random variation and to obtain better patient outcome measures; and 4) developing rules for producing better data. In addition to the comments made by Fitzmaurice in 1992, in the National Committee on Vital and Health Statistics (NCVHS) July 2000 report on "Uniform Data Standards for Patient Medical Record Information,"[84] the committee states "the major impediments to electronic exchange of patient medical record information are limited interoperability of health information systems, limited comparability of data exchanged among providers, and the need for better quality, accountability, and integrity of data."

The purpose of the study performed for this thesis was to contribute to the body of healthcare research that focuses on information quality issues in computer-based patient records and associated analytical databases. It is my hope that contributing to this body of knowledge will add to the efforts to meet the needs and challenges Fitzmaurice and the NCVHS have put before us.

The broad objective of this study was to evaluate the quality of imaging services (radiology) data in the Health Evaluation through Logical Processing (HELP) computer-based patient record system and an associated clinical data mart (Imaging Services Data Warehouse [ISDW]) at LDS Hospital (LDSH), in Salt Lake City, Utah. The ISDW contains information relevant to the processes of delivering imaging services. The information includes, but is not limited to, clinical indications for procedures, patient demographics, patient diagnoses, and procedure results. The information in the ISDW does not include the images themselves.

The specific aims of the research on which this thesis is based were to do the following:

- Identify quality problems pertaining to the information about the clinical indication for ordering an imaging procedure.
- Determine the probable sources of the imaging services information quality problems pertaining to the clinical indication for imaging procedures

- Quantify information quality errors pertaining to the clinical indication for ordering an imaging procedure.

CHAPTER II

MATERIALS AND METHODS

Study Design

The focus of the research on which this thesis was based was the evaluation of imaging services information quality in the HELP System and Imaging Services Data Warehouse. The basic design of the research was that of a "Medical Informatics Information Resource Evaluation Study." In their book, *Evaluation Methods in Medical Informatics*,[161] Friedman and Wyatt present three definitions of evaluation studies. These three definitions were adapted from other authors and were modified for use in medical informatics. Each definition offers a slightly different perspective on the meaning of "evaluation." The definition pertinent to this discussion is the one that Friedman and Wyatt adapted from House.[162] The definition adapted from House states: "Evaluation leads to the settled opinion that something about an information resource is the case, usually, but not always leading to a decision to act in a certain way."[161] Friedman and Wyatt believe the definition they adapted from House "emphasizes evaluation as a process leading to a deeper understanding of an information resource."[161] The above definition adapted from House also best describes the focus of this thesis research. In addition to providing three definitions of "evaluation," Friedman and Wyatt classified evaluation studies into several categories. Based on the categories described by Friedman and Wyatt, the study performed for this thesis research would be classified as a "retrospective correlational evaluation study." Correlational evaluation studies explore the relationship between a set of variables that are measured but not manipulated in any way and are designed to facilitate the development and refinement of information resources. The independent variables of the research performed for this thesis were the processes for the production, capture, storage, and utilization of imaging services data. The dependent variable was the quality of those data. The dependent variable was evaluated in the context of the information consumers' perspective. None of the variables were manipulated in the course of the study.

The data sources used for this study were the Diagnostic Imaging and Nuclear Medicine (written) request form, the HELP System, a printed copy of the completed order for an imaging procedure (white sheet), the Hospital Case mix System, the ISDW, and the radiologists' dictated report of their interpretation of the imaging procedure. The individual data sources are described in more detail below. Figure 3 is a high level flow chart of the processes for producing, capturing, storing, and utilizing imaging services data. Figure 4 is a high level flow chart of the sub processes for ordering an imaging procedure. These sub processes lead to the production and capture of the imaging services data on which the quality assessment in this thesis research was focused. A more detailed explanation of the ordering process is provided in Chapter III.

The methodology used in the evaluation of the imaging services information quality of the HELP System and ISDW incorporated aspects of all three information quality research methodologies described by Wang and Strong.[98] The three

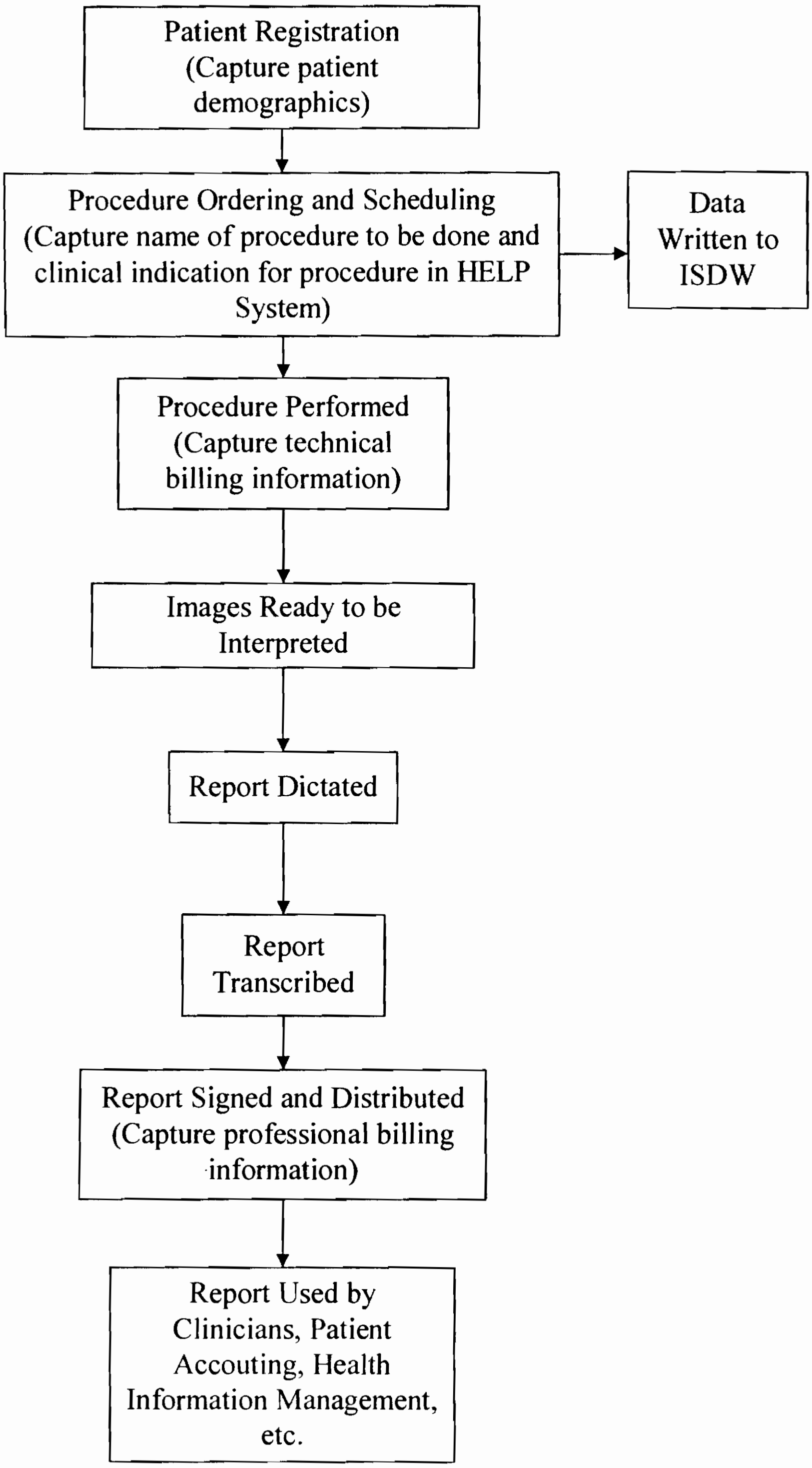


Figure 3. Processes for producing, capturing, storing, and utilizing imaging services data.

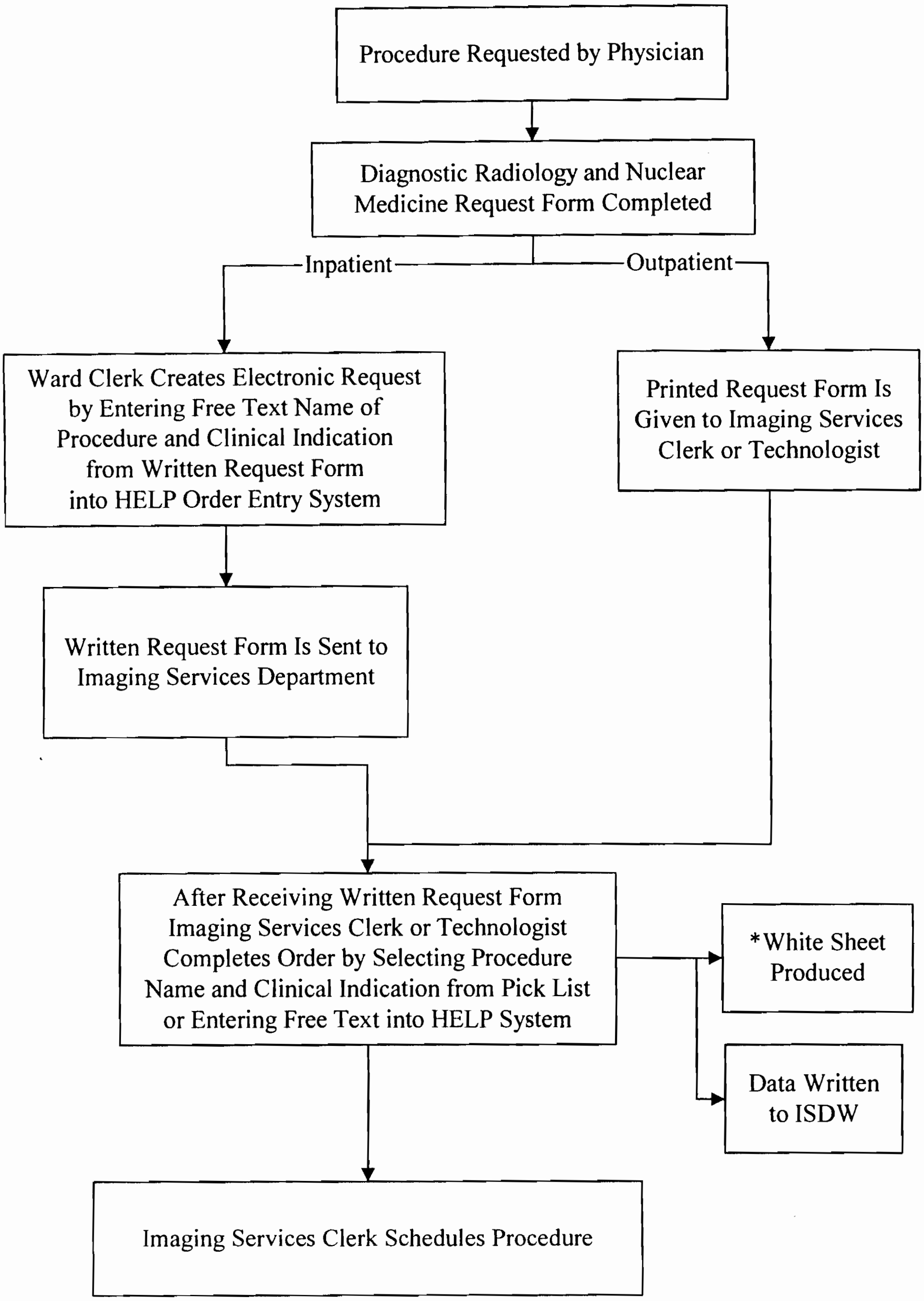


Figure 4. Subprocesses for ordering an imaging procedure.
(* Defined in section on "Data Sources for Information Quality Assessment")

methodologies, as described in the "Research in Information Quality" section of Chapter I of this thesis, are intuitive, theoretical, and empirical.

The intuitive method, which focuses on the evaluation of the various dimensions of IQ, was incorporated by performing an evaluation of the IQ dimensions of accuracy, consistency, and completeness of imaging services information. Accuracy was defined as the information consumers' perception as to whether the information about the clinical indication for a procedure in the various steps of the imaging procedure ordering process was an accurate representation of the true clinical indication. The gold standard for the true clinical indication for the inpatient population was the indication documented by the ordering physician (MD) on the original request form. For the outpatient population, the gold standard for the true indication was the first documented source of information accessible for review, which was the clinical indication documented by the imaging services scheduler (ISS). Thus, an example of evaluating for accuracy, in the inpatient population, is determining whether the clinical indication for the procedure, as captured in the ward clerk free text or the completed order in the HELP System, was the same as the clinical indication written on the original request form submitted by the requesting physician. The dimension of consistency was defined as the information consumers' perception as to whether the information on the clinical indication for a procedure, for a specific patient, in each data source, was consistent with the clinical indication in all the other data sources. This is an extension of the dimension of accuracy. Thus, in addition to determining whether the clinical indication as captured in the ward clerk free text or the completed order in the HELP system was consistent with the written request form (definition of accuracy dimension), a determination also was made as to whether the clinical indications, as captured in the ward clerk free text and the completed order in the HELP system, are consistent with each other and all other data sources evaluated in this research. The dimension of completeness was defined as the presence of a clinical indication in each step of the information flow in the ordering process for an imaging procedure. Thus, if the clinical indication for a procedure was present on the written request form, but not present in the HELP System, this was classified as incomplete information.

The theoretical method, which focuses on how data become deficient, was incorporated by investigating the process for ordering imaging procedures and the associated flow and use of imaging services information. The specific method used was open-ended interviews modeled after the techniques of qualitative research as described by Kaplan and Maxwell[163] and Friedman and Wyatt.[161] The people interviewed included producers, custodians, and consumers of imaging services information. In some cases, a particular person may be a producer as well as a consumer of information. For example, a physician requesting an imaging study produces information by providing the clinical indication for the study being requested. The same requesting physician is a consumer of the information in the radiologist's report of the results of the imaging study. The interviews included 8 information producers, 2 information custodians, and 14 information consumers.

The empirical method focuses on identifying quality attributes important to information consumers and analyzing data collected on these quality attributes to determine how information consumers assess the fitness of data for use. The empirical method was incorporated into this thesis research by focusing, in the interviews described above and the quantitative analysis of the data, on the information consumers' perception of the quality of the processes and the information. In their TDQM work, Wang and Strong stated the advantage of the empirical method in information quality research is that it captures the "voice of the consumer."[98] Friedman and Wyatt supported the statement of Wang and Strong, when they suggested that one of the key differences between a medical informatics evaluation study, and what is more commonly thought of as "traditional" medical informatics research is the focus of the questions being asked. In "traditional" research, the focus of the questions is on what the "researcher" wants to know. In contrast, in evaluation studies, the focus is on what the "customer (consumer)" wants to know.[161] In the case of this thesis, the questions to which answers were being sought were based on questions posed by imaging services information consumers. The information consumers wanted to know if the information they were using to perform their work and make decisions was accurate. The consumers (who also were producers in some cases) of the imaging services information included two radiology clerks, two radiology technologists, two radiologists, two referring physicians, one medical record coder, two patient account representatives, and three billing clerks/coordinators.

In addition to the primary evaluation in this thesis, which focused on the sources, nature, volume, and information consumers' perception of imaging services information quality problems, a secondary study was performed, which focused on the impact of imaging services IQ problems on the quality of imaging study reports.

In order for imaging services to provide added value to the process of clinical care, radiologists must perform a number of tasks. The first step in the radiologist's process of adding value to an individual patient's care is helping the referring clinician determine the appropriate imaging or interventional procedure for the patient's clinical condition and the information being sought. Once the appropriate procedure is identified and ordered, the radiologist must provide information about how best to prepare the patient for the procedure. When the patient arrives for the procedure, the radiologist must determine the most appropriate technique to use in performing it. After the procedure is complete, the radiologist performs two primary interpretation tasks. The first task is perceptual. The radiologist must identify all abnormalities present. The second task is "cognitive." For the cognitive task, the radiologist must make a decision about the diagnostic and therapeutic relevance of the abnormalities detected.[164] The majority of evidence has shown that each of these tasks is performed more effectively and efficiently if the radiologist is provided with timely, accurate, and relevant clinical and historical information about the patient.[164-169]

The secondary study was a pilot study and was intended to set the stage for future evaluations of the impact of the quality of imaging services information on its downstream uses in clinical care.

Study Site

The study was performed in the Imaging Services and Medical Informatics Departments at the LDS Hospital in Salt Lake City, Utah, from July 1997 to August 1997. The hospital is a 520-bed, private, tertiary care facility of Intermountain Health

Care (IHC), an integrated delivery system providing services in Utah, Idaho, and Wyoming. The LDS Hospital also is a major teaching center associated with the University of Utah School of Medicine. Complete inpatient and outpatient imaging services, including plain film radiography, ultrasonography, computed tomography, magnetic resonance imaging, nuclear medicine, and interventional procedures, are performed at LDS Hospital. At the time of the study performed for this thesis, the imaging services department was staffed by a group of 13 radiologists, 60 technologists, and 19 office personnel. Of the 19 office personnel, 4 were involved in ordering and scheduling imaging procedures. The annual volume of imaging procedures was approximately 110,000.

The Department of Medical Informatics at LDS Hospital was co-led by two informaticists with well over 30 years of combined experience in the field. The department is staffed by a combination of trained informaticists, computer programmers, systems analysts, and medical informatics graduate students. Many of the personnel have clinical medical backgrounds. The LDS Hospital Department of Medical Informatics routinely participates in ongoing research in the field of informatics and the development and support of clinical systems required for the daily processes of care delivery.

Patient Population

Primary Study

All patient-related information used in the study was obtained from paper-based medical records, the HELP System, the IHC Hospital Casemix System, and the

ISDW (a clinical data mart designed for quality assurance and improvement work). Institutional Review Board approval was obtained from both LDS Hospital and the University of Utah. No direct patient participation was required for this study, nor was individually identifiable patient information available to anyone involved, except for the primary investigator. The only criterion for patient selection was having an imaging procedure (inpatient or outpatient) at LDS Hospital during the time period of the evaluation. There were no age-, gender-, diagnosis-, or examination-related inclusion or exclusion criteria.

Fifteen inpatients and 15 outpatients were randomly selected each week, over a four-week period, by a clerk in the imaging department billing office as she performed her daily work. The clerk used the imaging procedure request forms submitted for billing as the source of patient selection. At the end of each day, after processing all procedure request forms for the day, the billing clerk randomly selected request forms for three inpatients and three outpatients to be included in the thesis research study. One-hundred-twenty patients were selected during the four-week period. Six of the patients were excluded because their exams were never performed, and thus the full compliment of information required for the study was not available. Therefore, the final study population included 114 patients (66 inpatients, 48 outpatients). The population consisted of 45 males and 69 females. In the inpatient population there were 33 males and 33 females. In the outpatient population there were 12 males and 36 females. The distribution of inpatients and outpatients does not add up to 60 in each group, as would be expected from the random selection process described above. The reason for this discrepancy was the initial misidentification, on the imaging request form, of a number of inpatients as outpatients. The age range of the combined inpatient and outpatient populations was 0 (newborn) to 93 years (average, 54 years, standard deviation, 21 years). The age range in the inpatient population was 18 to 92 years (average, 55 years, standard deviation, 22 years) and in the outpatient population was 0 to 93 years (average, 53 years, standard deviation, 19 years).

In the 114 patients included in the evaluation, 133 imaging procedures (81 inpatient, 52 outpatient) were performed. The type of imaging procedures included plain film radiography, computed tomography (CT), ultrasound (US), nuclear medicine (NM), magnetic resonance imaging (MRI), fluoroscopic studies such as upper gastrointestinal series (UGI) and barium enemas (BE), mammograms, and interventional procedures. The specific procedures and volumes are detailed in Table 1.

Secondary Study

The secondary study population included 40 patients (15 males, 25 females). After completion of the primary study, one group of 20 inpatients and one group of 20 outpatients were randomly selected from the primary study patient population. Each group of 20 patients consisted of two subgroups of 10 patients. In one subgroup of 10 patients, all of the raters scored (see explanation of scoring below) the clinical indication available to the radiologist at the time of interpretation, as being the same as the true indication for the procedure. In the other subgroup of 10 patients, all of the raters scored the clinical indication available to the radiologist as being different

Table 1

Primary Study Procedures and Volumes

Procedure Type	Outpatient	Inpatient
Abdomen Plain Films	2	4
Barium Enema	1	-
Chest X-Ray	4	40
Cholangiogram	-	1
C-Spine Plain Films	1	-
CT Abdomen	1	5
CT Brain	1	2
CT Cervical Spine	-	1
CT Face	3	-
CT Neck	1	-
CT Paracentesis	-	1
CT Pelvis	1	3
CT-Guided Abdominal Drainage	-	1
Face Plain Films	-	1
Feeding Tube Placement	1	3
Foot Plain Films	1	-
Hip Plain Films	-	1
Hysterosalpingogram	1	-
Inferior Vena Cava Filter	-	1
Intravenous Urogram	4	1
Knee Plain Films	1	2
Lumbar Spine Plain Films	1	1
Mammography	10	-
MRI Brain	1	1
MRI Lumbar Spine	2	-

Table 1. Continued

Procedure Type	Outpatient	Inpatient
Nephrostogram	-	1
NM Bone Scan	2	-
NM Brain Scan	1	-
NM Heart Perfusion Study	1	1
NM Lung Scan	2	-
NM Thyroid Therapy	-	1
Orbits Plain Films	-	1
Ribs Plain Films	-	1
Sacro-Iliac Joint Plain Films	1	-
Sinuses Plain Films	1	-
Skull Plain Films	-	1
Sternum Plain Films	-	1
Thoracic Spine Plain Films	-	1
UGI	2	-
US Abdomen	3	2
US Head (Neonatal)	-	1
US Obstetrical	2	-
Wrist Plain Films	-	1
Total Number of Exams	52	81

from the true indication. In the inpatient population there were 9 males and 11 females. In the outpatient population there were 6 males and 14 females. The age range of the total population was 19 to 92 years (average, 57 years, standard deviation, 20 years). The age range in the inpatient population was 19 to 92 years (average, 61 years, standard deviation 22 years), and in the outpatient population was 20 to 86 years (average, 53 years, standard deviation, 18 years). Forty imaging procedures (20 inpatient, 20 outpatient) were performed on the 40 patients included in the secondary study. The specific procedures and volumes are detailed in Table 2.

Qualitative Analysis Data: Process Mapping of Imaging Service Processes and Information Flow

The first step required in this thesis research was to understand the components of the LDS Hospital processes for ordering, performing, and interpreting an imaging procedure; the way in which data flowed to support the steps in these processes; the downstream uses of the data captured during these processes; and the processes for the downstream uses of the data. The reason for needing to understand the downstream uses of (and the processes for using) the data captured during the ordering, performing, and interpreting processes for imaging procedures was to support demonstration of an example of the impact of imaging services information quality problems on healthcare service delivery. Understanding the data flow included understanding the way in which the procedure type and the clinical indication for the imaging procedure were captured and encoded. Investigation of the use of the information and impact of the information quality problems focused

Table 2

Secondary Study Procedures and Volumes

Procedure Type	Outpatient	Inpatient
Chest X-Ray	1	13
CT Abdomen	-	1
CT Brain	-	1
CT Pelvis	-	1
CT Thoracic Spine	-	1
Feeding Tube Placement	-	1
Intravenous Urogram	1	-
Lumbar Spine Plain Films	1	-
Mammography	6	-
MRI Brain	1	-
Nephrostogram	-	1
NM Bone Scan	2	-
NM Lung Scan	1	-
NM Thyroid Therapy	-	1
UGI	2	-
US Abdomen	2	-
US Breast	1	-
US Obstetrical	2	-
Total Number of Exams	20	20

qualitatively on billing processes and quantitatively on interpretation of the imaging studies. The information pertinent to the impact investigation was the procedure type (CPT-4 code) and clinical indication for the procedure (ICD-9-CM code).

Open-ended interviews were conducted by the principal investigator to acquire the information necessary for understanding each of the issues mentioned above in both the inpatient and outpatient setting. Interviews varied in length from 30 to 60 minutes. The interviews were conducted with two referring physicians; two radiologists; two radiology department schedulers; one radiology department supervisor; two radiology department staff technologists; one radiology department billing coordinator; one radiology department billing clerk; the radiology practice billing coordinator; the LDSH Patient Account Services group leader and service representative; the LDSH Health Information Systems department coding supervisor; the LDSH Hospital Casemix System manager; an LDSH Department of Medical Informatics programmer responsible for the HELP System radiology applications and the ISDW; and the imaging services systems analyst from the IHC Information Systems department. In the interviews, a range of questions were asked about the ordering processes, the flow and use of the information, the people involved in the ordering and information use processes, perceived information quality problems, and possible sources of those problems. Some of the specific questions included the following:

1. What is the process for ordering an imaging procedure?

2. What is the process for capturing and encoding, in the HELP System, the type of imaging procedure (CPT-4 code), and clinical indication (ICD-9-CM code) for performing an imaging procedure?
3. What is the process for moving the encoded clinical indication from the HELP System into the ISDW?

The respondents were allowed to reply in any way and length in which they chose. The information acquired in the interviews of the imaging services information producers, custodians, and consumers was captured in handwritten notes. The notes were transcribed into textual narratives and used to create flow charts showing the steps in the ordering, performing, interpreting and reporting, and billing processes, including the flow of information. One flow chart was made for each of the processes. The flow charts represented a composite of the process steps reported by each interviewee.

Data Sources for Quantitative Information Quality Assessment

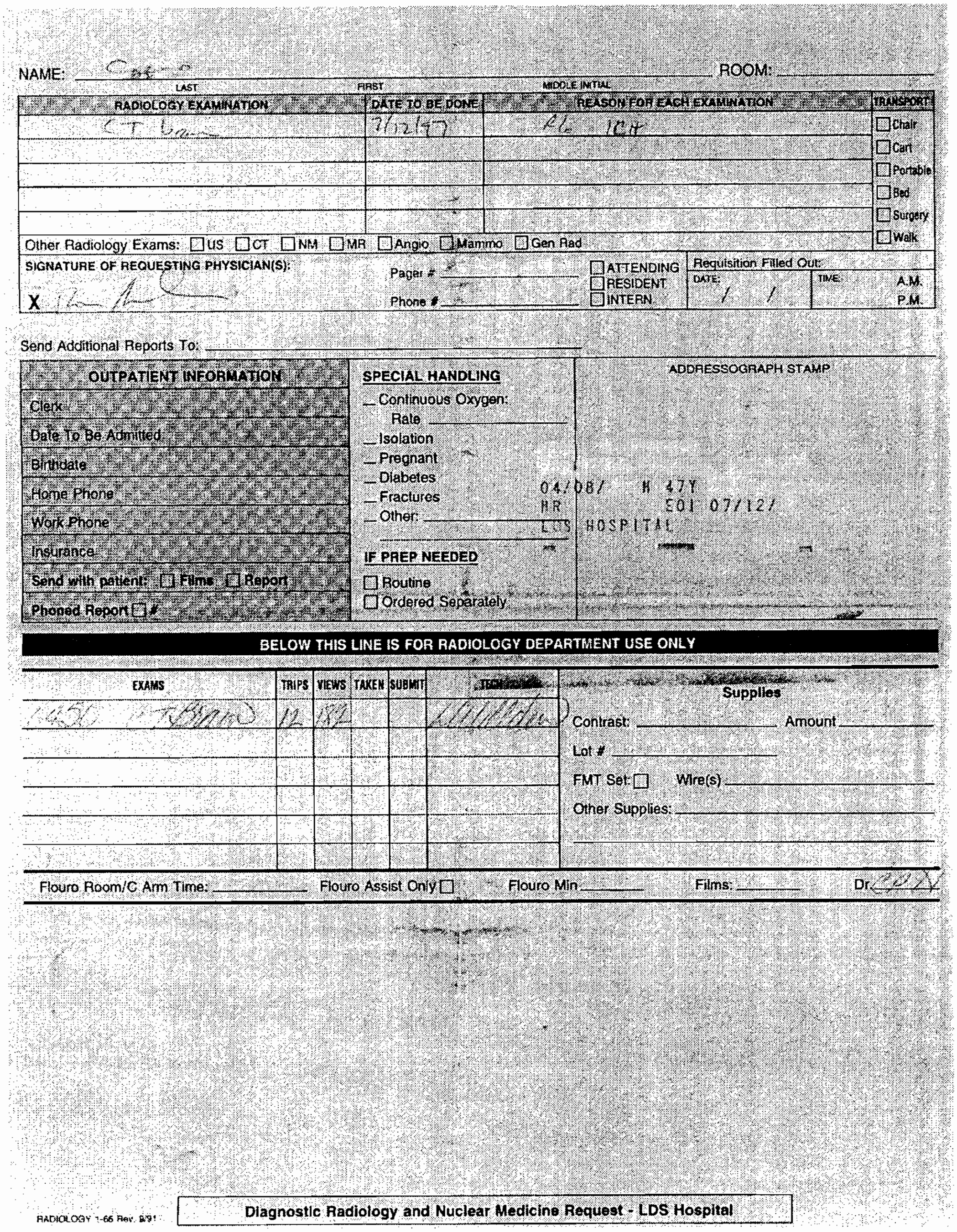
The second step required in the thesis research was to collect the data that would be used in the quantitative information quality assessment. The quantitative assessment was focused on the information about the clinical indication for the imaging procedure. The results of the interviews described above were used to determine the data sources and elements required for identifying and quantifying the intrinsic, contextual, and representational information quality problems, pertaining to the clinical indication for the procedure, in the HELP System and the ISDW. The documents and systems described below were the sources of the data used for the assessment.

Diagnostic Radiology and Nuclear Medicine Request Form

As flow charted in Figure 4, the imaging procedure ordering process required the completion of a "Diagnostic Radiology and Nuclear Medicine Request" form (hereafter, also referred to as the "imaging services request form" or simply the "request form"). The information provided on the request form includes patient name, requesting physician name, attending physician name (which may or may not be the same as the requesting physician), procedure being ordered, and clinical indication for the procedure. Figure 5 and Figure 6 are examples of the forms used for inpatient and outpatient requests, respectively. The information on the Diagnostic Radiology and Nuclear Medicine Request form is used by the ward clerk to enter the electronic procedure request and by the imaging services clerk or technologist to complete the electronic order for an imaging procedure. The request form also is used by the technologist as part of the documentation of the performance of the procedure and by the radiologist in the interpretation of the procedure.

HELP System

General information. The Health Evaluation through Logical Processing Hospital Information System was developed at LDS Hospital over the course of the last 35 years.[170] The core of the HELP System is a comprehensive, integrated, time-oriented, patient-focused clinical database. The database receives information from pharmacy, laboratory, surgery, radiology, intensive care, nursing, and other clinical



NAME: ______ LAST ______ FIRST ______ MIDDLE INITIAL ______ ROOM: ______

RADIOLOGY EXAMINATION	DATE TO BE DONE	REASON FOR EACH EXAMINATION	TRANSPORT
CT	7/12/97		☐ Chair
			☐ Cart
			☐ Portable
			☐ Bed
			☐ Surgery
			☐ Walk

Other Radiology Exams: ☐ US ☐ CT ☐ NM ☐ MR ☐ Angio ☐ Mammo ☐ Gen Rad

SIGNATURE OF REQUESTING PHYSICIAN(S): X ______ Pager # ______ Phone # ______ ☐ ATTENDING ☐ RESIDENT ☐ INTERN

Requisition Filled Out: DATE: / / TIME: A.M. P.M.

Send Additional Reports To: ______

OUTPATIENT INFORMATION

Clerk

Date To Be Admitted

Birthdate

Home Phone

Work Phone

Insurance

Send with patient: ☐ Films ☐ Report

Phoned Report ☐ #

SPECIAL HANDLING

_ Continuous Oxygen:

Rate ______

_ Isolation

_ Pregnant

_ Diabetes

_ Fractures

_ Other: ______

IF PREP NEEDED

☐ Routine

☐ Ordered Separately

ADDRESSOGRAPH STAMP

04/08/ M 47Y

HR E01 07/12/

LDS HOSPITAL

BELOW THIS LINE IS FOR RADIOLOGY DEPARTMENT USE ONLY

EXAMS	TRIPS	VIEWS	TAKEN	SUBMIT	TECH
1450 CT Brain	12	182			

Supplies

Contrast: ______ Amount ______

Lot # ______

FMT Set: ☐ Wire(s) ______

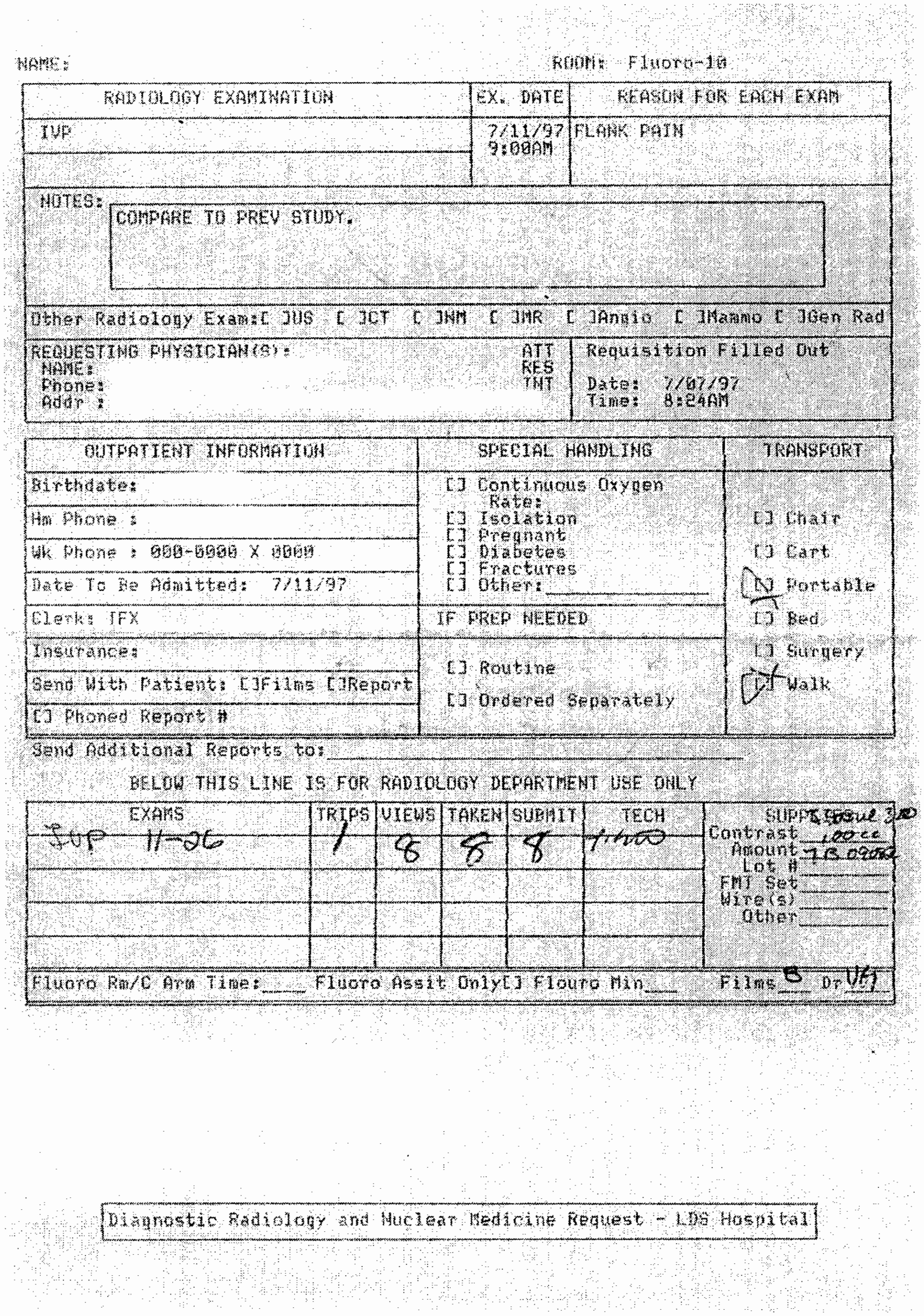
Other Supplies: ______

Flouro Room/C Arm Time: ______ Flouro Assist Only ☐ Flouro Min ______ Films: ______ Dr. ______

RADIOLOGY 1-66 Rev. 9/91

Diagnostic Radiology and Nuclear Medicine Request - LDS Hospital

Figure 5. Inpatient imaging procedure request form.



NAME: ROOM: Fluoro-10

RADIOLOGY EXAMINATION	EX. DATE	REASON FOR EACH EXAM
IVP	7/11/97 9:00AM	FLANK PAIN

NOTES: COMPARE TO PREV STUDY.

Other Radiology Exam:[ ]US [ ]CT [ ]NM [ ]MR [ ]Angio [ ]Mammo [ ]Gen Rad

REQUESTING PHYSICIAN(S): ATT RES INT
NAME:
Phone:
Addr :

Requisition Filled Out
Date: 7/07/97
Time: 8:24AM

OUTPATIENT INFORMATION	SPECIAL HANDLING	TRANSPORT
Birthdate:	[ ] Continuous Oxygen Rate:	
Hm Phone :	[ ] Isolation	[ ] Chair
Wk Phone : 000-0000 X 0000	[ ] Pregnant [ ] Diabetes	[ ] Cart
Date To Be Admitted: 7/11/97	[ ] Fractures [ ] Other:______	[X] Portable
Clerk: TFX	IF PREP NEEDED	[ ] Bed
Insurance:		[ ] Surgery
Send With Patient: [ ]Films [ ]Report	[ ] Routine	[X] Walk
[ ] Phoned Report #	[ ] Ordered Separately	

Send Additional Reports to:______

BELOW THIS LINE IS FOR RADIOLOGY DEPARTMENT USE ONLY

EXAMS	TRIPS	VIEWS	TAKEN	SUBMIT	TECH	SUPPLIES
IVP 11-26	1	8	8	8	[illegible]	Contrast [illegible] Amount [illegible] Lot # FMI Set Wire(s) Other

Fluoro Rm/C Arm Time:____ Fluoro Assit Only[ ] Flouro Min____ Films 8 Dr [illegible]

Diagnostic Radiology and Nuclear Medicine Request - LDS Hospital

Figure 6. Outpatient imaging procedure request form.

areas. The HELP System also contains a knowledge-base consisting of frame-based medical decision support modules (computer programs), the tools necessary for maintaining and continually growing these modules, and data- and time-drive applications for activating the modules.[171-175] The data-drive and time-drive applications automatically activate the knowledge-base modules when predefined data are stored in a patient record or specific "time triggers" are reached. Figure 7 is a schematic drawing of the HELP System.

Because the HELP System was designed with a focus on supporting clinical care processes, functionality for supporting administrative and financial aspects of the delivery of care are less well developed. To address these deficiencies, the HELP System has been interfaced with multiple other systems at LDS Hospital. Figure 8 shows the information systems with which the HELP System is interfaced.

Data dictionary. The data dictionary for the HELP System consists of a hierarchical representation of data elements. Each data element is identified by an 8-byte code known as PTXT, an acronym for Pointer-to-Text. Each PTXT code is linked to a textual representation of the data element (medical concept).[173,176] Clinical and patient demographic data are stored in the HELP System patient database as PTXT codes. The first byte in the PTXT code represents the data class (DC). Data classes correspond to different types of patient data. Imaging services data are represented by data class 20. Medical records data (e.g., admit and discharge diagnoses ICD-9-CM codes) are represented by data class 24.

The PTXT codes associated with diagnoses in data classes 20 and 24 have an associated code and textual description of the diagnoses from the *International*

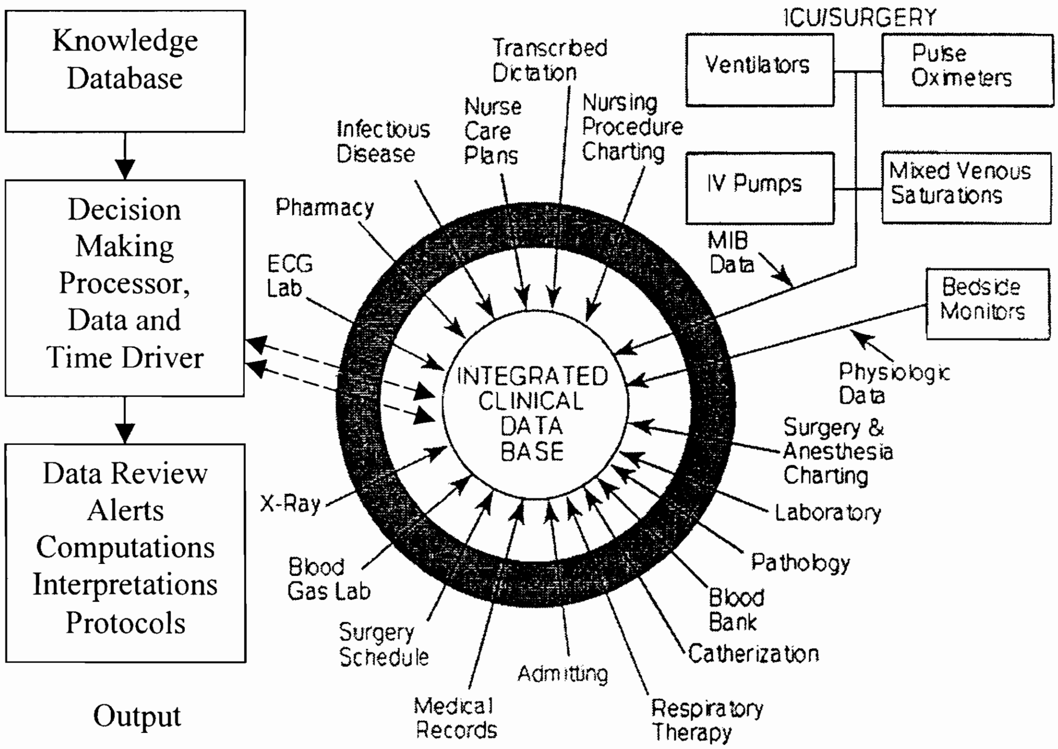


Figure 7. Schematic block diagram of the HELP System. The central database is shown in the middle. Data flow from many clinical data sources is shown by inward pointing arrows. As the data flows into the database, it passes through a dark "stippled" area. This stippled area is representative of the "data drive" capability of the system. As data flows in from the various sources, knowledge from the various medical decision support modules is applied to the data. (Reprinted with permission from reference 173)

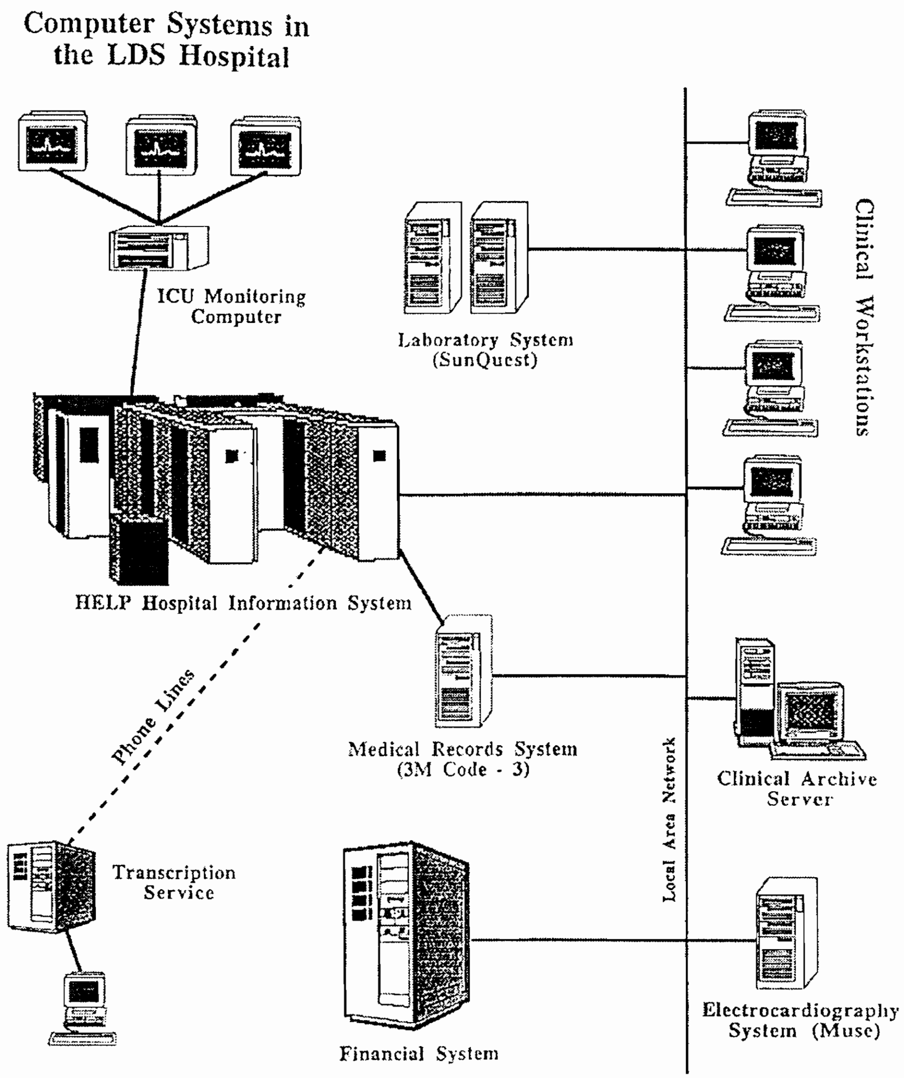


Figure 8. Information systems interfaced with the HELP System at LDS Hospital. (Reprinted with permission from reference 173)

Classification of Disease, 9th Revision, Clinical Modification. However, the textual description of the same diagnosis, differs between data class 20 and data class 24. The textual description in data class 24 is identical to the description in the ICD-9-CM code book. The textual description in data class 20 is a simplified version of the textual description from the ICD-9-CM code book. The simplified text was intended to facilitate use of these textual descriptions by people (e.g., radiology clerks) who were not specifically trained in medical terminology or coding practices.

Radiology information system (RIS). Embedded in the HELP System is a radiology information subsystem. The RIS is used primarily in the Imaging Services Department and provides functionality for order entry, charge capture, film tracking, and process tracking, such as the time exams are started and completed. In the Imaging Services Department, in addition to the RIS workstations for clerks and technologists, there also are radiologist workstations used for managing dictation, for signing of imaging procedure reports, and for accessing clinical data.[172,176]

White Sheet

As described in the ordering process flow chart (Figure 4), after the imaging services clerk or technologist completes the order, a sheet of paper with the information from the completed electronic order is printed from the HELP System. This piece of paper is referred to as the "white sheet." Figure 9 is an example of the "white sheet." The white sheet includes demographic information on the patient, information on the requesting and attending physician, information on patient location and transportation requirements, the free text indication for the procedure as

ROOM: ER BD: 47M
RAD#:
Request:
Attend:
SEQ # CODE EXAMINATION
* 182 0450 CT Brain w/o Cnt
Req Reason
LOC ETOH RI ICH

DATE: 12 JUL 97
AD DATE: 12 JUL 97 17:56
HSP#:
CLERK: ESX
Ward Clerk: KN

BI997193.0182T

TRANSPORT: STRETCHER
Admit Dx:ADMITTING DIAGNOSIS LAC ON HEAD
REQ: ORD:

Ins:SELF PAY
HANDLING: NO SPECIAL HANDLING
APPT TIME: STAT
Ord.Reason:
780.0 LOSS OF CONSCIOUSNESS

CONFIRMED
** NEW **

Req Exam:
CT BRAIN

Figure 9. Inpatient HELP System white sheet.

entered when the order is requested, and the encoded name of and reason for the procedure as captured when the order is completed in the HELP System. The white sheet is attached to the Diagnostic Radiology and Nuclear Medicine Request form and is given to the imaging services technologist for use, along with the request form, in performing and documenting the imaging procedure. The white sheet also is used, in conjunction with the Diagnostic Radiology and Nuclear Medicine Request form, by the radiologist in the interpretation of the imaging procedure. In some cases, the Diagnostic Radiology and Nuclear Medicine Request form and the white sheet inadvertently become separated from one another, and only one of them is delivered to the radiologist for use in the interpretation of the imaging procedure.

Imaging Services Data Warehouse

The Imaging Services Data Warehouse is an independent database designed to capture and store coded information relevant to the processes of delivering imaging services. The information stored in the ISDW for each imaging study performed includes, but is not limited to, patient demographics, time of study request and performance, clinical indication for study, identification of person performing study, type of study performed, and number of films used. The ISDW was created by members of the LDSH Medical Informatics Department, LDSH Imaging Services Department, and IHC Information Services Department prior to initiation of this thesis research. The purpose of the ISDW was to provide a central source for data that could be used to perform imaging services quality assurance and quality improvement work. The primary targets of the quality assurance and improvement work were processes within the radiology department, the ordering practices of staff physicians who use imaging services, and the quality of the products of the radiology department (e.g., the radiologists' report).[172]

The ISDW was built using the Oracle 7 relational database management system (Oracle Corp, Redwood Shores, California). Information in the ISDW comes primarily from the HELP System. Using the data-drive capability of the HELP System, whenever a string of radiology data (PTXT data class 20) is written to the native HELP database, at the same time, a copy of that data string also is sent to the ISDW. This same mechanism is used to capture nonimaging services data such as biopsy results, discharge diagnoses, and microbiology results on patients who have had imaging procedures (Figure 10).

A small subset of the data in the ISDW does not come directly from the HELP System. An example of the data not coming from the HELP System is coded data on the radiologists' impression as to whether the findings on the imaging procedure are consistent or inconsistent with the clinical indication for the exam. The coded data on the radiologists' impression about the consistency of the findings of the imaging study with the clinical indication for the study were used to assess the appropriateness of physician ordering patterns. Tools other than the HELP System data-drive were used to capture the coded impression data "directly" from the radiologists. The way in which the information from the radiologists was captured was through the use of a natural language parsing system that parsed the radiologists' dictation, looking for a code they dictated to indicate their impression of the appropriateness of the imaging procedure.

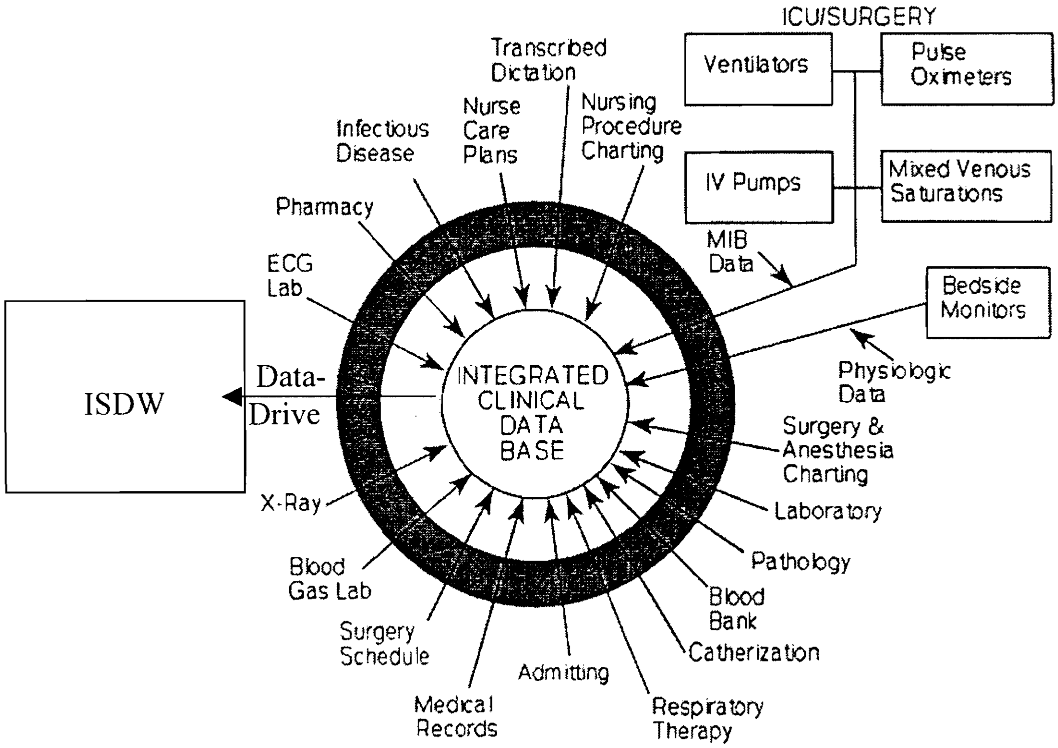


Figure 10. Relationship of the HELP System and Imaging Services Data Warehouse. The HELP System data-drive process is used to move imaging services data and other pertinent data such as biopsy results, discharge diagnoses, and microbiology results on patients who have had imaging procedures, from the HELP System into the ISDW. (adapted, with permission, from Figure 2 reference 173)

Appendix A is a high level entity-relationship diagram of the ISDW, and Appendix B is a summary of the data elements in the ISDW tables included in the entity-relationship diagram. At the time of completion of this thesis research (August 1997), work was in progress to integrate the ISDW with the IHC Enterprise Data Warehouse (EDW).

Information in the ISDW could be accessed over the IHC intranet using any Open Database Connectivity (ODBC) compliant query tool. Client tools used include desktop database management systems, spreadsheet programs, Structured Query Language (SQL) writing tools, and on-line analytical processing query tools. Because the Imaging Services Data Warehouse contains individually identifiable patient data, people wishing to access the ISDW must have the appropriate security clearance.

Hospital Casemix System

The Casemix System is an application developed internally at Intermountain Health Care. It resides on the AS 400 platform and has an interface to the IHC Enterprise Data Warehouse for which the Casemix System provides data for populating several tables. At the time of this thesis research, the Casemix System was not used to populate tables in the ISDW. The source systems for the Casemix System are the HELP System, AS 400 billing system, and the 3M-Code 3 medical records coding system (3M Health Information Systems, Salt Lake City, Utah). Each IHC facility has its own Casemix database. The LDSH Casemix database is updated nightly from each of the source systems. On a monthly basis, data from the LDSH

Casemix files are extracted and integrated with Casemix data from the other IHC facilities, thus creating a common Casemix database for all sites. The common Casemix database is used as the source of data that is moved into the EDW.

Radiology Reports

After radiologists have studied and interpreted an imaging procedure, he or she dictate his or her findings and conclusions into a centralized dictation system. The dictation is transcribed into the HELP System by a member of the transcriptionist pool and stored as a preliminary report in a free text electronic format. The radiologist reviews the electronic preliminary report through the HELP System RIS, makes the necessary corrections on the computer, and then electronically signs the report. The finalized electronic free text report replaces the electronic preliminary free text report in the HELP System. A printed version of the finalized report is stored in the patients' paper record.

Data Elements for Quantitative Information Quality Assessment

Primary Study

Because the process mapping was divided into inpatient and outpatient processes, the data elements also were divided into inpatient and outpatient groups. The inpatient data element names, sources, and descriptions are presented in Table 3. The information on the outpatient data elements is presented in Table 4.

Table 3

Primary Study Inpatient Data Elements

Name	Source	Description
True clinical indication	Imaging procedure request form	Clinical indication for procedure as documented by ordering physician
Ward clerk free text clinical indication	HELP System (White sheet)	True clinical indication as transcribed into HELP System
Person completing order in HELP System	ISDW	Role of person completing order (Radiology clerk or technologist)
White sheet ICD-9-CM code	HELP System (White sheet)	ICD-9-CM code for clinical indication as recorded in HELP System
White sheet ICD-9-CM text	HELP System (White sheet)	Text associated with ICD-9-CM code recorded in HELP System
ISDW ICD-9-CM code	ISDW	ICD-9-CM code for clinical indication as recorded in ISDW
ISDW ICD-9-CM text	ISDW	Text associated with ICD-9-CM code recorded in ISDW

Table 4

Primary Study Outpatient Data Elements

Name	Source	Description
True clinical indication	Imaging procedure request form	Clinical indication for procedure as documented by imaging services scheduler
Person completing order in HELP System	ISDW	Job of person completing order
White sheet ICD-9-CM code	HELP System (White sheet)	ICD-9-CM code for clinical indication as recorded in HELP System
White sheet ICD-9-CM text	HELP System (White sheet)	Text associated with ICD-9-CM code recorded in HELP System
ISDW ICD-9-CM code	ISDW	ICD-9-CM code for clinical indication as recorded in ISDW
ISDW ICD-9-CM text	ISDW	Text associated with ICD-9-CM code recorded in ISDW
Primary admit diagnosis ICD-9-CM code	Casemix system	ICD-9-CM code for primary admit diagnosis as recorded in Casemix

Secondary Study

Three data elements were used in the secondary study. They were the true clinical indication, the clinical indication captured in the HELP System, and the report of the radiologists' interpretation of the imaging procedure findings.

Research Questions for Quantitative Information Quality Assessment

Primary Study

The information consumers' perception of the accuracy, consistency, and completeness of the information pertaining to the clinical indication for an imaging procedure was acquired by obtaining answers to the following questions:

1. Does the ward clerk's free text description of the clinical indication in the HELP System order request differ from the physician's documented (on an imaging procedure request form) clinical indication for ordering an imaging procedure?
2. Does the ward clerk's free text description of the clinical indication in the HELP System order request differ from the text description of the clinical indication in the completed HELP System order?
3. Does the text description of the clinical indication in the completed HELP System order differ from the physician's documented (inpatient imaging procedure request form) or imaging services scheduler's documented (outpatient imaging procedure request form) clinical indication for ordering an imaging procedure?

4. Does the ward clerk's free text description of the clinical indication in the HELP System order request differ from the text description of the clinical indication in the ISDW?
5. Does the text description of the clinical indication for an imaging procedure (inpatient or outpatient) differ in the HELP System completed order and the ISDW?
6. Does the text description of the clinical indication for an imaging procedure in the ISDW differ from the physician's documented (inpatient imaging procedure request form) or imaging services scheduler's documented (outpatient imaging procedure request form) clinical indication for ordering an imaging procedure?
7. Does the ICD-9-CM code for the clinical indication for an imaging procedure (inpatient or outpatient) differ in the HELP System completed order and the ISDW?
8. Is there missing data in each of the steps (inpatient or outpatient) from making the initial request for exam to the storage of the encoded clinical indication in the ISDW?

Secondary Study

Answers to the following questions, from the information consumer's perspective, were used to conduct the quantitative analysis for the secondary study:

1. Is the information the radiologist received prior to interpreting the results of the imaging procedure consistent with or different from the true clinical indication for the procedure?

2. Does the imaging procedure report produced by the radiologist provide the clinical information being sought (as defined by the true clinical indication for the imaging procedure)?

Collection and Aggregation of Data Elements for Quantitative Information Quality Assessment

Primary Study

To facilitate initial aggregation and subsequent manipulation of the data into subgroups for quality evaluation, a relational database was built using Microsoft Access (Microsoft Corporation, Redmond, Washington). In the Access database, two "master tables" were built, one for inpatients and one for outpatients. At the end of each week of the four-week data collection period, the imaging procedure request forms and white sheets for the 30 patients selected that week were given to the primary investigator. The patient demographics (patient account number, name, age, and gender), general procedure information (sequence number and time to be done), and "true" clinical indication for the exam were obtained from the imaging procedure request forms and manually entered into the master tables in the Access database. Also manually entered into the Access database for inpatients were 1) the ward clerk free text clinical indication, 2) white sheet ICD-9-CM code, and 3) white sheet ICD-9-CM text. For outpatients, the additional manually entered information included the ICD-9-CM code and ICD-9-CM text for the reason for admission obtained from the Casemix System.

The Access database was then attached to the ISDW using an ODBC connection. Using SQL queries, data from the ISDW tables were retrieved and integrated with the data that had been manually entered into the Access database. The data retrieved from the ISDW tables included the role (radiology clerk or technologist) of the person who completed the procedure order in the HELP System, the ISDW ICD-9-CM code, and the ISDW ICD-9-CM text.

The data subgroup tables in the Access database were built from the master tables using SQL queries. Each subgroup table was designed to support comparison of the data element pairs required to answer the research questions outlined earlier in this chapter. Tables 5 and 6 are lists of the inpatient and outpatient subgroup tables, respectively. The data from these tables were exported into Microsoft Excel files (Microsoft Corporation, Redmond, Washington), which were used to create score sheets for the information consumer experts to use for "scoring" data accuracy, consistency, and completeness in the subgroups. Figure 11 is an example of a data quality score sheet for the primary study.

Secondary Study

The patients and procedures included in the secondary study were identified after completion of the primary study. When the patients were identified, two subgroup tables, one inpatient and one outpatient, were created from the master tables in the Access database using SQL queries. The information included in these subgroup tables was patient demographics, the true clinical indication for the procedure, and the white sheet ICD-9-CM text for the indication for the procedure. Information from these tables was exported into Excel files in a manner similar to that described for the primary study. The Excel files were used to create score sheets

Table 5

Access Database Inpatient Data Subgroup Tables

Table Name
1. MD Documented Indication - Ward Clerk Free Text Indication
2. MD Documented Indication - White Sheet ICD-9-CM Text for Indication
3. MD Documented Indication - ISDW ICD-9-CM Text for Indication
4. Ward Clerk Free Text Indication – White Sheet ICD-9-CM Text for Indication
5. Ward Clerk Free Text Indication – ISDW ICD-9-CM Text for Indication
6. White Sheet ICD-9-CM Code for Indication - ISDW ICD-9-CM Code for Indication
7. White Sheet ICD-9-CM Text for Indication - ISDW ICD-9-CM Text for Indication

Table 6

Access Database Outpatient Data Subgroup Tables

Table Name
1. Imaging Services Scheduler Documented Indication - White Sheet ICD-9-CM Text for Indication
2. Imaging Services Scheduler Documented Indication - ISDW ICD-9-CM Text for Indication
3. White Sheet ICD-9-CM Code for Indication – ISDW ICD-9-CM Code for Indication
4. White Sheet ICD-9-CM Text for Indication – ISDW ICD-9-CM Text for Indication
5. White Sheet ICD-9-CM Code for Indication – Admit Diagnosis ICD-9-CM Code for Indication

INPATIENT - WARD CLERK FREE TEXT TO WHITE SHEET ICD TEXT.xls

	A	B	C
1	WARDCLERK FTXT	WHITE SHEET ICD TEXT	SCORER #
2	F/U ASPIRATION LAKE POWELL WATER	ASPIRATION	S D M
3	STRICU ROUTINE AM CHEST	RESPIRATORY DISTRESS	S D M
4	POST-OP F/U	CHECK PROGRESS (POST OP)	S D M
5	F/U	MASTECTOMY F/U	S D M
6	RESP DISTRESS	RESPIRATORY DISTRESS	S D M
7	PREOP	PRE OP	S D M
8	SEPTIC SYNDROME, R/O LOWER ISCHEMIA, ALSO SEE ORDER	SEPTIC SYNDROME	S D M
9	R/O PNEUMONIA	PNEUMONIA	S D M
10	R/O LOWER ISCHEMIA, INTERCAR...HERNIA SEE ORDER FORM	LOWER ISCHEMIA	S D M
11	F/U INFILTRATES	INFILTRATES	S D M
12	RULE OUT PNEUMONIA	PNEUMONIA	S D M
13	F/U PULM FAT EMBOLI	EMBOLISM, PULMONARY	S D M
14	POST OP F/U	CHECK PROGRESS (POST OP)	S D M
15	POST OP F/U	CORONARY ARTERY DISEASE	S D M
16	POST OP AAA	CHECK PROGRESS (POST OP) S/P TCAG	S D M
17		OBSTRUCTION, URINARY	S D M
18		CHECK PROGRESS (POST OP)	S D M
19	R/O PN	PNEUMOTHORAX	S D M
20		PAIN, ABDOMINAL	S D M
21	ABD PAIN	AORTIC STENOSIS	S D M
22		DEGENERATED DISC DISEASE	S D M
23	POST OP F/U	CHECK PROGRESS (POST OP)	S D M
24		CAD	S D M
25	LINE PLACEMENT	CENTRAL LINE (TPN, CVP, SUBCLAVI)	S D M

Figure 11. Data quality score sheet for primary study.

for the secondary study. Figure 12 is an example of a data quality score sheet for the secondary study. Using the patient demographic and procedure information, printed copies of the imaging procedure final reports were obtained for each of the 40 procedures included in the analysis.

Scoring of Information Quality

Primary Study

The accuracy, consistency, and completeness of the clinical indication for the imaging procedure for each of the data element pairs in the data subgroups (Tables 5 and 6) was scored by three clinically experienced physicians. All of the physicians were board certified, and each practiced in a different specialty. The three physicians were a diagnostic radiologist (Scorer 1), a radiation oncologist (Scorer 2), and a general surgeon (Scorer 3). At the beginning of the scoring process, the raters were given a bound booklet that included a brief overview of the research project, written scoring instructions, and the complete set of inpatient and outpatient score sheets. Appendix C is a copy of the scoring instructions provided to the participants.

Scoring consisted of comparing the clinical information in each data element pair for a specific patient and procedure and determining whether the data elements in the pair were the same, different, or missing. The comparison was performed using the data quality score sheet (Figure 11). The raters were instructed to use a score of "S" if the information content in the data elements was the same, a score of "M" if no comparison could be made because there was no information present for one member

Inpatient Radiology Report Study

	A	B	C
1	PT ID	MD WRITTEN REASON	ANSWERS CLINICAL QUESTION
2		R/O INFILTRATE	Y N
3		POST OP F/U	Y N
4		F/U CP	Y N
5		F/U LUNG BX	Y N
6		RT ARM MYOPATHY AND TEAR ANUERYSM IN AXILLARY CUT	Y N
7		SEPTIC SYNDROME, R/O BOWEL ISCHEMIA, INCARCERATED HERNIA	Y N
8		SEPTIC SYNDROME, R/O BOWEL ISCHEMIA, INCARCERATED HERNIA	Y N
9		FOLLOW UP	Y N
10		POST-OP	Y N
11		F/U PQN	Y N
12		R/O INFILTRATE	Y N
13		POST OP F/U	Y N
14		POST OP	Y N
15		S/P CT DC'D	Y N
16		BACK TRAUMA	Y N
17		RT ARM WEAK	Y N
18		CHEST PAIN	Y N
19		THRYOID CA	Y N
20		S/P PACER INJECTION	Y N
21		NUTRITION	Y N

Figure 12. Data quality score sheet for secondary study.

of the data element pair, and a score of "D" if the information in the two data elements was different. Different was defined as a discrepancy between the information in the two components of the data element pairs that 1) described two different diagnoses or clinical problems, 2) would lead to a request for a different procedure for the two "different" clinical indications, or 3) would lead to two different interpretations of the imaging study findings.

No attempt was made to obtain consensus among the raters. There were two reasons for not seeking consensus: 1) Information quality was being measured from the perspective of the individual information consumer (each rater), and 2) one of the measurements was the level of interrater agreement on the quality of the information.

Secondary Study

The information being sought in the scoring for the secondary study was whether the information in the radiologists' reports provided the clinical information being sought by the requesting physician. At the time the secondary study was performed, two of the three original raters were unable to participate; thus scoring for the secondary study was performed by only one rater, the diagnostic radiologist. This was acceptable to the primary investigator because the secondary study was considered a pilot for future studies. The single rater was provided with the "true" clinical indication for the procedure and a copy of the final report for each procedure in the inpatient and outpatient groups. The instructions were to read the body of the report and the impression and use the score sheet (Figure 12) to provide a yes or no answer to the question of whether the information in the report answered the clinical question being asked.

Statistical Analysis of Quantitative Information Quality Assessment

Primary Study

The unit of analysis was the number of procedures. The data were analyzed using the SAS Statistical Program, Version 6.12 (SAS Institute, Cary, North Carolina). The results of the information quality scoring were evaluated separately for the inpatient and outpatient populations.

The scores were used to calculate the proportion of responses in each of the response categories (same, different, and missing), and 95% confidence intervals around these proportions. The confidence intervals are denoted by the upper confidence limit (UCL) and the lower confidence limit (LCL). A decision was made to calculate confidence intervals to demonstrate the relative precision of the proportions, instead of performing statistical tests to determine the significance of differences between the proportions. The basis for this decision was a fundamental assumption that any error of commission ("different" category) or omission ("missing" category) has potentially significant implications in the delivery of care; therefore, the presence or absence of statistical significance of any differences in frequency between the presence of good quality ("same" category) and poor quality ("different" and "missing" categories) information, within or between data element comparisons, is irrelevant. Furthermore, although confidence intervals are not a statistical test, if the need should arise, it is possible to infer, on the basis of overlap between confidence intervals and the inclusion or exclusion of the sample proportion in the overlapping segments of the confidence intervals, whether two results are likely to be truly different values.[177]

The proportions and confidence intervals were calculated for each of the three raters individually for each of the data element pairs listed in Tables 5 and 6. In addition, composite proportions and associated 95% confidence intervals were calculated for the three raters combined. The composite proportions were defined as the majority response for the three raters. Thus, if two raters scored the data elements as missing and one scored them as different, the composite score was recorded as "missing." Cases in which the three raters each selected a different response were removed from the composite analyses. The reason for calculating the composite proportions and associated confidence intervals was to enable the presentation of the scoring data in an easily viewed summary format that still would convey the information contained in the results.

The individual rater proportions of responses and 95% confidence intervals in each category for each data element pair were entered into three-by-three contingency tables. Using the data in the contingency tables, analysis of the interrater agreement on the scoring of the quality of the data was performed using the Kappa statistic. The Kappa statistic is a measurement of the observed agreement between raters above that expected by chance alone.[178] The probability of agreement between raters purely due to chance is designated by the notation P_C. The value of Kappa ranges from $-P_C/(1-P_C)$ to one. A value of zero indicates "pure chance" agreement and a value of one indicates perfect "true" agreement. Landis and Koch[179] provided the following guidelines for the interpretation of Kappa: Kappa greater than 0.75 denotes excellent agreement, Kappa greater than 0.4 and less than 0.75 denotes good agreement, and Kappa greater than zero and less than 0.4 denotes marginal agreement. Although Kappa is used as a measure of the probability of agreement above that resulting from chance, a Kappa less than zero indicates that not only is any agreement due to chance alone but also suggests that any observed disagreement between the raters is "true"; in other words there is a causal factor, other than chance, for the disagreement.[178]

Although the Kappa statistic generally is used to compare agreement between two raters, Fleiss[180] and Light[181] have described methods for computing an "overall" Kappa statistic for greater than two raters. The Kappa values reported in the results section of this thesis were computed as an "overall" Kappa, using the methods of Fleiss and Light, for all three raters who participated in the scoring of the data quality.

Two overall Kappa statistics were computed, one using all three categories (same, different, and missing) and one using only the same and different categories. The two Kappas are represented as Kappa 1 and Kappa 2, respectively. The reason the two Kappas were computed is because all three raters generally agreed quite well on the frequency of missing data, thus suggesting it is relatively "straightforward" to determine if data is missing. Based on the observation of the general agreement in the missing category, it was anticipated that the kappa including the missing category responses would artificially increase the level of overall agreement and, thus, mask the level of agreement between the raters in the same and different categories (i.e., frequency of accurate or consistent information and frequency of errors of commission). The null hypothesis being tested by the use of the Kappa statistic was that the agreement between the raters occurred purely by chance. The alternative hypothesis was that any observed agreement between the raters is not purely due to chance, and there is true consistency in their perception that there are information quality problems.

In addition to evaluating interrater agreement, the symmetry of the data in the contingency tables was evaluated. Symmetry in this context refers to the manner in which raters disagreed with each other. For the same reason two Kappa statistics were calculated; two tests of symmetry were conducted for each data element comparison. One test was conducted using all three response categories, and one test was conducted using only the same and different response categories. The analyses of symmetry were conducted using McNemar's test of symmetry and Bowker's extension to McNemar's test of symmetry.[182] Bowker's was used for the analyses including all three response categories, and McNemar's was used for the analyses including only the same and different response categories. Unlike the "overall" Kappa statistic, there is no methodology for calculating "overall" McNemar or Bowker's statistics. Therefore, the results of the Bowker's extension and the McNemar's statistics are reported for each pair of raters (i.e., rater 1 and rater 2, rater 1 and rater 3, rater 2 and rater 3). The null hypothesis being tested by the use of the McNemar and Bowker's statistics is that the disagreement between the raters occurred purely by chance. The alternative hypothesis is that the disagreement between raters is not due to chance, but rather is due to characteristics of the raters or some other influencing factor.

For the data element pairs that were present in both the inpatient and outpatient populations, an additional analysis was performed. The analysis was performed using Fisher's Exact Test for comparison of two proportions. The Fisher's Exact Test was used to determine the presence of association for the way in which the raters scored the data, and the inpatient or outpatient nature of the procedure, and thus the process for generating and capturing the data The null hypothesis for this analysis is that there is no association between the variables. The Kappa, McNemar's, Bowker's, and Fisher's Exact statistics were tested for significance at a level of $\alpha = 0.05$.

Secondary Study

The null hypothesis (H_0) being tested in the data analysis for the secondary study was that the frequency with which imaging study reports include the answer to the specific clinical question being asked is the same when comparing cases in which the information the radiologist receives is consistent with the true clinical indication for the procedure and those in which it is not consistent. The alternative hypothesis (H_1) was that the frequency with which imaging study reports include the answer to the specific clinical question being asked is significantly different when comparing cases in which the information the radiologist receives is consistent with the true clinical indication for the procedure and those in which it is not consistent. The hypothesis testing was performed by determining the presence or absence of an association between the raters' perception of the accuracy of information available to the radiologists at the time of interpretation of an imaging procedure and the raters' perception of whether the interpretation in the report provided the clinical information being sought. The analysis was performed using two-by-two contingency tables and the Fisher's Exact Test. The results of the quality scoring from the secondary study, the score of whether the clinical question being asked was answered in the report, were used to provide data for one "axis" of the contingency table. The other "axis," data on whether the information received by the radiologist was consistent with the true clinical indication, was obtained from the results of the primary study. One table was constructed for inpatient data, and one was constructed for outpatient data. The Fisher's Exact statistic was tested for significance at a level of $\alpha = 0.05$.

CHAPTER III

RESULTS

The results of the research performed for this thesis are presented in two sections. In the first section, the results of the qualitative analyses of the processes for ordering, performing, interpreting and reporting, and billing for imaging procedures are described. The descriptions include an explanation of the flow of information produced from these processes and a qualitative description of potential information quality issues. In the second section, the results of the primary and secondary study quantitative information quality analyses, pertaining to the clinical indication for the procedure and the "value" of the radiologist's report, respectively, are presented.

Qualitative Analyses Results

The processes involved in ordering, performing, interpreting and reporting, and billing for an imaging procedure generate information used for clinical decision making, quality improvement, outcomes research, and imaging department operations. Information quality problems can arise anywhere within the chain of these various processes. The narratives and flow charts describing the ordering, performance, interpretation and reporting, and billing processes are presented in the following order:

1. Procedure ordering
2. Procedure performance
3. Procedure interpretation and reporting
4. Facility (technical) and professional services billing

Procedure Ordering

The information produced from the ordering process is used in the performance, interpretation and reporting, and billing of imaging procedures. Information quality problems in the ordering process can impact any of these downstream processes. The processes for ordering imaging procedures differ for the inpatient and outpatient populations. Therefore, the processes for these two populations are described separately.

Inpatient ordering process. The first step in the inpatient procedure ordering process (Figure 13) is for a physician requesting an imaging procedure to complete a Diagnostic Radiology and Nuclear Medicine Request form (Figure 5). As described in Chapter II, the information provided on the request form includes patient demographics, requesting physician name, attending physician name, procedure being ordered, and clinical indication for the procedure. Of these pieces of information, the ones entered on the form by the requesting physician are their name, the procedure being ordered, and the clinical indication for the procedure. The pieces of information entered by the requesting physician are handwritten in a free text format. After completion of "their part" of the form, the requesting physician gives the request to a ward clerk. The ward clerk then stamps the request with the

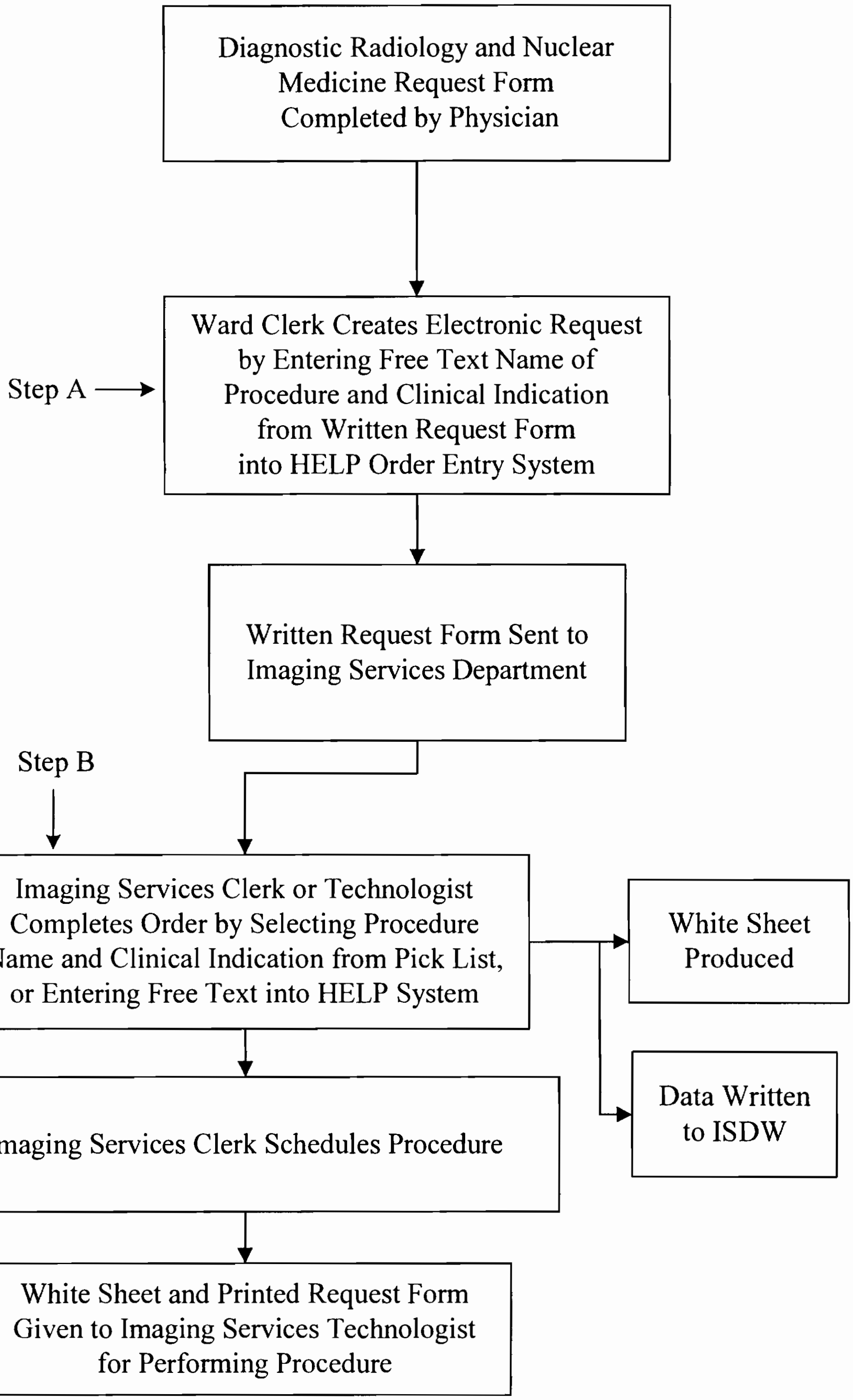


Figure 13. Inpatient procedure ordering process.

patient's demographic information from a stamper plate.

Using the request form as a guide, the ward clerk uses the Order Entry program on the HELP System to submit an electronic request for a procedure to the imaging services department. The two pieces of information pertinent to this research that are entered by the clerk into the HELP System are as follows:

1. Name of the procedure being requested (e.g., Abdominal CT)
2. Clinical indication for requesting the procedure (e.g., right lower quadrant pain)

These two pieces of information are entered into the HELP System by the clerk as free text.

Creating the request for the procedure in the HELP System (Figure 13, Step A), is the first place where a problem with information quality may arise. One quality problem that may occur is inaccuracy of the information. For example, a discrepancy between the true reason the physician ordered the procedure (the clinical indication written on the request form) and the free text reason in the HELP System may occur, if the ward clerk cannot read the physician's writing. If the ward clerk cannot read the clinical indication on the request form, does not make the effort to contact the physician and clarify the order, and enters what they believe the physician has written, a data accuracy problem will be created if the clerk's interpretation of what was written is incorrect. Another quality problem that may occur during the creation of the imaging services request is incompleteness of the information. For example, if the physician does not write a clinical indication for the procedure on the request form, the ward clerk may leave the clinical indication field blank in the HELP System.

After completing the order request in the HELP System, the ward clerk sends the Diagnostic Imaging and Nuclear Medicine Request paper form to the imaging services department via a pneumatic tube system or a hospital messenger. Upon receiving the written request form, a clerk or technologist in the imaging services department pulls up the electronic order request in the HELP System and completes the order (Figure 13, Step B). Completing the order includes three main tasks:

1. The first task is selecting the procedure type. Using the free text name of the procedure in the electronic request, the imaging services clerk or technologist selects a procedure name from an encoded pick list in the HELP System. The text in the pick list is linked to CPT-4 procedure and associated PTXT codes. If no matching term is present on the pick list, the imaging clerk or technologist enters a free text procedure name into the order completion screen.
2. The second task is selecting the clinical indication for the procedure. Using the free text description of the clinical indication for the procedure in the electronic request, the imaging services clerk or technologist selects from an encoded pick list in the HELP System, a clinical indication that most closely matches the free text reason submitted by the ward clerk. The text in the pick list is linked to ICD-9-CM diagnosis and associated PTXT codes. If no matching term is present on the pick list, the imaging clerk or technologist enters a free text clinical indication into the order completion screen.
3. The third task is producing the printed white sheet, attaching it to the written request form, and giving it to the radiology technologist to be used while performing the imaging procedure.

When the order is completed, the order information is written to the HELP System database. The procedure type is written to the database as the PTXT codes representing the CPT-4 procedure code and associated text. The clinical indication is written as the PTXT codes representing the ICD-9-CM diagnosis code and associated text. If an order was completed using free text rather than entries from the encoded pick lists, no PTXT codes for the CPT-4 or ICD-9-CM codes are stored in the HELP System for that order. However, the free text information is stored as PTXT codes. At the same time that the information from the completed order is written to the HELP database, the data-drive mechanism triggers a process that writes the order information to the ISDW.

The process of completing the order (Figure 13, Step B) is the second place where information quality problems may arise. If there is no exact match on the pick list for the free text clinical indication entered in the electronic order request, imaging services clerks or technologists either select from the pick list the reason they believe most closely matches the ward clerk's free text indication for the procedure, or enter the ward clerk's, or their own, free text into the order completion screen. In each of these scenarios, the clinical indication for the procedure recorded in the completed order in the HELP System may not accurately reflect the physician's reason for ordering the procedure. Furthermore, because the imaging services clerk or technologist who completes the order also has the written request form as a reference, if they notice that the indication on the request form and the clerk's free text indication are different, the imaging services personnel must try to determine if the ward clerk's indication is incorrect or if there was a change between the time the written request was created and the electronic request was entered. Depending on the success of the imaging services clerk or technologist's investigation, the discrepancy between the written request and ward clerk's free text clinical indication presents another opportunity for creating an inaccuracy or inconsistency in the information about the imaging procedure. In other cases, although there may be a pick list match for the ward clerk's free text reason, the free text reason is not specific (e.g., "Post-Op"). If the information is not specific, the radiologist who ultimately uses this information to interpret the imaging procedure will not know what clinical information the referring physician is seeking. If the radiologist does not know what information the referring physician is seeking, he or she may not tailor the procedure report appropriately. In still other cases, if the radiologist receives both the white sheet and the written request form as information sources during interpretation and there is conflicting information on these two forms, it may be the case that, in spite of his or her best efforts, the radiologist is unable to determine which source of information is accurate and thus will be unable to appropriately tailor the procedure report.

Outpatient ordering process. The first step in the outpatient ordering process (Figure 14) is for a clerk from the requesting physician's office to call the imaging services scheduling office to request an imaging procedure. Using information provided by the requesting physician, the physician's office clerk provides the imaging services scheduler with information including patient demographics, the type of procedure being ordered, and the clinical indication for the procedure. The second

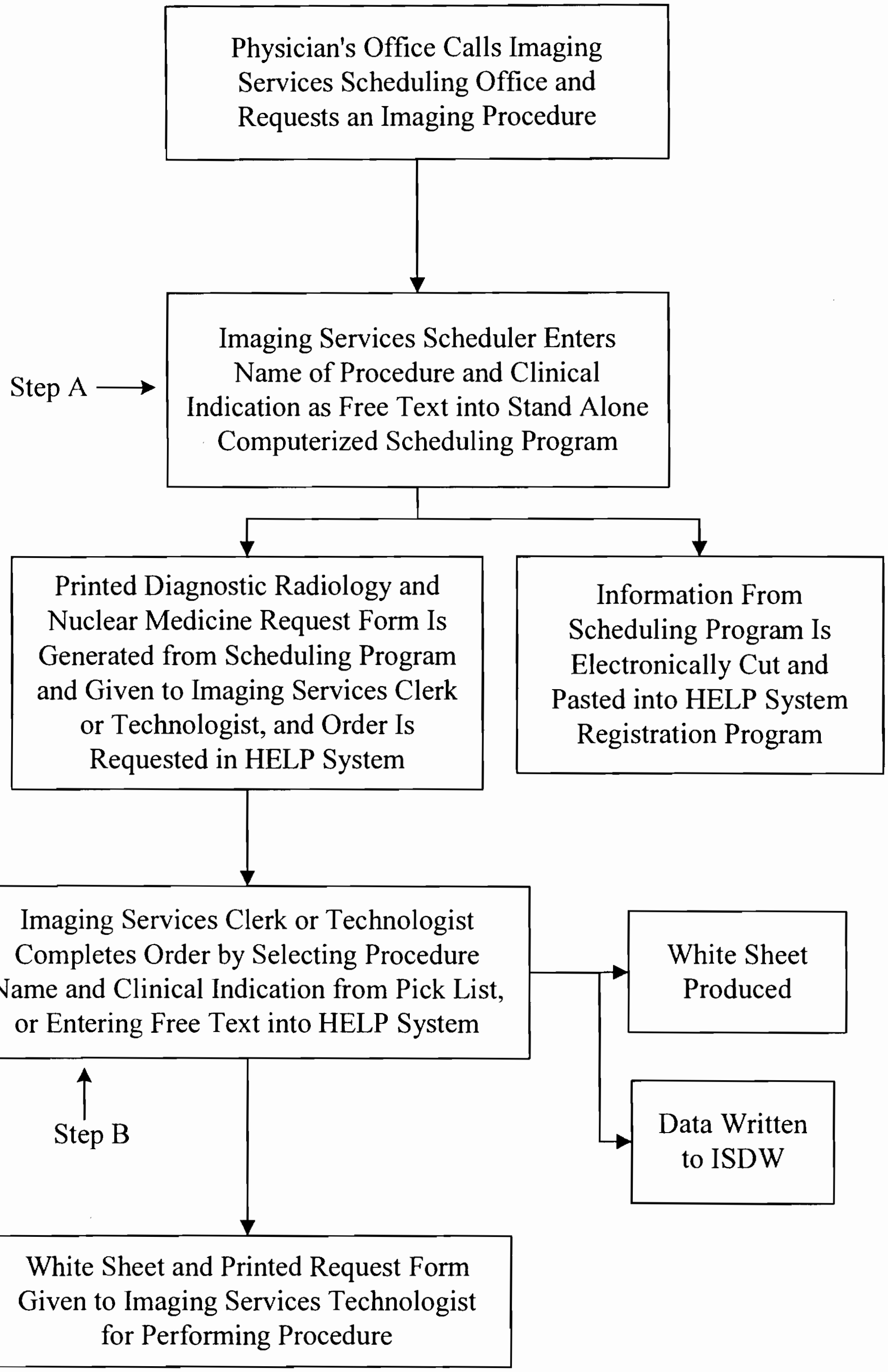


Figure 14. Outpatient procedure ordering process.

step is for the imaging services scheduler to enter the information obtained from the physician's office clerk into a stand-alone computerized scheduling program in a free text format (Figure 14, Step A). The third step is generation, by the imaging services scheduler, of a printed Diagnostic Radiology and Nuclear Medicine Request form (Figure 6) from the information in the scheduling program and an order request in the HELP System. In addition, the information from the scheduling program is electronically "cut and pasted" by the imaging services scheduler into the HELP System registration program. Thus, in the case of outpatients, the free text version of the clinical indication for the imaging procedure also is used as the admitting diagnosis in the HELP System. The admitting diagnosis is ultimately used for technical billing purposes for outpatients.

Although the "true" clinical indication for the procedure in the outpatient setting is the requesting physician's documented indication in his or her office chart, the requesting physician's office records were not accessible for review. Therefore, although recognizing the potential for additional information quality errors in the transmission of the clinical indication from the requesting physician to his or her office clerk and from the office clerk to the imaging services scheduler, because the most proximate (to the physician's documented indication) documented source of information about the clinical indication available for review was the imaging services scheduler free text indication on the request form, the scheduler documented indication was used as the gold standard for analysis of accuracy in the outpatient population.

In spite of the difference in gold standard for the "true" clinical indication for the imaging procedure in the inpatient and outpatient populations, in both cases, the requesting physician's clinical indication for the imaging procedure is being "transcribed" into an electronic request as free text information. Therefore, the "transcription" step (Figure 14, Step A) presents an opportunity for misinterpretation or incorrect data entry in the outpatient process, just as it does in the inpatient process. Thus the same potential information quality problems described for the "transcription" step in the inpatient ordering process may arise in the outpatient ordering process.

The fourth step in the outpatient ordering process is for the Diagnostic Radiology and Nuclear Medicine Request form generated from the stand-alone computerized scheduling program to be delivered to an imaging services clerk or technologist. The clerk or technologist uses the information on the request form to create and complete an order in the HELP System using a process similar to that used for inpatients. Completing the outpatient order (Figure 14, Step B) includes three main tasks:

1. The first task is selecting the procedure type. Using the free text name of the procedure on the Diagnostic Radiology and Nuclear Medicine Request form, the imaging services clerk or technologist selects a procedure name from the same encoded pick list in the HELP System that is used in the inpatient ordering process. If no matching term is present on the pick list, the imaging clerk or technologist enters a free text procedure name into the order completion screen.

2. The second task is selecting the clinical indication for the procedure. Using the free text description of the clinical indication for the procedure on the Diagnostic Radiology and Nuclear Medicine Request form, the imaging services clerk or technologist selects from an encoded pick list in the HELP System (same one used in the inpatient process), a clinical indication that most closely matches the free text reason on the request form. If no matching term is present on the pick list, the imaging services clerk or technologist enters a free text clinical indication into the order completion screen.
3. The third task is producing the printed white sheet, attaching it to the printed request form and giving it to the imaging services technologist to be used while performing the imaging procedure.

When the order is completed, as with the inpatient order information, the outpatient order information is written to the HELP System database, and the data-drive mechanism triggers a process that writes the same order information to the ISDW. For the same reasons discussed in the description of the inpatient ordering process, the completion of the imaging procedure order in the HELP System is the second place where information quality problems may arise in the outpatient ordering process.

Procedure Performance

The process for performing an imaging procedure is the same for inpatients and outpatients (Figure 15). After receiving the white sheet and Diagnostic Radiology and Nuclear Medicine Request form, the imaging services technologist takes the

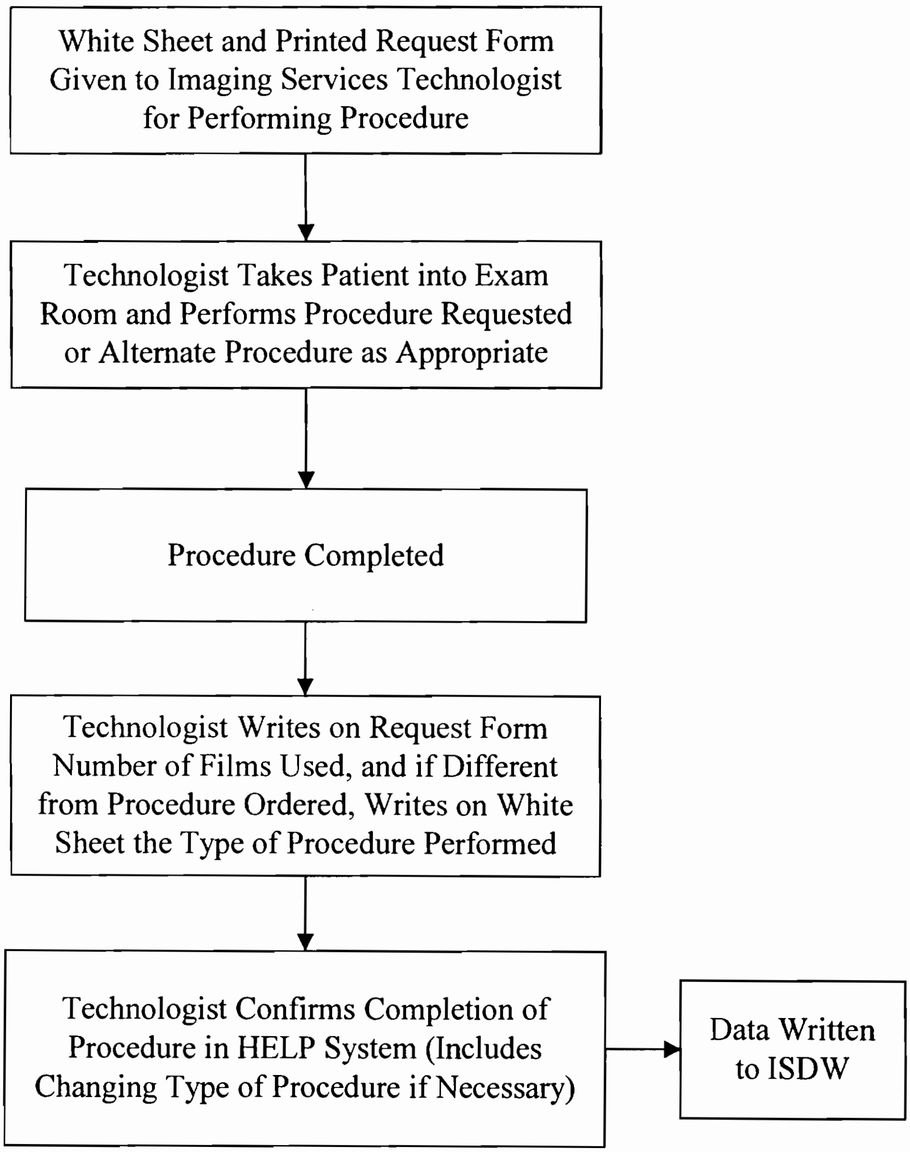


Figure 15. Procedure performance process.

patient into the exam room to perform the procedure indicated on the white sheet. If the patient is unable to tolerate the requested procedure, then, depending on the situation, the technologist either will consult the radiologist or use his or her own judgment in determining an alternate procedure to be performed. For example, if an erect chest x-ray is ordered but the patient is unable to stand, the technologist will perform a supine chest x-ray.

When the procedure is completed, the technologist writes the number of films used on the Diagnostic Radiology and Nuclear Medicine Request form and confirms completion of the procedure in the HELP System. If the procedure performed by the technologist differed from the procedure ordered in the HELP System, at the time of procedure confirmation the technologist is supposed to write on the white sheet the name of the procedure actually performed and modify the order in the HELP System. The order is modified in the HELP System by changing the type of procedure from the type ordered to the type actually performed. If the technologist makes this change, the information is changed not only in the HELP System but also in the ISDW due to the data-drive program. Knowing the actual procedure performed is important to the radiologist when interpreting the results of the study. However, the technologists do not always perform the step of correcting the information on the white sheet and in the HELP System when confirming completion of the procedure. The failure to correct the procedure type information creates an information quality problem that impacts the downstream processes of both the imaging procedure interpretation and the technical and professional services billing.

Procedure Interpretation and Reporting

Interpretation and reporting of imaging procedures also are the same in the inpatient and outpatient populations and require integration of several pieces of information including the films from the current procedure, films from prior procedures, the Diagnostic Radiology and Nuclear Medicine Request form, and the white sheet (Figure 16). After completion of an imaging procedure, the imaging services technologist delivers the Diagnostic Radiology and Nuclear Medicine Request form, the white sheet, and the films obtained during the current procedure to the imaging services file room. The file room personnel match the patient's demographic information with information in the HELP System RIS to determine if the patient has had previous imaging procedures. If films from previous procedures exist, the file room personnel retrieve them. After retrieving the films from previous imaging procedures, the "old" films along with the films from the current procedure, the Diagnostic Radiology and Nuclear Medicine Request form, and the white sheet are delivered to the radiologist. The radiologist uses each of these pieces of information to perform his or her interpretation of the current imaging procedure and dictate a preliminary report. The preliminary report is typed into the HELP System by an imaging services transcriptionist. The radiologist uses a HELP System Radiologist workstation to review the transcribed report for accuracy, to make any necessary corrections, and to electronically sign the report to produce the final version. The final report is stored both electronically in the HELP System and on paper in the patient's paper medical record.

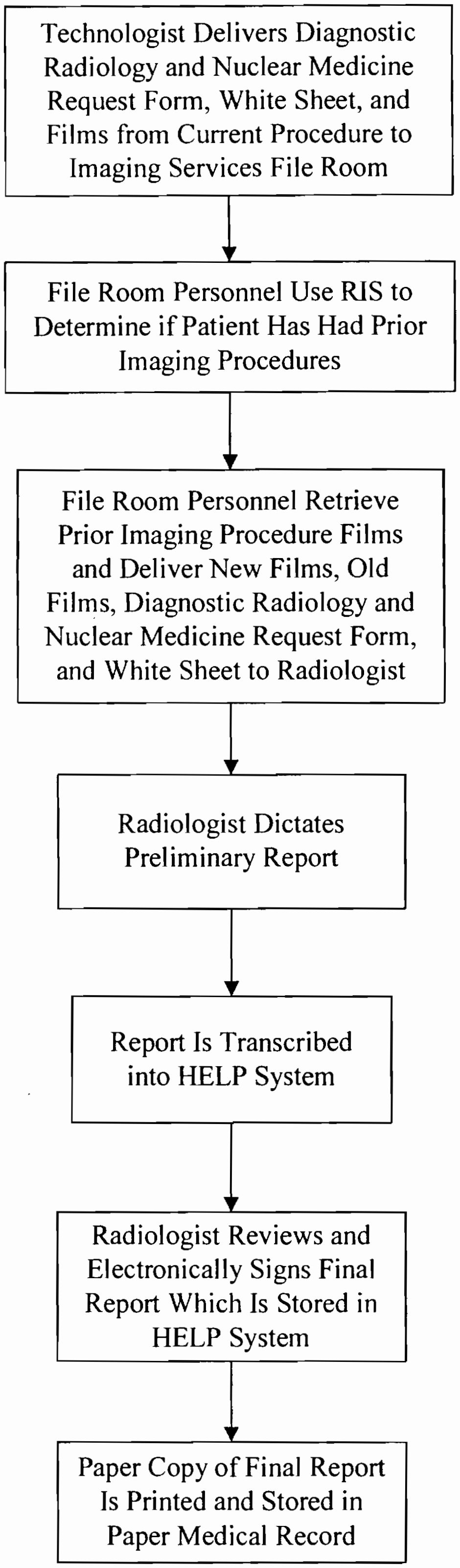


Figure 16. Procedure interpretation and reporting process.

In some cases, the films from previous imaging procedures, the white sheet, or the Diagnostic Radiology and Nuclear Medicine Request form may not be available to the radiologist when interpreting the current films. In these cases, the information quality problem is incomplete information, and it becomes more difficult for the radiologist to provide a clinically useful interpretation. In cases in which the information on the white sheet or Diagnostic Radiology and Nuclear Medicine Request form is inaccurate, or inconsistent, the radiologist may provide an erroneous interpretation.

Facility (Technical) and Professional Services Billing

As described above, problems with the quality of information pertaining to the procedure type and the clinical indication for the procedure may be produced during the procedure ordering and performing processes and may impact the radiologist's interpretation and report, as well as the billing processes. In addition, information quality problems pertaining to the procedure type and the clinical indication for the procedure may be created during the processes used for producing bills for facility (technical) and professional services. For example, the information stored in the HELP database as a result of the procedure ordering and performing processes may be changed during the procedure or diagnosis coding process that takes place in the imaging services billing department. The information quality problems produced during the billing processes can impact the downstream clinical and operational uses of the encoded information. In general, information quality problems produced during the billing processes occur after the radiologist has produced his or her report and thus do not directly impact the reporting process.

Procedure coding for facility (technical) and professional services billing. The process for assigning CPT-4 procedure codes for professional and technical services billing is the same for inpatients and outpatients (Figure 17). The data flow supporting this process begins after the imaging procedure has been completed. Confirmation in the HELP System of completion of the imaging procedure by the imaging services technologist results in IHC charge codes and CPT-4 procedure codes being written to the "Daily Hospital Log" (Figure 18) generated by the HELP System. The hospital log is printed nightly and contains information on all procedures ordered that day. The information contained in the Hospital Log includes patient demographics, procedure charge codes, CPT-4 procedure codes, ICD-9-CM diagnosis codes for the clinical indication for the procedure, and order status. The Hospital Log is used to track work in progress in the imaging services department. It provides a way for technologists, supervisors, and billing personnel to confirm the progress and status of each of the orders for imaging procedures.

In addition to the Hospital Log, a log called the "Radiologists' Log" (Figure 19) also is generated by the HELP System. The Radiologists' Log differs from the Hospital Log in that it contains insurance information and ICD-9-CM diagnosis codes (DC 24) for the patient's admitting diagnosis, in addition to the patient demographic information, CPT-4 procedure codes, and ICD-9-CM diagnosis codes for the clinical indication for the procedure, that are contained in the Hospital Log. Another difference between the Hospital Log and the Radiologists' Log is that the

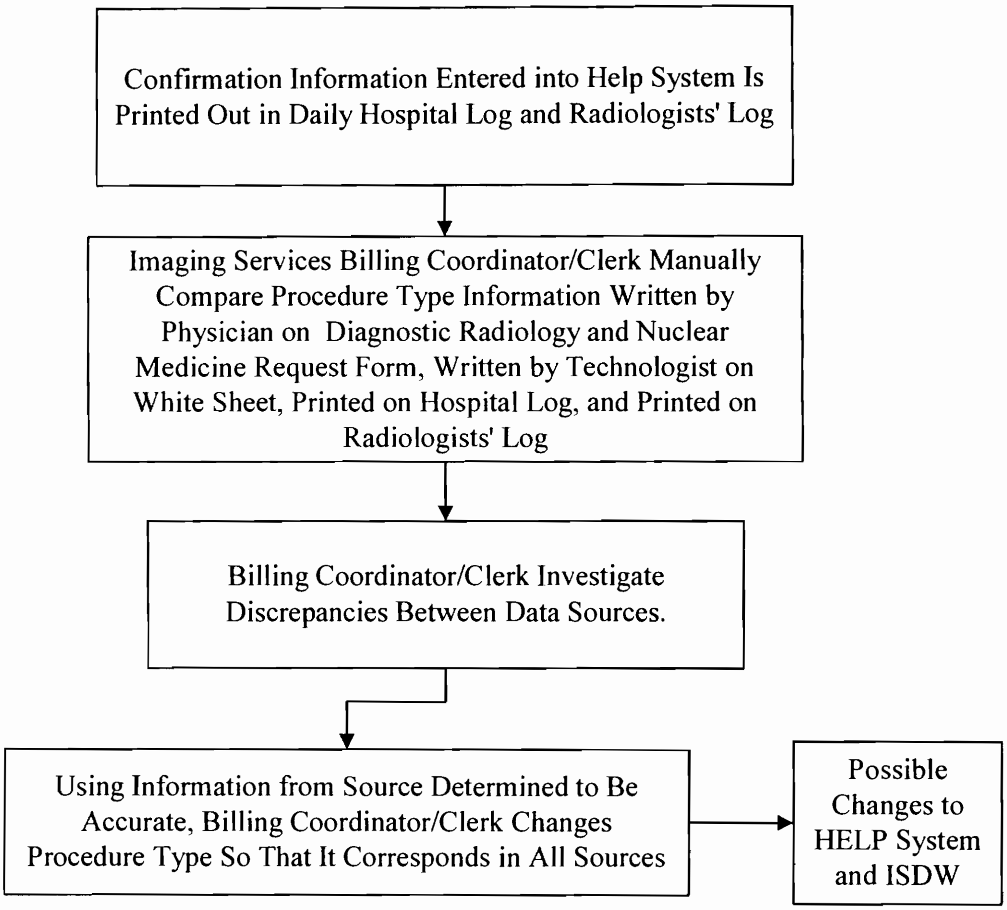


Figure 17. Procedure coding for facility (technical) and professional services billing.

Department : RADIOLOGY
Order Status : COMP Billing Status : ALL Sub Department : ALL Printed : 12/28/96.03:14

Dt.Seq	Order Time	RAD #	ENC #	Patient Name	Pt.Loc.	R MD	D MD	PRI	TRAN	Film Loc	Bill	Ord
12/01/96												
12/02/96												
02.199	12/02/96.12:44	********			DSCH	232	0	Today	PORTABLE	FILE ROOM	BILL	COMP
	Primary Procedure: 01620202			72020 Spine Cerv 1 View					Reason: V67.0 POST-OP CHECK			
	Secondary Items: 01664903			Portable Charge for (1)								
02.216	12/02/96.13:43	********				554	0	Today	WALK	FILE ROOM	BILL	COMP
	Primary Procedure: 01640002			74000 Abdomen/Kub 1 View					Reason: 789.0 PAIN, ABDOMINAL			
02.259	12/02/96.15:32	********			DSCH	1432	0	Today	WALK	FILE ROOM	BILL	COMP
	Primary Procedure: 01610203			71020 Chest 2 Views					Reason: 518.89 PRE-OP			
02.273	12/02/96.15:47	********			OPD	2371	0	Today	WALK	FILE ROOM	BILL	COMP
	Primary Procedure: 01660919			76091 Mammography Bilatera					Reason: 611.72 MASS, BREAST			
02.274	12/02/96.15:49	********			OPD	6620	0	Today	WALK	FILE ROOM	BILL	COMP
	Primary Procedure: 07282239			78223 NM Hepatobiliary Ima					Reason: 536.8 PAIN			
02.298	12/02/96.16:40	********			DSCH	2336	0	Today	CARRY	FILE ROOM	BILL	COMP
	Primary Procedure: 01660653			76065 Osseous Surv Infant					Reason: 780.9 PAIN			
12/03/96												
03.171	12/03/96.10:28	********				2542	0	Today	PORTABLE	FILE ROOM	BILL	COMP
	Primary Procedure: 01610104			71010 Chest 1 View					Reason: 518.89 PRE-OP			
	Secondary Items: 01664903			Portable Charge for (1)								
03.187	12/03/96.11:06	********				232	0	Today	WALK	FILE ROOM	BILL	COMP
	Primary Procedure: 01620202			72020 Spine Cerv 1 View					Reason: V67.0 POST-OP CHECK			
03.280	12/03/96.15:19	********			OPD	232	0	Today	WALK	FILE ROOM	BILL	COMP
	Primary Procedure: 01620202			72020 Spine Cerv 1 View					Reason: V67.0 POST-OP CHECK			
03.321	12/03/96.17:56	********			DSCH	6410	0	Today	PORTABLE	FILE ROOM	BILL	COMP
	Primary Procedure: 01636109			73610 LT Ankle 3 Views					Reason: 827.0 REDUCTION			
	Secondary Items: 01664903			Portable Charge for (1)								

Figure 18. Daily Hospital Log.

07/16/1997

--------------------------- Patient Information -------------------------------

Name : [redacted] [redacted]

Visit # : [redacted] Med Rec #: 00000000 Patient Type: 0001 Bill Type 0038

Address : [redacted] , SOUTH JORDAN , UT 84095

Phone : [redacted] Ssn: [redacted] Emp: HOMEMAKER

DOB : [redacted] Age: [redacted] Sex: F Admit Date: 07/16/97

--------------------------- Guarantor Information ----------------------------

Name : [redacted] Relation: SELF

Address : [redacted] , SOUTH JORDAN , UT 84095

Phone : [redacted] [redacted]

Employer: HOMEMAKER Phone: ( )

Emp. Add: , ,

--------------------------- Spouse/Parent Information -------------------------

Name : [redacted] Relation: HUSBAND

Address : [redacted] , SOUTH JORDAN , UT 84095

Phone : [redacted] [redacted]

Employer: RET/QUESTAR 1986 Phone: ( )

Emp. Add: , ,

--------------------------- Friend/Relative Information -----------------------

Name : Relation:

Address : , ,

Phone : ( )

--------------------------- Primary Insurance Information ---------------------

Primary Ins. : SENIORCARE BASIC/STD/HIGH HMO [redacted]

Address : PO BOX 11657 , SALT LAKE CITY , UT 84147

Phone : [redacted]

Policy Holder: [redacted] Relation: P

Contract # : [redacted] Group #:

--------------------------- Secondary Insurance Information -------------------

Secondary Ins: Carrier #:

Address : , ,

Phone : ( )

Policy Holder: Relation:

Contract # : Group #:

--------------------------- Service Information -------------------------------

Admit Diagnosis : CYST

Attending Physican : 00593 VANKOMEN, GEORGE J.

Primary Care Physican : 00593 VANKOMEN, GEORGE J.

16 0013 2901 610.0 SOLITARY CYST OF BREAST

76645 Breast(Gen) 1 10.00 XRAY 00593

Figure 19. Radiolgoists' Log.

only procedures included on the Radiologists' Log are those that have been interpreted, dictated, and signed by the radiologist.

Each day, the imaging services billing coordinator or clerk manually compares the procedure type on the Diagnostic Radiology and Nuclear Medicine Request form, the procedure type printed on the white sheet or written on the white sheet by the technologist if the procedure changed from the type originally printed on the white sheet and the procedure type information on the Hospital and Radiologists' Logs. If there is a discrepancy between the information in these four data sources, the imaging services billing personnel investigate to determine the "accurate" information. Using information from the source determined to be accurate (a subjective decision on the part of the billing coordinator or clerk), the billing coordinator or clerk changes the procedure type in the other three sources so that it corresponds in all four sources. As a result of this process, it is possible that the procedure type encoded in the HELP System and written to the ISDW at the time the procedure was ordered will be changed to another procedure type to reflect what is believed to be more accurate information. In addition, the procedure type indicated on the Hospital or Radiologists' Log also may be changed.

The facility (technical) bills are produced by the hospital billing service after the patient is discharged, using the CPT-4 procedure codes stored in the HELP System. The professional services bills are produced by the radiologists' billing service, using the CPT-4 procedure codes on the Radiologists' Log. For both types of bills, the CPT-4 procedure codes are entered onto a billing form specific to the patient's particular payer. Changes to the CPT-4 procedure codes in the multiple data sources described above, if not kept in sync, may lead to discrepancies between the CPT-4 codes used on the facility and professional bills. Discrepancies between the facility and professional bills for the same procedure on the same patient may result in violation of Medicare fraud and abuse regulations. Appendix D contains an example of how a discrepancy between the CPT-4 procedure codes on the facility and professional bills for the same patient may arise.

Diagnosis coding for facility (technical) and professional services billing – general information. The basic steps in the process for diagnosis coding, as outlined in Figure 20 and described in this section, are similar for facility and professional billing. After the patient is discharged, the imaging services department billing coordinator or billing clerk manually compares the clinical indication for the imaging procedure on the Diagnostic Radiology and Nuclear Medicine Request form, the HELP System ICD-9-CM encoded clinical indication on the white sheet, and the ICD-9-CM code for the indication as noted on the Hospital Log. Using the assumption that the clinical indication on the request form is the "true" indication for the procedure if there is a discrepancy between the three sources mentioned above, the billing coordinator or clerk will change the encoded reason in the HELP System and on the Hospital Log so that they correspond with the ordering physician's written reason for the procedure.

Neither the imaging services billing coordinator nor billing clerk reviews the clinical indication for the procedure on the Radiologists' Log as a routine part of his or her work. There is no specific procedural reason why this is the case; it is purely based on historical practices. Therefore, in contrast to the process for the procedure

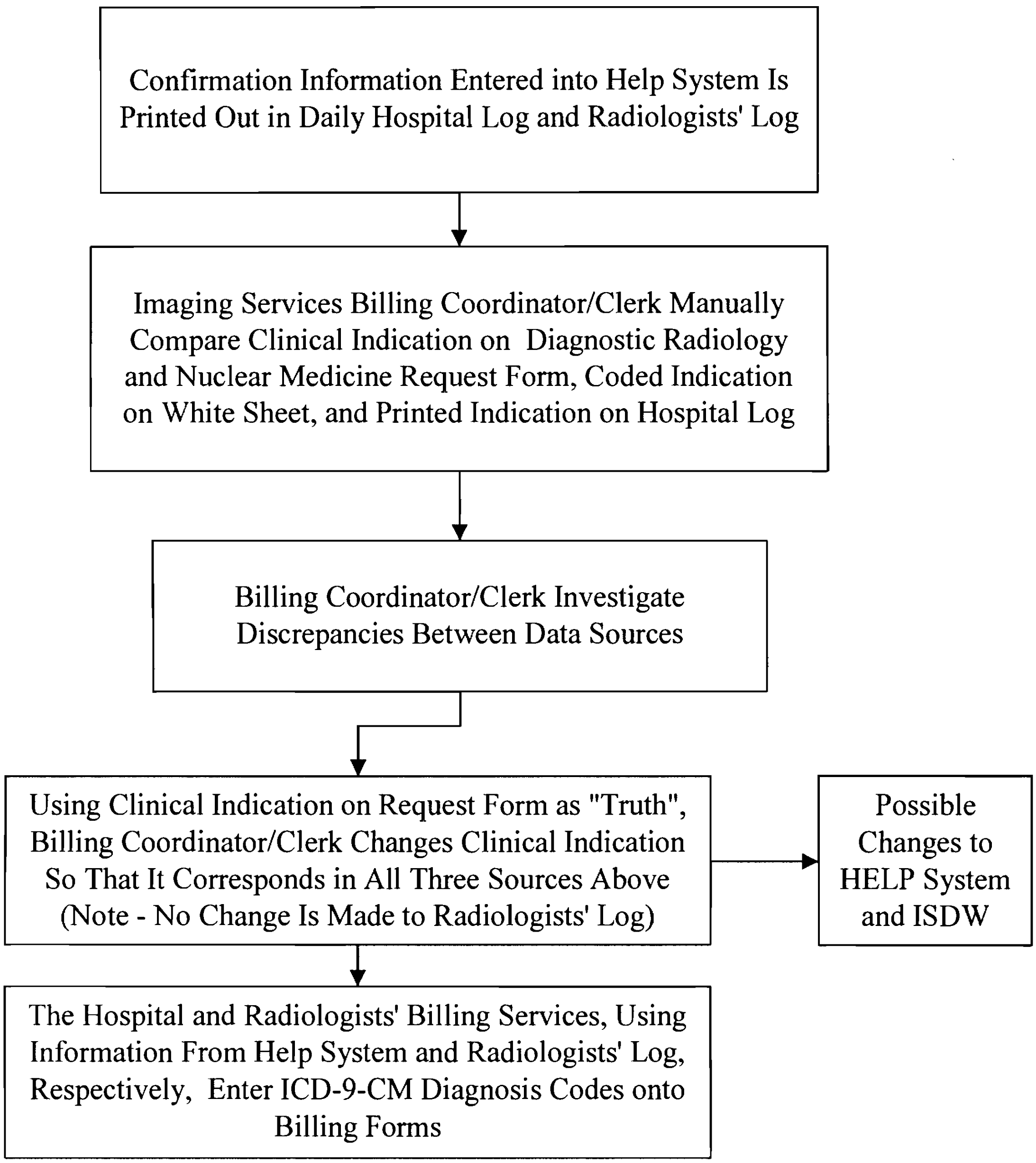


Figure 20. Diagnosis coding for facility (technical) and professional services billing – general information.

type coding, information pertaining to the clinical indication for an imaging procedure is changed on the Hospital Log but not on the Radiolgists' Log. If the radiologist's report for an imaging procedure has not been signed by the radiologist, the procedure type and clinical indication for the procedure will not have been written to the Radiologists' Log at the time the billing coordinator or clerk make changes in the HELP System and on the Hospital Log. Therefore, when the procedure report is signed and the procedure type and clinical indication information is written to the Radiologists' Log, the indication for the procedure on the Radiologists' Log will match the indication in the HELP System and on the Hospital Log. If, on the other hand, the report already has been finalized and thus the information about the procedure already written to the Radiologists' Log at the time the above changes are made in the HELP System and on the Hospital Log, the HELP System and Hospital Log information will differ from the Radiologists' Log information.

The ICD-9-CM diagnosis codes stored in the HELP System and on the Radiologists' Log, which may differ after the imaging services department billing office procedures, are added to the information on the facility (technical) and professional services bills, respectively. Because the ICD-9-CM codes for facility and professional services bills are obtained from different data sources, as with the CPT-4 procedure codes, changes to the ICD-9-CM diagnosis codes in the multiple data sources described above, if not kept in sync, may lead to discrepancies between the ICD-9-CM codes used on the facility and professional bills. Appendix E contains an example of how a discrepancy between the ICD-9-CM diagnosis codes on the facility and professional bills for the same patient may arise.

Although the basic steps in the diagnosis coding process are similar for facility and professional billing, the processes involved in each of the steps, as described in the next several sections, differ between the two types of billing.

Diagnosis coding for facility (technical) services billing. The process for diagnosis coding for facility services billing differs for inpatients and outpatients. Therefore, the processes for inpatient and outpatient populations are described separately. Moreover, the process for outpatients differs according to payer and procedure type. The process steps that differ as a function of payer differences are noted in the description of the outpatient coding process.

Inpatient diagnosis coding for facility (technical) services billing. The billing forms mentioned in the description of the procedure coding process contain ICD-9-CM diagnosis codes, as well as CPT-4 procedure codes (in actuality the forms also contain ICD-9-CM procedure codes; however, those codes are not pertinent to this discussion). As with the CPT-4 procedure codes, Health Information Systems (HIS) enters the ICD-9-CM diagnosis codes onto the billing forms after the patient is discharged. For all patients, except those billed using DRGs, the ICD-9-CM diagnosis code is based on the clinical indication associated with the completed imaging services order in the HELP System. In the case of patients billed using DRGs, the ICD-9-CM diagnosis codes are not derived from imaging services information. For patients billed using DRGs, the ICD-9-CM codes entered onto the billing forms are based on the admitting diagnosis, discharge diagnoses, and complication diagnoses in the HELP System and paper chart (Figure 21).

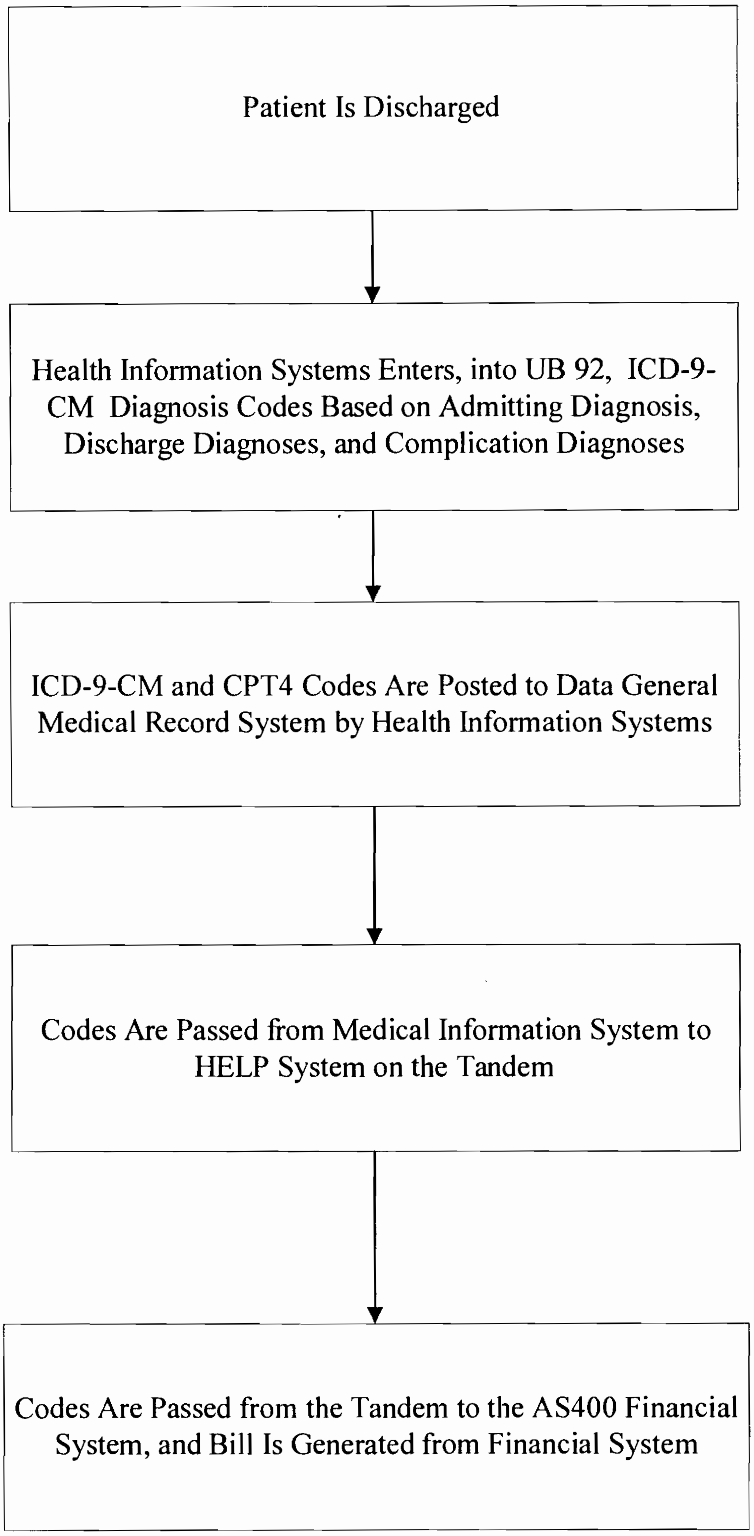


Figure 21. Inpatient diagnosis coding for facility (technical) services billing.

After entering the ICD-9-CM and CPT-4 codes onto the billing forms, the coded information is posted by HIS to the 3M Code-3 Medical Record System (Figure 8). The information from the Medical Record System is then passed to the HELP System on the Tandem and subsequently to the AS400 financial system. The bill is generated from the AS400.

The accuracy of the procedure-related clinical indication information (ICD-9-CM diagnosis code) that is stored in the HELP System and the consistency of that information with other data sources directly impacts on the consistency between the facility and professional bills, which is important with respect to not violating Medicare fraud and abuse regulations.

Outpatient diagnosis coding for facility (technical) services billing. Depending on the type of insurance the patient has, either Health Information Systems or Patient Account Services (PAS) performs the diagnosis coding for facility services billing for outpatients (Figure 22). Health Information Systems performs the coding for patients whose insurance is provided by Medicare, Medicaid, and Champus; for all Emergency Department (ED) patients; and for patients who undergo high risk (e.g., interventional radiology) procedures. Patient Account Services performs coding for all patients that do not fall into the categories listed as being coded by HIS.

Within four days after the patient is "discharged" (i.e., has had the imaging procedure performed), HIS or PAS assign an ICD-9-CM diagnosis code to the patient's billing record. However, the information HIS and PAS use to determine what diagnosis code(s) to assign differs. The HIS information sources are the free text admit diagnosis in the HELP System registration program and the report of the

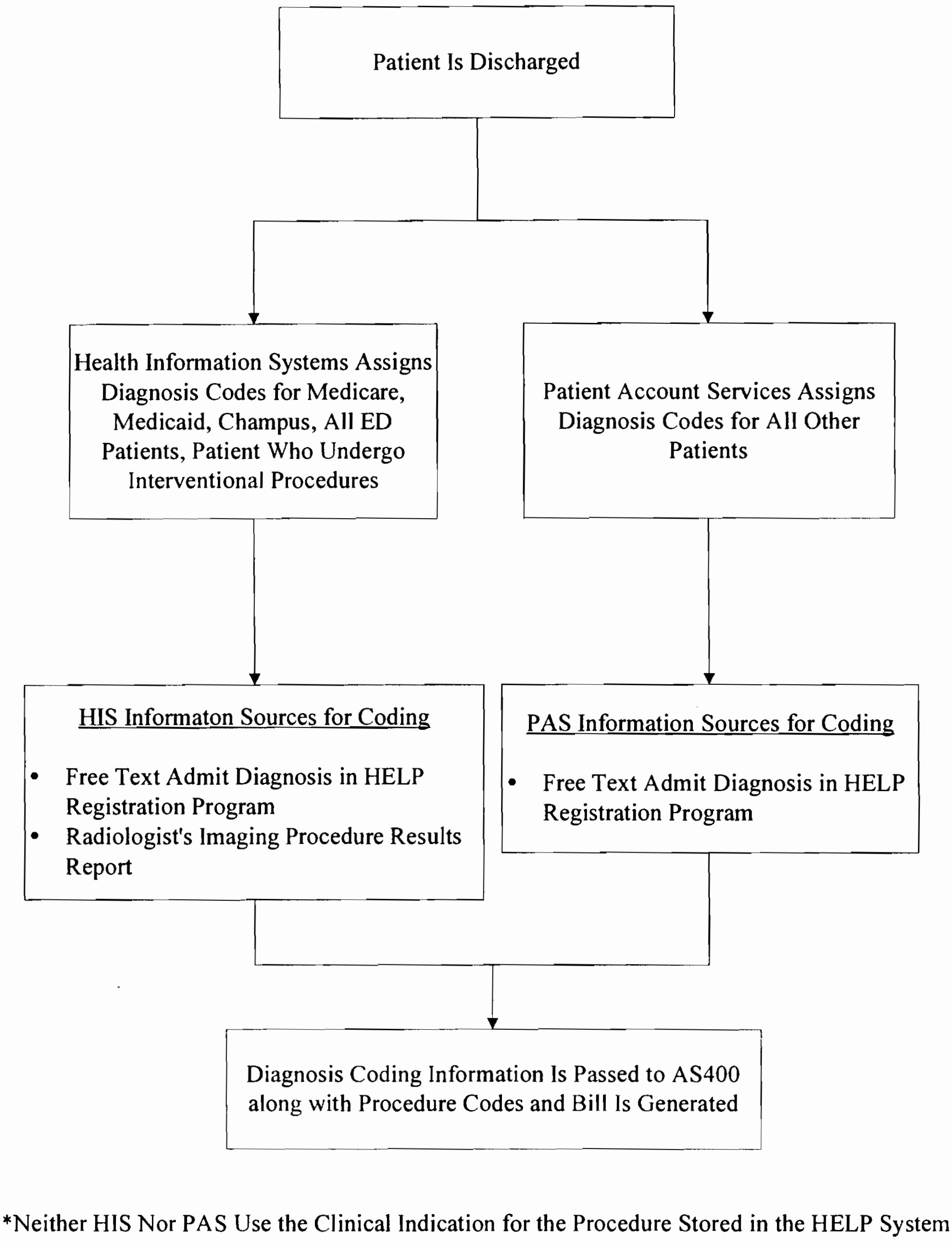


Figure 22. Outpatient diagnosis coding for facility (technical) services billing.

radiologist's interpretation of the imaging procedure results. The report is used by the coders reading it and inferring the symptoms and/or diagnoses for which the imaging procedure was performed. Using the information from the report in conjunction with the admitting diagnosis from the HELP System, the coders determine which ICD-9-CM diagnosis code(s) to assign the admitting diagnosis and thus for billing. The only information source used by PAS is the free text admit diagnosis code in the HELP registration program. Patient Account Services does not review other portions of the patients' records or the imaging procedure report in determining the correct ICD-9-CM diagnosis code to assign for billing. Of note is that neither HIS nor PAS uses as a data source the final clinical indication for the procedure stored in the HELP System as a result of the completed ordering process. Although the free text admitting diagnosis in the HELP System is the same as the free text clinical indication in the initial imaging procedure order request for outpatients, because of the outpatient ordering and billing process, as described above, it is possible that the final coded indication in the HELP System for an imaging procedure may differ from the free text admitting diagnosis and ultimately the coded admit diagnosis in the HELP System. Thus, it is possible that the ICD-9-CM diagnosis code on the bill for technical services will differ from the ICD-9-CM diagnosis code stored in the HELP System as the clinical indication for the procedure. For this reason, as well as due to the process for professional services billing as described below, it is possible the ICD-9-CM diagnosis code on the facility and professional bills will differ.

Diagnosis coding for professional services billing. The process for assigning ICD-9-CM diagnosis codes for professional services bills is the same for inpatients and outpatients (Figure 23). The Radiologists' Log is sent to the radiologists' professional billing office in a printed paper format. At the radiologists' professional billing office, the patient account coordinator evaluates the information on the Radiologists' Log for the clinical indication for the procedure and the procedure requested. If, in the opinion of the professional billing office account coordinator, the clinical indication and procedure type on the log do not coincide (e.g., procedure ordered is abdominal ultrasound, and the clinical indication is "rule out peptic ulcer disease") or the ICD-9-CM code is a "V" code (an ICD-9-CM code type used for supplementary classification of factors influencing health status and contact with health services) and thus the bill is unlikely to get reimbursed, the professional billing office account coordinator changes the ICD-9-CM code so that it is "correct." This is done without seeking any additional patient information other than that included on the Radiologists' Log, such as the admitting diagnosis. Thus as stated above, due to the nature of the professional services billing process, there is potential for the ICD-9-CM diagnosis codes for the reason for the procedure for a particular patient to differ in the HELP System, on the facility bill, and on the professional bill.

Quantitative Analyses Results

Primary Study General Information

In the primary study, the quantitative analysis of information quality focused on the dimensions of accuracy, consistency, and completeness of the information pertaining to the clinical indication for ordering an imaging procedure. The definitions of the dimensions of information quality, the data sources used, and the

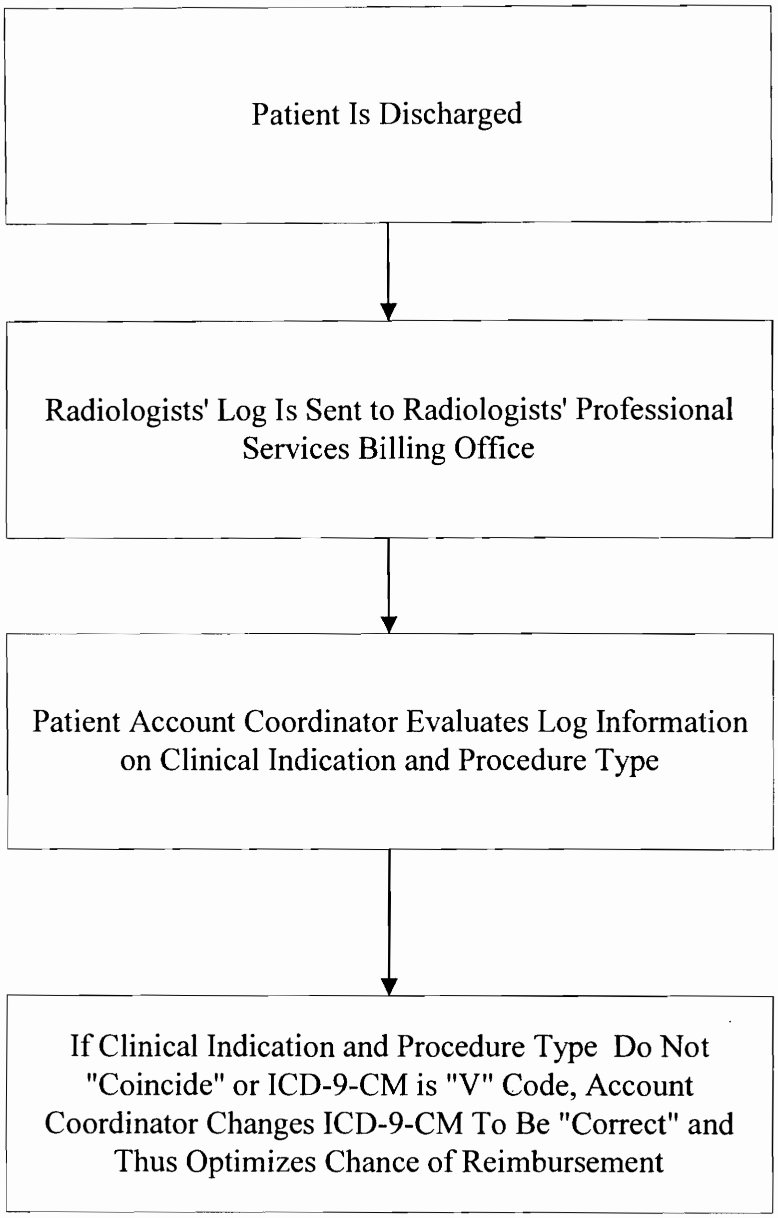


Figure 23. Diagnosis coding for professional services billing.

measurement methods and data analyses performed are described in the Materials and Methods section of this thesis.

The results of the quantitative analysis performed in the primary study are presented in three sections below. The first two sections present the inpatient and outpatient results, respectively. As noted in Chapter II, in the description of the statistical analyses, for those data element pairs that were present in both the inpatient and outpatient populations, an additional analysis was performed. The additional analysis is presented in the third section.

Primary Study Inpatient Results

The results of the first three data element comparisons in the inpatient section are an indication of the information consumers' perception of the accuracy and completeness of the data. In these data element comparisons, the cases in which the data elements are scored as the same represent the cases with accurate data. Cases in which the data elements are scored as different represent the cases with inaccurate data. The proportion of cases scored as having missing data represent the frequency of cases with incomplete data.

The subsequent four data element comparisons in the inpatient results section are an indication of the information consumers' perception of the consistency and completeness of the data. In these data element comparisons, the cases in which the data elements are scored as the same represent the cases with consistent data. Cases in which the data elements are scored as different represent the cases with inconsistent data. The proportion of cases scored as having missing data represent the frequency of cases with incomplete data.

MD documented indication compared to ward clerk free text indication. The composite results in Table 7 demonstrate that overall the scorers perceived the information in these two data sources to be the same 52.5% of the time, different 22.5% of the time, and missing 25.0% of the time. The individual results show that scorers 1 (diagnostic radiologist) and 2 (radiation oncologist) had similar scores in the same and different categories. However, scorer 3 (general surgeon) tended to rate the information as the same, less often, and as different more often than both scorer 1 and scorer 2. In contrast to the scoring in the same and different categories, all three scorers indicated a similar frequency of missing data.

The Kappa statistic result is shown in Table 8. The result suggests that 56.8% of the difference between perfect agreement between the raters and agreement expected by chance alone is accounted for by true agreement. Using the criteria of Landis and Koch,[179] the Kappa of 0.568 denotes "good" agreement between the raters. The Kappa statistic is statistically significant with a probability of $P < 0.0001$.

The results of the McNemar's statistic are shown in Table 9. The results suggest that scorer 1 and scorer 2 generally disagreed by chance. In contrast, the disagreement between scorer 3 and both scorer 1 and 2 was not by chance. In other words, there is some characteristic of the raters or some other influencing factor that impacted the way in which they disagreed.

MD documented indication compared to white sheet ICD-9-CM text indication. The composite results in Table 10 show that as a group, the scorers perceived the

Table 7

MD Documented Indication Compared to Ward Clerk Free Text Indication – Individual and Composite Rater Scores

Response	Same			Different			Missing		
	Proportion	LCL	UCL	Proportion	LCL	UCL	Proportion	LCL	UCL
Scorer1	0.543	0.412	0.668	0.210	0.123	0.336	0.247	0.152	0.376
Scorer2	0.543	0.412	0.668	0.173	0.095	0.294	0.284	0.182	0.415
Scorer3	0.350	0.236	0.484	0.388	0.269	0.522	0.263	0.164	0.394
Composite	0.525	0.394	0.653	0.225	0.134	0.353	0.250	0.154	0.380

Table 8

MD Documented Indication Compared to Ward Clerk Free Text Indication – Kappa Statistic

	Kappa	LCL	UCL	P
Kappa 2	0.568	0.416	0.719	< 0.0001

Table 9

MD Documented Indication Compared to Ward Clerk Free Text Indication – McNemar's Statistics

	McNemar's	P
Scorer 1 vs 2	0.333	0.564
Scorer 1 vs 3	15.000	< 0.001
Scorer 2 vs 3	14.000	< 0.001

Table 10

MD Documented Indication Compared to White Sheet ICD-9-CM Text Indication – Individual and Composite Rater Scores

Response	Same			Different			Missing		
	Proportion	LCL	UCL	Proportion	LCL	UCL	Proportion	LCL	UCL
Scorer1	0.494	0.366	0.623	0.494	0.366	0.623	0.012	0.002	0.087
Scorer2	0.395	0.276	0.528	0.580	0.448	0.702	0.025	0.005	0.107
Scorer3	0.123	0.060	0.236	0.864	0.749	0.931	0.012	0.002	0.087
Composite	0.358	0.244	0.491	0.630	0.497	0.746	0.012	0.002	0.087

information to be the same 35.8% of the time, different 63.0% of the time, and missing 1.2% of the time. As in the previous comparison, the results in Table 10 demonstrate that scorers 1 and 2 had similar scores in the same and different categories. In addition, also as seen in the previous comparison, scorer 3 tended to rate the information as the same less often and as different more often than either scorer 1 or 2. In contrast to the previous comparison, although scorers 1 and 3 continued to perceive the same frequency of missing data, scorer 2 indicated data were missing in twice as many cases as both scorers 1 and 3.

The result of the Kappa statistic is shown in Table 11. The Kappa statistic indicates that 38.5% of the difference between perfect agreement between the raters, and agreement expected by chance alone is accounted for by true agreement. The Kappa of 0.385 suggests the true agreement is only “marginal.” The Kappa statistic is statistically significant with a probability of $P < 0.0001$.

The results of the McNemar’s statistic are shown in Table 12. The results of the test for symmetry in the comparison of the MD documented indication and the white

Table 11

MD Documented Indication Compared to White Sheet
ICD-9-CM Text Indication – Kappa Statistic

	Kappa	LCL	UCL	P
Kappa 2	0.385	0.257	0.512	< 0.0001

Table 12

MD Documented Indication Compared to White Sheet
ICD-9-CM Text Indication – McNemar's Statistics

	McNemar's	P
Scorer 1 vs 2	4.571	0.033
Scorer 1 vs 3	30.000	< 0.001
Scorer 2 vs 3	22.000	< 0.001

sheet ICD-9-CM text indication show that the disagreement between all three scorers occurred for reasons other than chance.

MD documented indication compared to ISDW ICD-9-CM text indication. The composite results in Table 13 demonstrate that as a group, the scorers viewed the information to be the same 21.0% of the time, different 59.3%, and missing 19.8% of the time. In the case of the comparison of these two data elements, there was a wider variation between the three scorers, than in the two previous comparisons, in the same and different scores. However, in spite of the greater variation, the trend for scorer 3 to rate the information as the same less often and as different more often than either scorer 1 or 2 persisted. In the missing category, all three scorers had the same perception of the degree of missing data.

The result of the Kappa statistic is shown in Table 14. The Kappa statistic indicates that none of the difference between perfect agreement between the raters and agreement expected by chance alone is accounted for by true agreement. Furthermore, the value of Kappa is not statistically significant in this case.

The results of the McNemar statistics are shown in Table 15. The results for the McNemar statistics suggest that the disagreement between all three scorers was not by chance. This is consistent with the Kappa results.

Ward clerk free text indication compared to white sheet ICD-9-CM text indication. The composite results in Table 16 show the scorers perceived the information to be the same in 34.6% of cases, different in 40.7% of cases, and missing in 24.7% of cases. The trend in the relative scores between the individual scorers (scorer 1 and 2 results being similar in all three categories, scorer 3 results

Table 13

MD Documented Indication Compared to ISDW ICD-9-CM
Text Indication – Individual and Composite Rater Scores

Response	Same			Different			Missing		
	Proportion	LCL	UCL	Proportion	LCL	UCL	Proportion	LCL	UCL
Scorer1	0.235	0.142	0.363	0.568	0.436	0.691	0.198	0.114	0.322
Scorer2	0.543	0.412	0.668	0.235	0.142	0.363	0.222	0.132	0.349
Scorer3	0.025	0.005	0.107	0.778	0.651	0.868	0.198	0.114	0.322
Composite	0.210	0.123	0.336	0.593	0.460	0.713	0.198	0.114	0.322

Table 14

MD Documented Indication Compared to ISDW
ICD-9-CM Text Indication – Kappa Statistic

	Kappa	LCL	UCL	P
Kappa 2	- 0.032	- 0.174	0.111	0.6688

Table 15

MD Documented Indication Compared to ISDW ICD-9-CM
Text Indication – McNemar's Statistics

	McNemar's	P
Scorer 1 vs 2	21.552	< 0.001
Scorer 1 vs 3	17.000	< 0.001
Scorer 2 vs 3	42.000	< 0.001

Table 16

Ward Clerk Free Text Indication Compared to White Sheet ICD-9-CM Text Indication – Individual and Composite Rater Scores

Response	Same			Different			Missing		
	Proportion	LCL	UCL	Proportion	LCL	UCL	Proportion	LCL	UCL
Scorer1	0.420	0.298	0.552	0.333	0.222	0.466	0.247	0.152	0.376
Scorer2	0.370	0.254	0.503	0.370	0.254	0.503	0.259	0.161	0.389
Scorer3	0.160	0.086	0.279	0.605	0.472	0.724	0.235	0.142	0.363
Composite	0.346	0.234	0.479	0.407	0.287	0.540	0.247	0.152	0.376

being lower in the same category and higher in the different category, and all three scorer results being similar in the missing category) continued in the comparison of these two data elements.

The result of the Kappa statistic is shown in Table 17. The Kappa statistic indicates that 38.7% of the difference between perfect agreement between the raters and agreement expected by chance alone is accounted for by true agreement. The Kappa of 0.387 suggests the true agreement is only "marginal." The Kappa statistic is statistically significant with a probability of $P < 0.0001$.

The results of the McNemar statistics are shown in Table 18. The results for the McNemar statistics indicate that scorer 1 and scorer 2 generally disagreed by chance and that the disagreement between scorer 3 and both scorers 1 and 2 was not by chance.

Ward clerk free text indication compared to ISDW ICD-9-CM text indication. The composite results in Table 19 show the scorers perceived the information to be the same in 23.5% of cases, different in 35.8% of cases, and missing in 40.7% of

Table 17

Ward Clerk Free Text Indication Compared to White Sheet ICD-9-CM Text Indication – Kappa Statistic

	Kappa	LCL	UCL	P
Kappa 2	0.387	0.241	0.533	< 0.0001

Table 18

Ward Clerk Free Text Indication Compared to White Sheet ICD-9-CM Text Indication – McNemar's Statistics

	McNemar's	P
Scorer 1 vs 2	1.333	0.248
Scorer 1 vs 3	19.174	< 0.001
Scorer 2 vs 3	15.211	< 0.001

Table 19

Ward Clerk Free Text Indication Compared to ISDW ICD-9-CM Text Indication – Individual and Composite Rater Scores

Response	Same			Different			Missing		
	Proportion	LCL	UCL	Proportion	LCL	UCL	Proportion	LCL	UCL
Scorer1	0.247	0.152	0.376	0.346	0.234	0.479	0.407	0.287	0.540
Scorer2	0.370	0.254	0.503	0.198	0.114	0.322	0.432	0.309	0.564
Scorer3	0.012	0.002	0.087	0.580	0.448	0.702	0.407	0.287	0.540
Composite	0.235	0.142	0.363	0.358	0.244	0.491	0.407	0.287	0.540

cases. As in the case of the comparison between the MD Documented Indication and the ISDW ICD-9-CM Text Indication, there was wide variation between the three scorers in the same and different scores, but the trend for scorer 3 to rate the information as the same less often and as different more often than either scorer 1 or 2 persisted, as did the consistency between the scores in the missing category.

The result of the Kappa statistic is shown in Table 20. The Kappa statistic indicates that 6.7% of the difference between perfect agreement between the raters and agreement expected by chance alone is accounted for by true agreement. The Kappa of 0.067 suggests the true agreement is only "marginal." This result is not statistically significant.

The results of the McNemar statistics are shown in Table 21. The results for the McNemar statistics indicate that the disagreement between all three scorers was not by chance.

White sheet ICD-9-CM code indication compared to ISDW ICD-9-CM code indication. The composite results in Table 22 demonstrate that as a group, the scorers viewed the information to be the same in 71.6% of cases, different in 16.0% of cases, and missing in 12.3% of cases. In the comparison of these two data elements, which differ from the others in that they are numeric codes rather than textual representations of the indication for the procedure, the ratings of all three scorers are nearly identical in all three categories.

The result of the Kappa statistic is shown in Table 23. The Kappa statistic further supports the presence of near perfect agreement. The result indicates that 96.8% of the difference between perfect agreement between the raters and agreement

Table 20

Ward Clerk Free Text Indication Compared to ISDW
ICD-9-CM Text Indication – Kappa Statistic

	Kappa	LCL	UCL	P
Kappa 2	0.067	- 0.100	0.234	0.2158

Table 21

Ward Clerk Free Text Indication Compared to ISDW
ICD-9-CM Text Indication – McNemar's Statistics

	McNemar's	P
Scorer 1 vs 2	8.333	0.004
Scorer 1 vs 3	19.000	< 0.001
Scorer 2 vs 3	29.000	< 0.001

Table 22

White Sheet ICD-9-CM Code Indication Compared to ISDW ICD-9-CM
Code Indication – Individual and Composite Rater Scores

Response	Same			Different			Missing		
	Proportion	LCL	UCL	Proportion	LCL	UCL	Proportion	LCL	UCL
Scorer1	0.728	0.598	0.828	0.148	0.077	0.266	0.123	0.060	0.236
Scorer2	0.716	0.585	0.819	0.160	0.086	0.279	0.123	0.060	0.236
Scorer3	0.716	0.585	0.819	0.160	0.086	0.279	0.123	0.060	0.236
Composite	0.716	0.585	0.819	0.160	0.086	0.279	0.123	0.060	0.236

Table 23

White Sheet ICD-9-CM Code Indication Compared to ISDW
ICD-9-CM Text Indication – Kappa Statistic

	Kappa	LCL	UCL	P
Kappa 2	0.968	0.834	1.000	< 0.0001

expected by chance alone is accounted for by true agreement. According to the guidelines of Landis and Koch, the Kappa of 0.968 suggests the true agreement is "excellent." The Kappa statistic is statistically significant with a probability of $P < 0.0001$.

The results of the McNemar statistics are shown in Table 24. The results for the McNemar statistics indicate that the small amount of disagreement between scorers 1 and 2 and 1 and 3 was purely due to chance. No result is recorded for the comparison between scorers 2 and 3, because they were in complete agreement on all scores in the comparison of these two data elements.

White sheet ICD-9-CM text indication compared to ISDW ICD-9-CM text indication. The composite results in Table 25 show the scorers perceived the information to be the same in 44.4% of cases, different in 37.0% of cases, and missing in 18.5% of cases. Once again, there was wide variation between the three scorers in the same and different scores, but the trend for scorer 3 to rate the information as the same less often and as different more often than either scorers 1 or 2 persisted, as did the similarity of the scores in the missing category. Moreover, in those data element comparisons in which there is wide variation between all three

Table 24

White Sheet ICD-9-CM Code Indication Compared to ISDW ICD-9-CM Text Indication – McNemar's Statistics

	McNemar's	P
Scorer 1 vs 2	1.000	0.317
Scorer 1 vs 3	1.000	0.317
Scorer 2 vs 3	-	-

Table 25

White Sheet ICD-9-CM Text Indication Compared to ISDW ICD-9-CM Text Indication – Individual and Composite Rater Scores

Response	Same			Different			Missing		
	Proportion	LCL	UCL	Proportion	LCL	UCL	Proportion	LCL	UCL
Scorer1	0.444	0.320	0.576	0.370	0.254	0.503	0.185	0.104	0.308
Scorer2	0.630	0.497	0.746	0.185	0.104	0.308	0.185	0.104	0.308
Scorer3	0.198	0.114	0.322	0.617	0.484	0.735	0.185	0.104	0.308
Composite	0.444	0.320	0.576	0.370	0.254	0.503	0.185	0.104	0.308

scorers in the same and different categories, there appears to be a trend for scorer 2 to rate the information as the same in both data elements more often and different less often than scorer 1.

The result of the Kappa statistic is shown in Table 26. The Kappa statistic indicates that 29.2% of the difference between perfect agreement between the raters and agreement expected by chance alone is accounted for by true agreement. The Kappa of 0.292 suggests the true agreement is only "marginal." The Kappa statistic is statistically significant with a probability of $P < 0.0001$.

The results of the McNemar statistics are shown in Table 27. The results for the McNemar statistics indicate that the disagreement between all three scorers was due to reasons other than chance.

Primary Study Outpatient Results

The results of the first two data element comparisons in the outpatient section are an indication of the information consumers' perception of the accuracy and completeness of the data. As with the first set of inpatient results, in the first two data element comparisons for the outpatient population, the cases in which the data

Table 26

White Sheet ICD-9-CM Text Indication Compared to ISDW
ICD-9-CM Text Indication – Kappa Statistic

	Kappa	LCL	UCL	P
Kappa 2	0.292	0.152	0.431	< 0.0001

Table 27

White Sheet ICD-9-CM Text Indication Compared to ISDW ICD-9-CM Text Indication – McNemar's Statistics

	McNemar's	P
Scorer 1 vs 2	15.000	< 0.001
Scorer 1 vs 3	20.000	< 0.001
Scorer 2 vs 3	35.000	< 0.001

elements are scored as the same represent the cases with accurate data. Cases in which the data elements are scored as different represent the cases with inaccurate data. The proportion of cases scored as having missing data represents the frequency of cases with incomplete data.

The subsequent three data element comparisons in the outpatient results section are an indication of the information consumers' perception of the consistency and completeness of the data. In these data element comparisons, the cases in whi ch the data elements are scored as the same represent the cases with consistent data. Cases in which the data elements are scored as different represent the cases with inconsistent data. The proportion of cases scored as having missing data represents the frequency of cases with incomplete data.

<u>Imaging services scheduler documented indication compared to white sheet ICD-9-CM text indication.</u> The composite results in Table 28 demonstrate that overall, the scorers perceived the information in these two data sources to be the same 53.8% of the time and different 46.2% of the time. None of the scorers rated any of the information as missing. The individual results show that scorers 1 and 2 had

Table 28

Imaging Services Scheduler Documented Indication Compared to White Sheet ICD-9-CM Text Indication – Individual and Composite Rater Scores

Response	Same			Different			Missing		
	Proportion	LCL	UCL	Proportion	LCL	UCL	Proportion	LCL	UCL
Scorer1	0.577	0.414	0.725	0.423	0.275	0.587	0.000	0.000	0.099
Scorer2	0.558	0.396	0.709	0.442	0.291	0.604	0.000	0.000	0.099
Scorer3	0.327	0.195	0.493	0.673	0.507	0.805	0.000	0.000	0.099
Composite	0.538	0.377	0.691	0.462	0.309	0.623	0.000	0.000	0.099

similar scores in the same and different categories. However, scorer 3 tended to rate the information as the same less often and as different more often than both scorer 1 and scorer 2. In contrast to the scoring in the same and different categories, all three scorers indicated a similar frequency of missing data.

The Kappa statistic result is shown in Table 29. The result suggests that 59.0% of the difference between perfect agreement between the raters and agreement expected by chance alone is accounted for by true agreement. Using the criteria of Landis and Koch,[179] the Kappa of 0.590 denotes "good" agreement between the raters. The Kappa statistic is statistically significant with a probability of $P < 0.0001$.

The results of the McNemar's statistics are shown in Table 30. The results suggest that scorer 1 and scorer 2 generally disagreed by chance. In contrast, the disagreement between scorer 3 and both scorers 1 and 2 was not by chance. No Bowker's statistics are presented in Appendix G for this data comparison. Because there was no disagreement in the missing category, no additional information would have been derived from calculating the Bowker's in addition to the McNemar's test

Table 29

Imaging Services Scheduler Documented Indication Compared to White Sheet ICD-9-CM Text Indication – Kappa Statistic

	Kappa	LCL	UCL	P
Kappa 2	0.590	0.433	0.747	< 0.0001

Table 30

Imaging Services Scheduler Documented Indication Compared to White Sheet ICD-9-CM Text Indication – McNemar's Statistics

	McNemar's	P
Scorer 1 vs 2	0.200	0.655
Scorer 1 vs 3	13.000	< 0.001
Scorer 2 vs 3	10.286	< 0.001

statistic.

Imaging services scheduler documented indication compared to ISDW ICD-9-CM text indication. The composite results in Table 31 show that as a group the scorers perceived the information to be the same 46.2% of the time, different 36.5% of the time, and missing 17.3% of the time. The trends in the relative individual scores are the same as in the data element comparison between the imaging services scheduler and the ISDW ICD-9-CM codes.

The result of the Kappa statistic is shown in Table 32. The Kappa statistic indicates that 50.2% of the difference between perfect agreement between the raters

Table 31

Imaging Services Scheduler Documented Indication Compared to ISDW ICD-9-CM Text Indication – Individual and Composite Rater Scores

Response	Same			Different			Missing		
	Proportion	LCL	UCL	Proportion	LCL	UCL	Proportion	LCL	UCL
Scorer1	0.500	0.343	0.658	0.327	0.195	0.493	0.173	0.082	0.329
Scorer2	0.538	0.377	0.691	0.288	0.165	0.453	0.173	0.082	0.329
Scorer3	0.269	0.150	0.434	0.558	0.396	0.709	0.173	0.082	0.329
Composite	0.462	0.309	0.623	0.365	0.226	0.531	0.173	0.082	0.329

Table 32

Imaging Services Scheduler Documented Indication Compared to ISDW ICD-9-CM Text Indication – Kappa Statistic

	Kappa	LCL	UCL	P
Kappa 2	0.502	0.330	0.675	< 0.0001

and agreement expected by chance alone is accounted for by true agreement. The Kappa of 0.502 suggests this is "good" true agreement. The Kappa statistic is statistically significant with a probability of $P < 0.0001$.

The results of the McNemar's statistics are shown in Table 33. As in the previous data element comparison, the results indicate that scorer 1 and scorer 2 disagreed by chance and that there were reasons other than chance that led scorer 3 to disagree with both scorer 1 and scorer 2.

White sheet ICD-9-CM code indication compared to ISDW ICD-9-CM code indication. The composite results in Table 34 show the scorers perceived the

Table 33

Imaging Services Scheduler Documented Indication Compared to ISDW ICD-9-CM Text Indication – McNemar's Statistics

	McNemar's	P
Scorer 1 vs 2	0.667	0.414
Scorer 1 vs 3	12.000	< 0.001
Scorer 2 vs 3	14.000	< 0.001

Table 34

White Sheet ICD-9-CM Code Indication Compared to ISDW ICD-9-CM Code Indication – Individual and Composite Rater Scores

Response	Same			Different			Missing		
	Proportion	LCL	UCL	Proportion	LCL	UCL	Proportion	LCL	UCL
Scorer1	0.885	0.739	0.954	0.038	0.008	0.160	0.077	0.025	0.213
Scorer2	0.885	0.739	0.954	0.038	0.008	0.160	0.077	0.025	0.213
Scorer3	0.885	0.739	0.954	0.038	0.008	0.160	0.077	0.025	0.213
Composite	0.885	0.739	0.954	0.038	0.008	0.160	0.077	0.025	0.213

information to be the same in 88.5% of cases, different in 3.8% of cases, and missing in 7.7% of cases. In the case of the comparison of these two data elements, the relative scores of all three scorers were identical. This is supported by the results of the Kappa statistic.

The result of the Kappa statistic is shown in Table 35. The Kappa statistic indicates that 100% of the difference between perfect agreement between the raters and agreement expected by chance alone is accounted for by true agreement. The

Table 35

White Sheet ICD-9-CM Code Indication Compared to
ISDW ICD-9-CM Code Indication – Kappa Statistic

	Kappa	LCL	UCL	P
Kappa 2	1.000	0.837	1.000	< 0.0001

Kappa statistic is statistically significant with a probability of $P < 0.0001$.

Because there was no disagreement, no McNemar or Bowker's statistics were calculated.

White sheet ICD-9-CM text indication compared to ISDW ICD-9-CM text indication. The composite results in Table 36 demonstrate that as a group, the scorers viewed the information to be the same 61.5% of the time, different 21.2%, and missing 17.3% of the time. The individual scores once again demonstrate that scorer 3 rated the information as the same less often and as different more often than either scorers 1 or 2 and that all three scorers had the same perception of the degree of missing data.

The result of the Kappa statistic is shown in Table 37. The Kappa statistic indicates that 44.0% of the difference between perfect agreement between the raters and agreement expected by chance alone is accounted for by true agreement. The Kappa statistic is significant with a probability of $P < 0.0001$.

The results of the McNemar statistics are shown in Table 38. The results for the McNemar statistics suggest that the disagreement between scorer 1 and scorer 2 was by chance and that between scorer 3 and the other two scorers was not by chance.

Table 36

White Sheet ICD-9-CM Text Indication Compared to ISDW ICD-9-CM Text Indication – Individual and Composite Rater Scores

Response	Same			Different			Missing		
	Proportion	LCL	UCL	Proportion	LCL	UCL	Proportion	LCL	UCL
Scorer1	0.673	0.507	0.805	0.154	0.070	0.307	0.173	0.082	0.329
Scorer2	0.635	0.469	0.774	0.192	0.095	0.350	0.173	0.082	0.329
Scorer3	0.442	0.291	0.604	0.385	0.243	0.550	0.173	0.082	0.329
Composite	0.615	0.450	0.757	0.212	0.109	0.373	0.173	0.082	0.329

Table 37

White Sheet ICD-9-CM Text Indication Compared to ISDW ICD-9-CM Text Indication – Kappa Statistic

	Kappa	LCL	UCL	P
Kappa 2	0.440	0.268	0.613	< 0.0001

Table 38

White Sheet ICD-9-CM Text Indication Compared to ISDW ICD-9-CM Text Indication – McNemar's Statistics

	McNemar's	P
Scorer 1 vs 2	0.667	0.414
Scorer 1 vs 3	12.000	< 0.001
Scorer 2 vs 3	8.333	0.004

White sheet ICD-9-CM code indication compared to admit diagnosis ICD-9-CM code indication. The composite results in Table 39 show the scorers perceived the information to be the same in 36.5% of cases, different in 40.4% of cases, and missing in 23.1% of cases. As in the case of the comparison between the White Sheet ICD-9-CM code indication and the ISDW ICD-9-CM code indication, the relative scores of all three scorers were identical.

The result of the Kappa statistic, shown in Table 40, indicates that 100% of the difference between perfect agreement between the raters and agreement expected by chance alone is accounted for by true agreement. The Kappa of 1.000 suggests the true agreement is "excellent." This result is statistically significant with a probability of $P < 0.0001$.

Because of the complete agreement, no Bowker or McNemar statistics were calculated.

Table 39

White ICD-9-CM Code Indication Compared to Admit Diagnosis ICD-9-CM Code Indication – Individual and Composite Rater Scores

Response	Same			Different			Missing		
	Proportion	LCL	UCL	Proportion	LCL	UCL	Proportion	LCL	UCL
Scorer1	0.365	0.226	0.531	0.404	0.259	0.568	0.231	0.122	0.393
Scorer2	0.365	0.226	0.531	0.404	0.259	0.568	0.231	0.122	0.393
Scorer3	0.365	0.226	0.531	0.404	0.259	0.568	0.231	0.122	0.393
Composite	0.365	0.226	0.531	0.404	0.259	0.568	0.231	0.122	0.393

Table 40

White Sheet ICD-9-CM Code Indication Compared to Admit Diagnosis ICD-9-CM Code Indication – Kappa Statistic

	Kappa	LCL	UCL	P
Kappa 2	1.000	0.820	1.000	< 0.0001

Additional Inpatient and Outpatient Results

The results reported in the inpatient and outpatient sections above include the individual and composite proportions and confidence intervals, Kappa 2 (calculated using only the same and different response categories), and McNemar's statistic (calculated using only the same and different response categories). The reason that Kappa 1 (calculated using all three response categories) is not reported in the presentation of the data above is that as expected, when removing the data on the missing category (where all three scorers clearly agreed in all cases), the level of agreement remained the same or decreased when the missing data scores were removed. Therefore, Kappa 2 proved to be a more accurate representation of the level of agreement between the raters on the frequency of accurate or consistent information and the frequency of errors of commission.

The reason for reporting only the McNemar's statistics is similar to the reason for reporting only Kappa 2. Although the values of the Bowker's Extension statistics is higher than or the same as those of the McNemar's statistics for the same scorer comparison for the same data element pair in all cases, the statistical significance of the two statistics for each scorer comparison within the same data element comparison did not differ. Therefore, although either the Bowker's or McNemar's statistics could be reported in the summaries below, for consistency (using only the same and different category data) with the reporting of Kappa 2, the McNemar's statistics are reported.

The complete results on agreement (Kappa 1 and Kappa 2) and the Bowker's Extension results for all data element comparisons are presented in Appendix F for inpatients and Appendix G for outpatients.

Primary Study Comparison of Inpatient and Outpatient Results

One of the questions posed in this evaluation of information quality problems in the HELP System and ISDW was whether there was a difference in the information consumers' perception of the quality of information when comparing the inpatient and outpatient populations. The relevance of the answer to this question is, that if a difference does exist, the likely cause was the differences described in the ordering, performance, and billing procedures between the populations. To explore this question, an analysis of association between the rater responses and the inpatient-outpatient nature of the data was performed on those data element pairs present in both the inpatient and outpatient data sets, using Fisher's Exact Test for difference in proportions. Two Fisher's Exact Tests were performed: The first test, with the resulting probability value denoted as P_O, was performed on the entire data set (same, different, and missing) for each scorer. The second test, for which the resulting probability value is denoted as P_M, was performed only on the data scores in the missing category. Because the scores in the missing categories were nearly always consistent between the raters, within the inpatient and outpatient populations, this test was performed to determine whether the information consumers' perception of the proportion of missing data between the inpatient and outpatient populations showed the same consistency. If so, this would support a conclusion that determining the completeness of the data evaluated in this thesis research was relatively straightforward and not significantly influenced by the factors that did lead to differing perceptions of the accuracy and consistency of the data.

MD documented/imaging services scheduler documented indication compared to white sheet ICD-9-CM text indication. The MD documented indication and imaging services scheduler documented indication are the inpatient and outpatient representations, respectively, of the same data element. In both cases, these data elements were compared to the white sheet ICD-9-CM text indication for the requested imaging procedure. The results of the comparison of these two data element pairs are presented in Table 41.

In the individual rater results, scorer 3 rates a significantly higher proportion of "different" and a lower proportion of "same" in the outpatient results when compared to the inpatient results. In contrast, there is no significant difference in the results of either scorer 1 or scorer 2. The proportions of "missing" do not test different between the inpatient and outpatient results.

MD documented/imaging services scheduler documented indication compared to ISDW ICD-9-CM text indication. The results of the comparison of these two data element pairs are presented in Table 42. In the individual rater results, both scorer 1

Table 41

MD Documented/Imaging Services Scheduler Documented Indication Compared to White Sheet ICD-9-CM Text Indication – Fisher's Exact Test

	Scorer1		Scorer2		Scorer3	
	Inpt	Outpt	Inpt	Outpt	Inpt	Outpt
Same	0.577	0.494	0.558	0.395	0.327	0.124
Different	0.423	0.494	0.442	0.58	0.673	0.864
Missing	0.000	0.012	0.000	0.025	0.000	0.012
P_O	0.681		0.117		0.007	
P_M	1.000		0.520		1.000	

Table 42

MD Documented/ Imaging Services Scheduler Documented Indication Compared to ISDW ICD-9-CM Text – Fisher's Exact Test

	Scorer1		Scorer2		Scorer3	
	Inpt	Outpt	Inpt	Outpt	Inpt	Outpt
Same	0.500	0.235	0.539	0.543	0.269	0.025
Different	0.327	0.568	0.289	0.235	0.558	0.778
Missing	0.173	0.198	0.173	0.222	0.173	0.198
P_O	0.006		0.681		0.000	
P_M	0.820		0.518		0.820	

and scorer 3 rate a significantly higher proportion of "different" and a lower proportion of "same" in the outpatient results when compared to the inpatient results. Scorer 2 shows no significant difference in the results of the inpatient and outpatient scores. As in the previous comparison, the proportions of "missing" do not test different between the inpatient and outpatient results.

White sheet ICD-9-CM text indication compared to ISDW ICD-9-CM text indication. The results of the comparison of these two data element pairs are presented in Table 43. The individual rater results again show both scorer 1 and scorer 3 rate a significantly higher proportion of "different" and a lower proportion of "same" in the outpatient results when compared to the inpatient results. Scorer 2 shows no significant difference in the results of the inpatient and outpatient scores. Furthermore, also as in the previous comparison, the proportions of "missing" do not test different between the inpatient and outpatient results.

Table 43

White Sheet ICD-9-CM Text Indication
Compared to ISDW ICD-9-CM Text
Indication – Fisher's Exact Test

	Scorer1		Scorer2		Scorer3	
	Inpt	Outpt	Inpt	Outpt	Inpt	Outpt
Same	0.673	0.444	0.635	0.630	0.442	0.198
Different	0.154	0.370	0.192	0.185	0.385	0.617
Missing	0.173	0.185	0.173	0.185	0.173	0.185
P_O	0.014		1.000		0.008	
P_M	1.000		1.000		1.000	

White sheet ICD-9-CM code indication compared to ISDW ICD-9-CM code indication. The results of the comparison of these two data element pairs are presented in Table 44. The individual rater results show both scorer 2 and scorer 3 rate a significantly higher proportion of "different" and a lower proportion of "same" in the outpatient results when compared to the inpatient results. In this case, scorer 1 shows no significant difference in the results of the inpatient and outpatient scores. The proportions of "missing," once again, do not test different between the inpatient and outpatient results.

Secondary Study Results

The radiologists' reports of the results of the imaging procedures were considered "high quality" when they contained the answer to the clinical information being sought, as indicated by the "true" clinical indication for the procedure. The results of the Fisher's Exact Test did not show statistically significant evidence for an association between the accuracy of the information the radiologist received and the

Table 44

White Sheet ICD-9-CM Code Indication Compared
to ISDW ICD-9-CM Code Indication
– Fisher's Exact Test

	Scorer1		Scorer2		Scorer3	
	Inpt	Outpt	Inpt	Outpt	Inpt	Outpt
Same	0.885	0.728	0.885	0.716	0.885	0.716
Different	0.039	0.148	0.039	0.161	0.039	0.161
Missing	0.077	0.124	0.077	0.124	0.077	0.124
P_O	0.075		0.047		0.047	
P_M	0.565		0.565		0.565	

quality of the report. For both the inpatient and outpatient populations, the probability value for the Fisher's Exact Test was 1.0.

CHAPTER IV

DISCUSSION

Summary and Conclusions

> In attempting to arrive at the truth, I have applied everywhere for information, but in scarcely an instance have I been able to obtain hospital records fit for any purpose of comparison. If they could be obtained, they would enable us to decide many other questions. They would show subscribers how their money was being spent and what amount of good was really being done with it.[183]

In this 1873 quote, Florence Nightingale pointed out the potential value of using medical records to evaluate the processes and outcomes of healthcare delivery. She also clearly stated that one of the barriers to evaluating the processes and outcomes of care was the quality of the (medical) records. Although there may have been several components to the poor quality to which Florence Nightingale was referring, it is likely that the quality of the information itself was a significant component. Information is increasingly being recognized as one of the most important assets that healthcare organizations have to support daily patient care, daily business operations, and the evaluation and management of processes and outcomes. Along with this growing recognition and treatment of information as an asset is an increasing awareness of the challenges of measuring, improving, and maintaining the quality of information.

The purpose of this thesis research was to contribute to the body of knowledge that will support the development of processes for addressing information quality problems in healthcare. The broad objective of the research was to evaluate the quality of imaging services data in the HELP System and the Imaging Services Data Warehouse at LDS Hospital. There were two components to the research performed for this thesis. The primary evaluation focused on the sources, nature, volume, and information consumers' perspective of imaging services information quality problems. The specific aims of the primary research were to 1) identify quality problems pertaining to the information about the clinical indication for ordering an imaging procedure, 2) determine the probable sources of the imaging services information quality problems pertaining to the clinical indication for imaging procedures, and 3) quantify information quality errors pertaining to the clinical indication for ordering an imaging procedure. The secondary evaluation, a pilot study, focused on the impact of imaging services IQ problems on the quality of imaging study reports. The specific aim of the secondary study was to determine if the quality of the imaging study report was affected by the quality of the information received by the interpreting radiologist about the indication for the study. By using a combination of qualitative and quantitative methods, the specific aims for both the primary and secondary evaluations were achieved.

Primary Study Qualitative Analyses Conclusions

The qualitative analyses were performed to gain an understanding of 1) the components of the LDS Hospital processes for ordering, performing, and interpreting an imaging procedure, 2) the way in which data flowed to support the steps in these processes, 3) the downstream uses of the data captured during these processes, and 4) the processes for the downstream uses of the data. The analyses were divided into the following process categories: procedure ordering, procedure performance, procedure interpretation and reporting, and facility (technical) and professional services billing. The results of the qualitative analyses identified from the information producer and consumers' perspectives the type of information quality problems that occur and demonstrated the presence of variability between the inpatient and outpatient populations in the way an individual process is performed. The results also demonstrated that variability in the process steps and the people who performed them occurred not only between the populations but also within each of the populations. For example, the diagnosis coding process for technical services billing differed between the inpatient and outpatient populations, as well as within the outpatient population. The variability within the outpatient population depended on who the patient was insured by and what kind of imaging procedure the patient had undergone. In addition to the process issues, the qualitative analyses also identified technical issues that presented potential sources of information quality problems. The primary technical issue identified as a potential source of information quality problems, as described in the description of the HELP System in Chapter II, was the presence of two different ICD-9-CM text to PTXT mappings (i.e., DC 20 and 24) for the same ICD-9-CM code. The significance of the process variations and technical issues was that they led to information quality problems such as those presented in Appendix D and Appendix E.

Primary Study Quantitative Analyses Conclusions

The quantitative analyses were designed to obtain the information consumers' perspective on the accuracy, consistency, and completeness of the information pertaining to the clinical indication for an imaging procedure. Accuracy was defined as the information consumers' perception as to whether the information about the clinical indication for a procedure in the various steps of the imaging procedure ordering process was an accurate representation of the true clinical indication. The gold standard for the true clinical indication for the inpatient population was the indication documented by the ordering physician (MD) on the original request form. For the outpatient population, the gold standard for the true indication was the first documented source of information accessible for review, which was the clinical indication documented by the imaging services scheduler (ISS). In the evaluation of accuracy, the frequency of accurate data was defined as the proportion of cases in which the data elements, one of which was the true indication, were scored as the same. The frequency of inaccurate data was represented by the proportion of cases in which the data elements were scored as different. The dimension of consistency was defined as the information consumers' perception as to whether the information on the clinical indication for a procedure in each data source other than the gold standard was consistent with the clinical indication in all the other data sources. This was an extension of the dimension of accuracy. The frequency of consistent data was defined as the proportion of cases in which the data elements in the data sources other than the gold standard were scored as the same. The frequency of inconsistent data was represented by the proportion of cases in which the data elements were scored as different. The dimension of completeness was defined as the presence of a clinical indication in each step of the information flow in the ordering process for an imaging procedure. The frequency of incomplete data was represented by the proportion of cases scored as having missing data.

The summary results of the inpatient and outpatient data analyses are shown in Tables 45, 46, 47, and 48. Each row in these tables represents one data element comparison performed across all cases in the study population. The data element comparisons are represented by the letters "A" through "H." In parentheses next to the identifying letter is the table number from Chapter III in which the complete results for that comparison can be found. The table columns represent the accuracy, inaccuracy, consistency, inconsistency, and incompleteness of the data in each category of data element comparisons.

In some cases, one of the two data elements being compared in one patient population (inpatient or outpatient) existed in only that population and therefore there was no equivalent comparison between the two populations. In those cases where a particular data element comparison was performed in only one of the two populations, no results are shown for whichever population the data element comparison was not performed.

In the inpatient population, the frequency of cases with inaccurate information ranged from 22.5% to 63.0%, the frequency of cases with inconsistent information ranged from 16.0% to 40.7%, and the frequency of cases with incomplete information ranged from 1.2% to 40.7%. In the outpatient population, the frequency of cases with inaccurate information ranged from 36.5% to 46.2%, the frequency of cases with

Table 45

Summary Results of Inpatient Data Element Comparisons
For Accuracy and Completeness

	Accurate	Inaccurate	Incomplete
A (Table 7)	52.5%	22.5%	25.0%
B (Table 10)	35.8%	63.0%	1.2%
C (Table 13)	21.0%	59.3%	19.8%

A = MD Documented Indication Compared to Ward Clerk Free Text Indication
B = MD Documented Indication Compared to White Sheet ICD-9-CM Text Indication
C = MD Documented Indication Compared to ISDW ICD-9-CM Text Indication

Table 46

Summary Results of Outpatient Data Element Comparisons
For Accuracy and Completeness

	Accurate	Inaccurate	Incomplete
A	-	-	-
B (Table 28)	53.8%	46.2%	0.0%
C (Table 31)	46.2%	36.5%	17.3%

A = No equivalent to inpatient population
B = ISS* Documented Indication Compared to White Sheet ICD-9-CM Text Indication
C = ISS Documented Indication Compared to ISDW ICD-9-CM Text Indication
*Imaging Services Scheduler data element is equivalent to "MD" data element in the inpatient population

Table 47

Summary Results of Inpatient Data Element Comparisons
For Consistency and Completeness

	Consistent	Inconsistent	Incomplete
D (Table 16)	34.6%	40.7%	24.7%
E (Table 19)	23.5%	35.8%	40.7%
F (Table 22)	71.6%	16.0%	12.3%
G (Table 25)	44.4%	37.0%	18.5%
H	-	-	-

D = Ward Clerk Free Text Indication Compared to White Sheet ICD-9-CM Text Indication
E = Ward Clerk Free Text Indication Compared to ISDW ICD-9-CM Text Indication
F = White Sheet ICD-9-CM Code Indication Compared to ISDW ICD-9-CM Code Indication
G = White Sheet ICD-9-CM Text Indication Compared to ISDW ICD-9-CM Text Indication
H = No equivalent to outpatient population

Table 48

Summary Results of Outpatient Data Element Comparisons
For Consistency and Completeness

	Consistent	Inconsistent	Incomplete
D	-	-	-
E	-	-	-
F (Table 34)	88.5%	3.8%	7.7%
G (Table 36)	61.5%	21.2%	17.3%
H (Table 39)	36.5%	40.0%	23.1%

D = No equivalent to inpatient population
E = No equivalent to inpatient population
F = White Sheet ICD-9-CM Code Indication Compared to ISDW ICD-9-CM Code Indication
G = White Sheet ICD-9-CM Text Indication Compared to ISDW ICD-9-CM Text Indication
H = White Sheet ICD-9-CM Code Indication Compared to Admit Diagnosis ICD-9-CM Code Indication

inconsistent information ranged from 3.8% to 40.4%, and the frequency of cases with incomplete information ranged from 0% to 23.1%. These results confirmed the presence of information quality problems in the HELP System and ISDW, in all three dimensions of IQ evaluated. In previous studies of information quality in CPRs and data warehouses,[107,108,117,120,121,125,128,129] the definition and frequency of accurate information varied, and none of the studies defined or measured the dimension of consistency. The results for the dimension of accuracy in the previous studies ranged from 5% to 51% of records being inaccurate. The results for completeness in these same studies ranged from 0% to 70% of records having missing information. Thus, the results of this thesis research are consistent with the results of previous studies.

The data elements evaluated in this research can be used for a variety of purposes including automated clinical decision support, imaging services outcomes measurement, and operational processes such as billing. The presence of the information quality problems identified brings into question the ability to perform or rely on the results of any of the activities that use the data elements evaluated. The results also suggest the need to reassess and redesign, where process issues are identified, the imaging services processes for procedure ordering, procedure performance, procedure interpretation and reporting, and facility (technical) and professional services billing

Analysis of the association between the inpatient and outpatient nature of the data and the individual scorers' perception of the accuracy, consistency, and completeness of the data was performed for the four data element comparisons in which the same data elements existed in both populations. The results of the analyses showed no association for the dimension of completeness. The results did show an association between the IQ dimension of accuracy and the IQ dimension of consistency and the inpatient-outpatient nature of the data. The association showed a tendency for the accuracy and the consistency for the data element pairs compared in the two populations to be rated as higher in the inpatient population than in the outpatient population (Tables 41, 42, 43, and 44). These results likely reflect a better practice in the inpatient processes than the outpatient processes that pertain to these data elements.

Interrater agreement about the accuracy, consistency, and completeness of the information was predominantly marginal (Kappa less than 0.4) in the inpatient population and good (kappa between 0.4 and 0.75) to excellent (kappa greater than 0.75) in the outpatient population.[179] The marginal agreement in the inpatient population was primarily driven by scorer 3, who was a general surgeon. Scorer 3 consistently scored the results differently from both of the other two scorers, who were a diagnostic radiologist and a radiation oncologist. The influence of scorer 3 on the agreement statistics for the inpatient population is further supported by the results of the evaluation of disagreement between the scorers for the inpatient population results. The results of disagreement demonstrate that scorers 1 and 2 consistently disagreed by chance, and scorer 3 consistently disagreed from both scorer 1 and scorer 2 for reasons other than chance. The difference in perception between scorer 3 and the other two scorers is likely a reflection of the difference in specialty. In the experience of the author of this thesis, surgeons, who are usually in the position of providing the information on the clinical indication for the imaging procedure, generally have a different perspective from diagnostic radiologists and radiation oncologists on the nature of the information required for performing and interpreting imaging procedures. The difference in perspective is that radiologists and radiation oncologists believe they require a greater range and depth of clinical information than surgeons believe radiologist and radiation oncologists require.

Secondary Study Qualitative Analyses Conclusions

The qualitative analysis results that pertain to the secondary study revealed process issues that negatively impacted all three dimensions of information quality being evaluated (accuracy, consistency, and completeness). Although not specifically evaluated in the quantitative analysis, unavailability of a data source (e.g., lost white sheet or prior studies for comparison) to a radiologist interpreting an imaging study was an additional quality issue identified for the dimension of completeness. For example, even though the white sheet on a particular patient may have all necessary information, the interpreting radiologist may not receive the white sheet at the time required for interpretation of the imaging study. The significance of this finding is that research has shown a negative impact on the effectiveness and efficiency of radiologists' ability to identify abnormalities and make decisions about the diagnostic and therapeutic relevance of detected abnormalities, in the absence of timely, accurate, and relevant clinical and historical information about patients.[164-169]

Secondary Study Quantitative Analyses Conclusions

The quantitative analysis in the secondary study focused only on looking for an association between the accuracy of the clinical information received by the interpreting radiologists and the quality of the reports they produced. As stated in the Results section of this thesis (Chapter III), no statistically significant association was found. The absence of a significant association in the secondary study performed for this thesis is inconsistent with the majority of previous research performed on the topic of the effects of clinical and historical information received by radiologists and the quality of the reports they produce. [164-169] Two possible reasons for the absence of a statistically significant association exist: 1) the small sample size (N = 40) used for the pilot study, or 2) in those cases where the information on the imaging study request form and/or white sheet received by the interpreting radiologist was inaccurate, the radiologist may have used other means to acquire clinical information about the patient, such as reviewing the patient's chart or contacting the requesting physician, and thus ultimately had accurate clinical information.

Practical Application of Results

One of the most important questions to ask about any research is, "What is the practical application of the results?" In the case of the research performed for this thesis, the results were used to support the work being performed by the Intermountain Health Care Imaging Compliance Committee. The role of the Imaging Compliance Committee was to ensure that the processes and practices of the IHC Imaging Services Division, including those for imaging procedure requests, coding, and billing, are consistent with the federal regulations and guidelines administered by the United States Office of the Inspector General (OIG) and the Health Care Financing Administration (HFCA). The Imaging Compliance Committee reports to the IHC Radiology Advisory Committee and IHC Imaging Guidance Council. The Radiology Advisory Committee was responsible for policies and procedures that pertain to the facility/technical aspects of IHC Imaging Services, and the Imaging Guidance Council is responsible for the professional/clinical aspects of IHC Imaging Services.

Based on the results of this thesis research and the results of investigations performed by subsequently formed IHC CPT and ICD-9-CM Coding Task Forces, the Imaging Compliance Committee made the following recommendations to the Radiology Advisory Committee and Imaging Guidance Council:

1. Intermountain Health Care should adopt a single radiology information system for all locations, inpatient and outpatient (ambulatory), that provide imaging services.
2. Clinicians must provide a clearly documented clinical indication (signs, symptoms, suspected diagnosis, etc.) for all imaging procedures being requested.
3. Clinicians must clearly document what imaging procedure they are requesting.
4. IHC Health Information Services should accept responsibility for both the facility (technical) and professional coding processes.
5. Facility and professional billing only should occur after the imaging procedure is completed, including having a signed dictation by the interpreting radiologist.

It was the belief of the Compliance Committee, that following these recommendations would improve the quality of information used for coding and billing practices and thus ensure all that physician professional services provided would be properly documented, that all bills would accurately reflect the services provided, and that only accurately and properly documented services would be billed. In addition, because the same information that was used in coding and billing for services performed was used in benchmarking and performing evaluations of delivery processes and outcomes, the improvement in information quality was anticipated to enhance the quality and reliability of these evaluations.

At the time of writing this thesis, IHC had adopted recommendations 1, 2, and 3. Recommendation 4 had not been adopted, and steps were being taken to move in the direction of recommendation 5.

In addition to the recommendations of the Imaging Compliance Committee, several changes to the HELP System could be made to improve information quality. In the imaging service ordering process, data may be entered either from a pick list or by using free text. Evaluations of the imaging services processes demonstrated that the use of free text by the ward clerks in the order request process in some cases led to information quality problems. Specifically, if there was no match for the ward clerk free text on the pick list used by imaging clerks to complete the order, this would negatively impact accuracy and consistency of the information. This fact points to the potential value of eliminating the option for the use of free text data entry and requiring the use of predefined pick lists. Although requiring selection of the indication from a predefined pick list would not completely eliminate the potential for inaccurate reasons being entered into the system, it would support the standardization of terms used for clinical indications and potentially decrease the likelihood of having "useless" reasons such as post-op, entered into the system. An additional benefit of this step would be a more standard (use of a pick list) mechanism for entry of the information about the clinical indication for the exam. The value of standard data entry processes can be inferred from the work of Wyatt,[119] Dambro and Weiss,[118] and Wilton and Pennisi,[128] who demonstrated that the use of multiple mechanisms for data entry negatively entry impacts data quality. Because some people enter free text as a way to avoid having to spend the time searching a predefined list for the term they want, in order not to frustrate the user it will be necessary to ensure that searching the pick list is intuitive and easy.

A second change that could be made to the HELP System that would potentially improve the quality of Imaging Services data would be to have automatic, field-level data validation performed. Specifically, I would recommend checking for blank fields and not allowing an order to be completed if a required field is blank. This would decrease the frequency of incomplete information. The third recommended change to the HELP System would be standardization of the PTXT dictionary for ICD-9-CM codes. The use of two different dictionaries for two different data classes created a situation in which the same code could have two different text descriptions and thus create consistency problems.

Study Limitations

The three primary limitations and potential sources of bias for this study are the sample sizes, the medical specialty and clinical experience of the scorers, and the absence of control of the process by which radiologists acquired the clinical information on which the secondary study was based.

The sample size for the primary study was 114 patients and for the secondary study was 40 patients. A criticism could be leveled that these sample sizes are too small to determine the statistical significance of the study results. There are three reasons why these sample sizes may, in fact, be considered adequate. First, with respect to the primary study, as stated in the Materials and Methods section of this thesis (Chapter II), statistical significance of the information quality measures was not evaluated. The reason was that an assumption was made that any degree of information quality errors has potentially significant implications in the delivery of care; therefore, the presence or absence of statistical significance in the frequency of errors is irrelevant. Second, because this research was performed as a retrospective correlational evaluation study, there was no expected "effect size" or expected population mean that could be included in a sample size calculation. Third, in those portions of the analyses in which statistical tests were used, the tests used were nonparametric. The relevance of the use of nonparametric tests is that nonparametric tests can be performed with sample sizes as small as 10 without the same concerns about the impact on validity of the results that one would have with parametric tests. For these reasons, the sample sizes were chosen based on the practical limits of the volume of work that could be reasonably expected from the scorers, who were volunteering their time.

The assessment of the quality of the information evaluated in this research was based on the subjective impression of the three scorers participating in the study. On the one hand, the subjectivity of the scoring process could be criticized as presenting uncontrolled opportunities for bias. The potential sources of bias include the medical specialty and individual type of clinical experiences of the scorers. However, the definition of information quality emphasizes the basic concept that good quality information is "information that is fit for use by information consumers" and therefore is dependent on what quality means to the individual consumer. For this reason, the presence of individual bias of the scorers is an integral aspect of the assessment of the quality of the information.

The secondary study assessed the association between the quality of the clinical information received by the interpreting radiologist and the quality of their dictated report. The assessment of the quality of the clinical information received by the radiologist was based only on the information they received that was provided through the data sources evaluated in this study. Because of the design of this study, there was no way to ensure the radiologists did not have additional information. Therefore, in those cases in which there was no evidence of association between poor quality clinical information and poor quality reports (i.e., the radiologist did not have good quality information and the quality of the report was high), if the radiologist in fact had additional, good quality information, the conclusions of the assessment may have been incorrect.

Recommendations

Measuring, improving, and maintaining the quality of information depend as much on organizational culture, organizational politics (who "owns" the information), and process as it does on technical factors.[31, 34, 184, 185] One of the core principles of an organizational approach to managing and improving information quality is treating information as an asset. Treating information as an asset can be achieved by developing and implementing an organization-wide information quality program using the basic principles of continuous quality improvement. The primary components of such a program include having inventories of the types, meaning, locations, and quality of the data that exist in the organization (i.e., having a metadata repository); understanding the processes for creating, storing, and using the data; having standard data definitions and a well-defined, continuously evolving data model; having standard processes for ensuring data quality as part of all organizational processes and projects (Figure 24); and having processes, including ongoing education, for ensuring the data are turned into information and used appropriately.

The components of an organizational information quality program can be categorized into a continuously repeating series of steps consisting of 1) assessments of various aspects of the quality of the information and associated information-related processes and information systems, 2) prioritization of the information quality issues identified during the assessment, 3) data cleansing, 4) technical and/or process redesign, and 5) education.[31,34,184,185] Although these steps begin sequentially, over time, as multiple specific information quality improvement projects are undertaken, the various components of the program begin operating in parallel.

Assessment Phase

The assessment phase of the information quality management process includes an in-depth review of the organization's data model(s); metadata; business rules; each

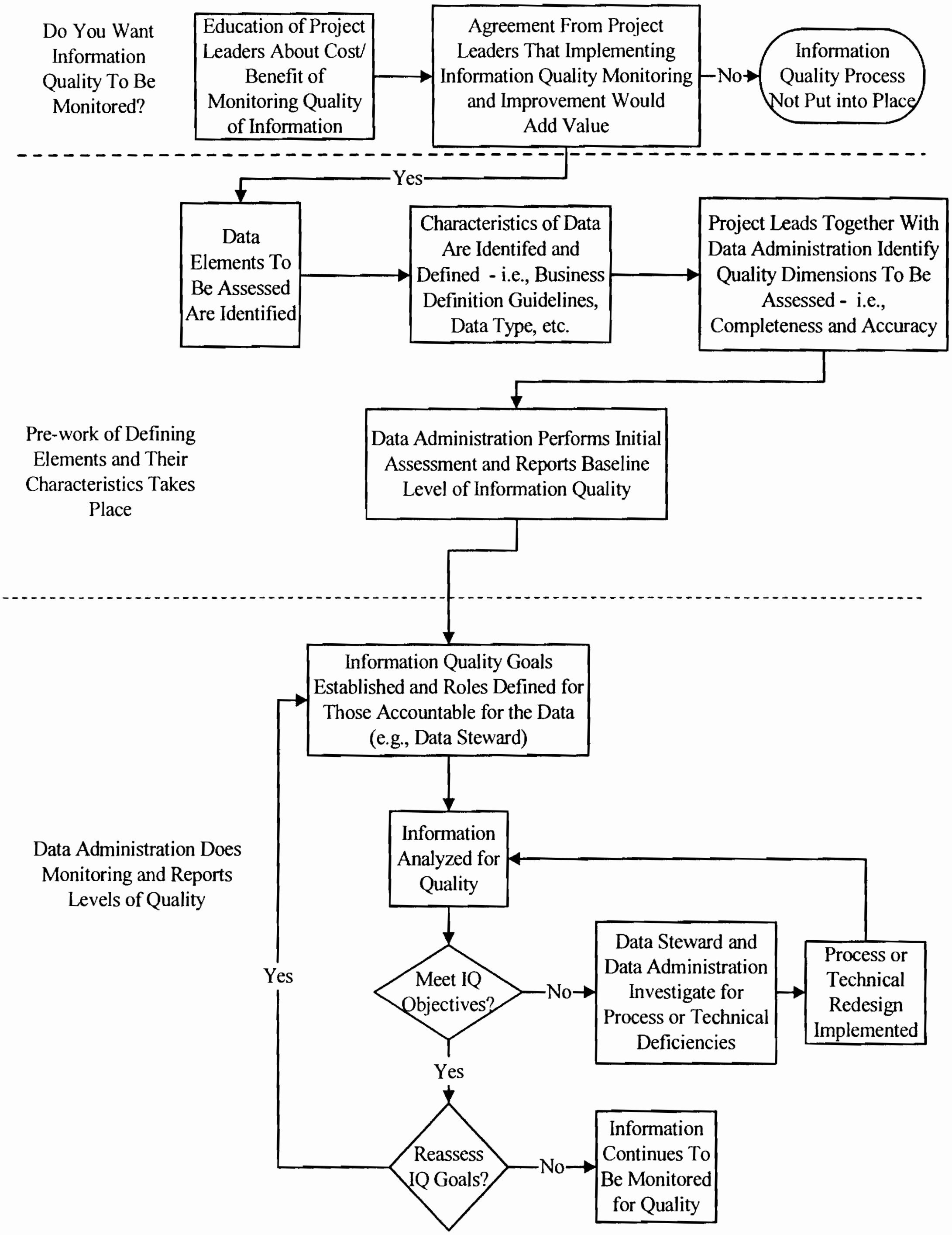


Figure 24. Example of a standard process for ensuring information quality.

data element (name, location, data type, definition, etc.) and its consistency with existing data standards (organizational and national); and the processes for capturing, storing, integrating, and using the data. The objectives of this phase are to gain an understanding of the overall reliability of the organization's information resources and the way in which information is used, to identify specific areas of poor information quality, and to determine what the cost is to the organization and its customers, including clinicians and patients, when information is of poor quality (e.g., inaccurate, incomplete, inconsistent, etc.) or is used improperly. In determining costs, it is important to consider both "obvious" and "nonobvious" costs associated with poor quality data. Obvious costs include issues such as the cost of finding "bad" data and the rework required to clean it up and lost revenue due to inaccurate information being sent to payers who subsequently reject the claims. Nonobvious costs become apparent when organizations consider questions such as the following: How many patients incur physical or emotional harm as a result of poor quality data? How many care providers spend minutes, hours, or days looking for information, correcting information, or correcting actions they took based on incorrect information? How many processes and/or treatments could have been improved if good quality information was available? Each of these questions addresses issues that lead to increased costs in the delivery of healthcare. Assessment of the cost of poor quality information helps measure the value of an information quality program, and thus support the business case for implementing and maintaining such a program.

Prioritization of Assessment Findings

After completion of the assessment, the identified information quality issues are prioritized on the basis of the importance of the affected data to the organization's business strategies and the cost-benefit ratio of having poor quality data versus resolving the information quality issue.

Data Cleansing

Data cleansing is used to address issues related directly to characteristics of the data itself.[31] This includes, but is not limited to, characteristics such as consistency of data element names, validity of data values, and completeness of data in various data fields in databases. Data cleansing is most efficiently and effectively achieved with electronic tools for data extraction, transformation, and loading (from source systems into target systems). These tools can be used to make changes to the information in an electronic system to address issues related to the data characteristics mentioned above. For example, if during the assessment phase, it was learned that a standard measurement term such as "kilograms" was being represented in several different ways (i.e., kg, kgrms, klg) in different source systems, the extraction, transformation, and loading tools could be used to find all fields with the range of names for kilograms and standardize the term across all source systems to be consistent with an organizational standard. Although performing data cleansing using electronic tools will save significant amounts of time when compared to performing the same cleansing process manually, unfortunately, there may be data sources in the organization that are not accessible to electronic tools, and will need to be cleansed manually.

Technical and Process Changes

In addition to the data cleansing process, it is necessary to address the technical system and "information process" issues that lead to poor quality data. Technical issues may be related either to software applications that do not function as they were intended to (i.e., bugs), to applications that become outdated as the intended use of the application changes, or to poorly designed or implemented applications. An example of poor implementation impacting information quality would be a healthcare organization that has multiple sites at which abstracting and encoding of medical records is performed and in which information from the various sites is used to perform aggregated, organization-wide outcome studies. If the organization makes a choice to implement an abstracting and encoding application as a "stand-alone" application with individual databases at each site, the organization creates a situation in which either standard coding edits and data definitions are absent or it becomes difficult to ensure the use of the standards, if present. The result of this type of distributed implementation and databases often is inconsistency between the same data elements in the different databases. The data inconsistencies result in an inability to support integration of the data for the organization-wide outcome studies. Had the organization chosen to implement the system from a central server, with standard edits and a single database to be used by all sites, the chances of creating consistent, high quality data would be significantly increased.

There is a great deal of truth to the expression that "information is only as good as the process that creates it." The processes evaluated in this research, for ordering imaging services procedures, are a real-life example of processes that, by their very nature, present the opportunity for creating poor quality data. The example of the variation in imaging procedure ordering processes is but one of many that can be found in healthcare organizations.

Many decisions about technical system selection and implementation and processes for capturing, storing, and using data have been made in many organizations, based on the immediate needs of the organization. The decisions seem logical at the time; however, the reason the software or processes often do not work at a later time is that the organizations have not considered both the immediate and potential long-term uses of the data. The result of not selecting and implementing technology, or not designing information-related processes with both present and future needs in mind, is having to redesign the processes at a later date, in order to maintain or improve the quality of the information. An information quality program will help identify the "problematic" technology and "defective processes," and guide the assignment of resources to address these issues and, thus, prevent future information quality problems. Achieving this goal will require setting up standard accountability structures, policies, and procedures within departments and at organization-wide levels.

Education

In addition to data cleansing and technical system and process refinement or redesign, education is essential to the success of an information quality program. At a very basic level, it is important for all information producers to understand the importance and uses of the information they produce. Furthermore, it is important for information custodians to understand the processes for producing, storing, and disseminating the information and for information consumers to understand the "history" of the information they are using to make decisions and the "correct" ways in which to use the information. Education for information consumers is particularly important in the areas of data analysis and presentation, including a focus on understanding how to ask the "right" questions of the data, to get the information necessary for the decisions they are making. An organized educational program, focused on increasing knowledge about how to create, store, deliver, and use information is an important component of an IQ program.

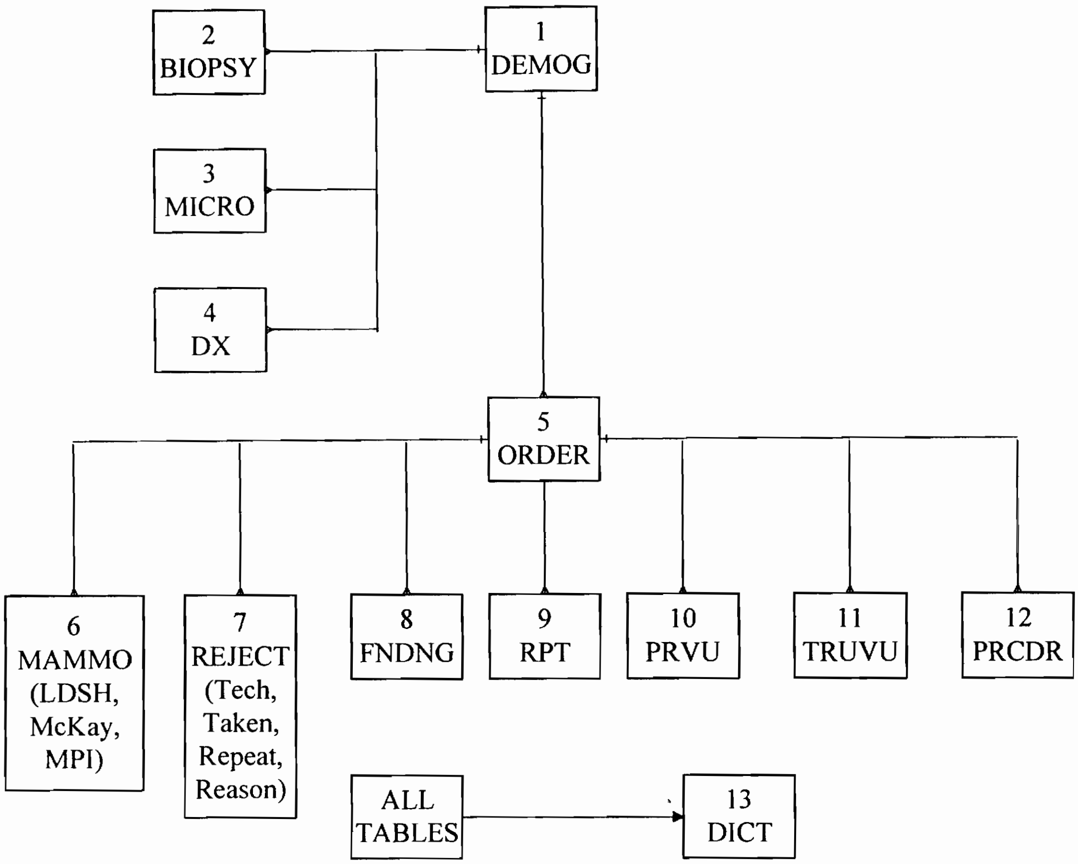
There are those who advocate that the healthcare industry cannot wait for "perfect" data and, thus, need to learn to use imperfect data while we are working on information quality.[100,143] If imperfect data are going to be used, it is essential for users to be aware of the limitations of the information they are using and not to use the imperfect data as an "absolute" judgment system. Furthermore, at the same time imperfect data are being used, efforts must be put forth to manage and improve the quality of the data.

Quality is generally not an accident. It begins with planning and is sustained through management. Just as the quality of clinical processes and outcomes is the responsibility and obligation of every member of a healthcare organization, so is the quality of information. An organized, structured program is required to address all aspects of information quality, including achieving, maintaining, and improving it. Information Quality Programs must involve individuals from all areas and levels of the organization.

If those of us in healthcare value our data and information, we must embrace the information quality process as a part of our daily work. Quality improvement is a continuous process that is achieved by integrating quality management beliefs, principles, and methods into the culture of an organization. Although, for lack of a better term, the word "program" has been used to describe the recommended approach to managing information quality, quality is not a program; it is a mindset, a belief, and a culture. Just as this belief has been adopted in healthcare with respect to clinical quality, it can and should be adopted with respect to information quality. Healthcare organizations need to have specific resources assigned and responsible for information quality, to have organization-wide educational efforts focused on the principles of information quality, and to ensure the adoption of those principles by members of the organization. The people in the healthcare industry are doing a great deal to improve clinical and operational quality, and thus the way in which services are delivered to patients and other consumers of healthcare. I believe we can further our efforts towards providing high quality, efficient healthcare through efforts targeted at information quality improvement.

APPENDIX A

ISDW ENTITY-RELATIONSHIP DIAGRAM



1. Patient Demographics
2. Biopsy Data
3. Microbiology Results Data
4. Discharge Diagnosis Data
5. Radiology Order Data
6. Mammography Data
7. Reject Analysis Data
8. Exam Findings Data
9. Exam Report Data
10. Professional Relative Value Units
11. Technical Relative Value Units
12. Procedure Data
13. Master Dictionary

Figure 25. ISDW entity-relationship diagram.

APPENDIX B

ISDW TABLES AND DATA ELEMENTS

Table 49

Rdlgy_Demog - (Patient Demographics)

Data Element Name	Data Element Definition
PT_ID	Patient account/encounter number
ADMIT_TIME	Date and time of admission
FCLTY_ID	Hospital/facility identification number
GENDER	Gender
BIRTH_DT	Birth date
UNITRECNUM	Unit record/reference number
PT_NM	Patient name
ATTNDNG_DR	Attending physician identification number
SSN	Social security number
MRN	Medical record number
RDLGY_NUM	Radiology record number
DRG	Principal diagnostic related group
ER_STAT	Emergency department status
ADMIT_DX	Admitting diagnosis (ICD-9-CM code)
PT_TYP	Clinical type
PROC_TYP	Billing type
PHON_NUM	Phone number
INSUR1	Primary medical insurance carrier
INSUR2	Secondary medical insurance carrier

Table 50

Rdlgy_Biopsy - (Biopsy Results)

Data Element Name	Data Element Definition
PT_ID	Patient account/encounter number
FCLTY_ID	Hospital/facility identification number
SLIDE_NUM	Slide number
BIOPSY_DT	Data and time of Tandem storage
UNITRECUM	Unit record/reference number
TOPOLOGY	HELP ptxt to "T" code (SNOMED Code)
MORPHOLOGY	HELP ptxt to "M" code (SNOMED Code)
ETIOLOGY	HELP ptxt to "E" code (SNOMED Code)
FUNCTION	HELP ptxt to "F" code (SNOMED Code)
DISEASE	HELP ptxt to "D" code (SNOMED Code)

Table 51

Rdlgy_Micro_Specimen - (Microbiology Specimen Collection)

Data Element Name	Data Element Definition
PT_ID	Patient account/encounter number
FCLTY_ID	Hospital/facility identification number
ACCESSION_ID	Accession identifier
MICRO_DT	Date and time of Tandem storage
SPECIMEN	PTXT code for specimen collected
UNITRECNUM	Unit record/reference number

Table 52

Rdlgy_Micro - (Microbiology Results)

Data Element Name	Data Element Definition
PT_ID	Patient account/encounter number
FCLTY_ID	Hospital/facility identification number
ACCESSION_ID	Accession identifier
TEST	Type of test – isolate, gram stain, etc.
BACTERIA	PTXT code for bacteria cultured
VOLUME	PTXT code for volume of growth

Table 53

Rdlgy_DX - (Discharge Diagnoses from Code-3 System)

Data Element Name	Data Element Definition
PT_ID	Patient account/encounter number
FCLTY_ID	Hospital/facility identification number
PTXT	HELP ptxt code to diagnosis
DX_IDX	Order of diagnosis (admit, primary, secondary)
UNITRECUM	Unit record/reference number
DSCH_DT	Discharge date and time

Table 54

Rdlgy_Order - (Radiology Order Data)

Data Element Name	Data Element Definition
TIME_TBD	Date and time exam to be done
SQNC_NUM	Examination sequence number
FCLTY_ID	Hospital/facility identification number
PT_ID	Patient account/encounter number
PRCDR_CHRGCODE	Primary procedure charge code
RQST_DR	Requesting physician number
PT_RM	Patient room at time of order
ORDER_STAT	Order status
TIME_ORDER	Date and time order was placed
TIME_DONE	Date and time order was completed/confirmed
RESN1	Principal reason for the exam (ICD-9 code)
RESN2	Secondary reason for the exam (ICD-9 code)
ORDER_CLRK	Ordering clerk social security number
CHNG_CLRK	Order change clerk social security number
CONFIRM_TECH1	Principal confirming technician social security number
CONFIRM_TECH2	Secondary confirming technician social security number
CONFIRM_TECH3	Tertiary confirming technician social security number
MODALITY	Modality/department of exam
BODY_REGN	Body region of exam
TRANSPRT	Transport code

Table 55

Rdlgy_Mammo – ('Tickle File' - Volatile Tandem - Mammography Data - LDSH Only)

Data Element Name	Data Element Definition
TIME_TBD	Date and time exam to be done
SQNC_NUM	Examination sequence number
FCLTY_ID	Hospital/facility identification number
EXAM_DX	Exam diagnosis code
DISPO	Disposition code
CALC	Calcification code
OPAC	Opacity code
COMP	Composition code
PREV	Previous exam code

Table 56

Rdlgy_Mammo_Mkay– ('Tickle File' - Volatile Tandem – Mammography Data – McKay Only)

Data Element Name	Data Element Definition
FCLTY_ID	Hospital/facility identification number
ACT_TIME	Date and time of next action
PT_ID	Patient account/encounter number
UNITRECNUM	Unit record/reference number
FOLLOW_UP	Follow up code
DISPOSITION	Disposition code
DISTRIBUTION	Distribution code
CONTACTS	Number of patient contacts/contact attempts
RPT_DOC	Reporting physician number
REFER_DOC	Referring physician number
TIME_TBD	Date and time exam to be done
PRCDR_CHRGCODE	Procedure charge code
RPT_TIME	Transcription date and time
COMP_DATE	Date and time order was completed/confirmed
COMP_PLACE	Place order was completed/confirmed

Table 57

Rdlgy_MPI_Mammo – (MPI File Mammography Data)

Data Element Name	Data Element Definition
TIME_TBD	Date and time exam to be done
SQNC_NUM	Examination sequence number
FCLTY_ID	Hospital/facility identification number
COMP	Composition code
SURG	Surgery code
CALC	Calcification code
OPAC	Opacity code
PREV	Previous exam code
ASSES	Assessment code
AWORK	Additional work code

Table 58

Rdlgy_Rej_Tech - (Technicians Involved in Reject Analysis)

Data Element Name	Data Element Definition
TIME_TBD	Date and time exam to be done
SQNC_NUM	Examination sequence number
FCLTY_ID	Hospital/facility identification number
REJ_TECH	Technician ID (Social security number)

Table 59

Rdlgy_Rej_Tkn - (Size and Number of Films Taken)

Data Element Name	Data Element Definition
TIME_TBD	Date and time exam to be done
SQNC_NUM	Examination sequence number
FCLTY_ID	Hospital/facility identification number
F_SIZE	Size of film
QTY	Number of films taken

Table 60

Rdlgy_Rej_Rept - (Size and Number of Films Repeated)

Data Element Name	Data Element Definition
TIME_TBD	Date and time exam to be done
SQNC_NUM	Examination sequence number
FCLTY_ID	Hospital/facility identification number
F_SIZE	Size of film
QTY	Number of films repeated

Table 61

Rdlgy_Rej_Resn - (Reason for Film Rejection/Repeat)

Data Element Name	Data Element Definition
TIME_TBD	Date and time exam to be done
SQNC_NUM	Examination sequence number
FCLTY_ID	Hospital/facility identification number
RESN	Reason for film rejection/repeat

Table 62

Rdlgy_Fndng - (Exam Findings Data)

Data Element Name	Data Element Definition
TIME_TBD	Date and time exam to be done
SQNC_NUM	Examination sequence number
FCLTY_ID	Hospital/facility identification number
FNDNG_LINK	Link number for internal join
FNDNG_CODE	Finding, audit, etc. ptxt code
FNDNG_VAL	Finding, audit, etc. value

Table 63

Rdlgy_Rpt - (Exam Report Data)

Data Element Name	Data Element Definition
TIME_TBD	Date and time exam to be done
SQNC_NUM	Examination sequence number
FCLTY_ID	Hospital/facility identification number
EVENT_NUM	Report or addenda number record refers to
TIME_DICT	Dictation date and time
TIME_RPT	Transcription date and time
TIME_CORRECT	Correction date and time
TIME_SIGN	Finalization date and time
DICT_DR	Dictating physician number
SIGN_DR	Finalizing physician number
RPT_CLRK	Transcriptionist social security number
RPT_CHAR-CNT	Character count of finalized report

Table 64

Rdlgy_PRVU - (Professional Relative Value Units)

Data Element Name	Data Element Definition
PRCDR_CHRGCODE	Procedure charge code
RVU	Relative value unit value

Table 65

Rdlgy_TRVU - (Technical Relative Value Units)

Data Element Name	Data Element Definition
FCLTY_ID	Hospital/facility identification number
PRCDR_CHRGCODE	Procedure charge code
IRVU	Inpatient relative value unit value
ORVU	Outpatient relative value unit value
ERVU	Emergency room patient relative value unit value

Table 66

Rdlgy_Prcdr - (Primary and Secondary Procedure Data)

Data Element Name	Data Element Definition
TIME_TBD	Date and time exam to be done
SQNC_NUM	Examination sequence number
FCLTY_ID	Hospital/facility identification number
PRCDR_CHRGCODE	Procedure charge code
QTY	Quantity of the procedure ordered

Table 67

Rdlgy_Dict - (Master Dictionary)

Data Element Name	Data Element Definition
FCLTY_ID	Hospital/facility identification number
CONCEPT	Concept classification of definition
CCODE	Alpha-numeric code of definition
NCODE	Numeric code of definition
TXT	Textual definition

APPENDIX C

SCORING INSTRUCTIONS

The data for this study were acquired from the following sources:

- Hand written imaging study requests
- HELP System
- IHC Imaging Services Data Warehouse
- Casemix System

Scoring will be based on comparisons between multiple sets of data elements. The data sheets have a scoring column containing the letters S, D, and M. Please indicate your score by circling the letter that applies to the comparison for that row. The key for the letters is as follows:

- "S" = The "value" of the data elements is the same
- "D" = The "value" of the data elements is different
  - Different for ICD-9-CM codes - Do the numbers in the two columns differ?
  - Different for ICD-9-CM text - This will be defined as a discrepancy between the text descriptions of the medical indications for the exam that could result in any of the following
    - The two text descriptions indicate signs or symptoms that suggest different diagnoses prior to the imaging study.
    - The two text descriptions directly identify different diagnoses prior to the imaging study.
    - The two text descriptions are of signs, symptoms, or diagnoses that would be evaluated by two different exams.

- The two text descriptions are of signs, symptoms, or diagnoses that could result in the findings in the imaging exam being interpreted differently.

❖ "M" = This will be used for comparisons that cannot be made due to missing data.

I would appreciate it if you would complete the data scoring no later than February 23, 1998. When you have completed scoring the data, please contact me and I will arrange to get your data worksheets from you. You can contact me in one of the following manners:

Mobile phone - 541-912-2461

Home phone – 801-466-2928

E-mail – dmemel@peacehealth.org

If you have any questions during the scoring process, please do not hesitate to contact me at the numbers or e-mail address above. Thanks again.

APPENDIX D

PROCESS STEPS RESULTING IN DISCREPANCY BETWEEN CPT-4 PROCEDURE CODES ON FACILITY AND PROFESSIONAL BILLS

The following example demonstrates the way in which CPT-4 procedure codes may end up being discrepant on facility and professional services bills. The example shows the text of the procedure type information present on the Diagnostic Radiology and Nuclear Medicine Request form, on the white sheet, in the HELP System, and on the Hospital and Radiologists' Logs at various points in the imaging services processes beginning with the ordering of an imaging procedure, and ending with the generation of a bill. The inpatient process is used as the subject of the example. The procedure ordered was a two view chest x-ray. The procedure performed was a one view chest x-ray.

- Information at time procedure order completed:

Diagnostic Radiology and Nuclear Medicine Request Form	Chest 2 view
White sheet	Chest 2 view
HELP System	Chest 2 view
Hospital Log	Chest 2 view
Radiologists' Log	No information yet

- Information at time procedure performed:

Diagnostic Radiology and Nuclear Medicine Request Form	Chest 2 view
White sheet (changed on white sheet, but not in HELP System)	Chest 1 view
HELP System	Chest 2 view
Hospital Log	Chest 2 view
Radiologists' Log	No information yet

- Information produced when imaging services billing coordinator reviews information and makes corrections before report is signed:

Diagnostic Radiology and Nuclear Medicine Request Form	Chest 2 view
White sheet	Chest 1 view
HELP System (information in HELP System changed)	Chest 1 view
Hospital Log	Chest 1 view
Radiologists' Log	No information yet

In this case, after the report is signed, the Radiologists' Log will have a procedure type of "Chest 1 view," and there will be no discrepancy between the indication on the professional and facility bills.

- Information produced when imaging services billing coordinator reviews information and makes corrections after report is signed:

Diagnostic Radiology and Nuclear Medicine Request Form	Chest 2 view
White sheet	Chest 1 view
HELP System	Chest 1 view
Hospital Log	Chest 1 view
Radiologists' Log (forgets to change this log)	Chest 2 view

The bill generated for the facility charges will be coded for performance of a one view chest, because this is the procedure type stored in the HELP System. The bill generated for the professional services will be coded for the performance of a two view chest, because the imaging services billing coordinator neglected to change the Radiologists' Log, and the billing clerk at the professional services billing office uses the information on the Radiologists' Log as the source for their billing codes. The result will be a discrepancy between the facility and professional bills and the potential for a fraud and abuse claim.

APPENDIX E

PROCESS STEPS RESULTING IN DISCREPANCY BETWEEN ICD-9-CM DIAGNOSIS CODES ON FACILITY AND PROFESSIONAL BILLS

The following example demonstrates the way in which ICD-9-CM diagnosis codes may end up being discrepant in different data sources. The example shows the text of the clinical indication information based on the physician's handwritten reason on the Diagnostic Radiology and Nuclear Medicine Request form, the ward clerk's free text reason in the HELP System, the HELP System coded reason, and the Hospital and Radiologists' Logs coded reasons at various points in the imaging services processes beginning with the ordering of an imaging procedure, and ending with the generation of a bill. The inpatient process is used as the subject of the example. The procedure ordered was a two-view chest x-ray. The procedure performed was a one-view chest x-ray. The true clinical indication is chest injury.

- Information at time procedure order completed:

Diagnostic Radiology and Nuclear Medicine Request Form	Chest injury
Ward clerk free text (could not read MD writing)	Pneumonia
HELP System coded reason	Pneumonia
Hospital Log	Pneumonia
Radiologists' Log	No information yet

- Information at time procedure performed:

Diagnostic Radiology and Nuclear Medicine Request Form	Chest injury
Ward clerk free text	Pneumonia
HELP System coded reason	Pneumonia
Hospital Log	Pneumonia
Radiologists' Log	No information yet

- Information produced when imaging services billing coordinator reviews information and makes corrections before report is signed:

Diagnostic Radiology and Nuclear Medicine Request Form	Chest injury
Ward clerk free text	Pneumonia
HELP System coded reason	Chest injury
Hospital Log	Chest injury
Radiologists' Log	No information yet

In this case, after the report is signed, the Radiologists' Log will have a clinical indication of chest injury, and there will be no discrepancy between the indication on the professional and facility bills.

- Information produced when imaging services billing coordinator reviews information and makes corrections after report is signed:

Diagnostic Radiology and Nuclear Medicine Request Form	Chest injury
Ward clerk free text	Pneumonia
HELP System coded reason	Chest injury
Hospital Log	Chest injury
Radiologists' Log	Pneumonia

The bill generated for the facility charges will be coded with a clinical indication of chest injury, because this is the indication stored in the HELP System. The bill generated for the professional services will be coded with a clinical indication of pneumonia, because: 1) The Radiologists' Log has already been generated, 2) the imaging services billing coordinator does not change the

Radiologists' Log, and 3) the billing clerk at the professional services billing office uses the information on the Radiologists' Log as the source for their billing codes. The result will be a discrepancy between the facility and professional bills and the potential for a fraud and abuse claim.

APPENDIX F

INPATIENT AGREEMENT AND SYMMETRY STATISTIC RESULTS

Table 68

MD Documented Indication Compared to Ward Clerk
Free Text Indication – Kappa Statistics

	Kappa	LCL	UCL	P
Kappa 1	0.732	0.641	0.822	< 0.0001
Kappa 2	0.568	0.416	0.719	< 0.0001

Table 69

MD Documented Indication Compared to Ward Clerk
Free Text Indication – Bowker's Statistics

	Bowker's	P
Scorer 1 vs 2	3.000	0.392
Scorer 1 vs 3	16.000	< 0.001
Scorer 2 vs 3	18.000	< 0.001

Table 70

MD Documented Indication Compared to White Sheet
ICD-9-CM Text Indication – Kappa Statistics

	Kappa	LCL	UCL	P
Kappa 1	0.403	0.293	0.522	< 0.0001
Kappa 2	0.385	0.257	0.512	< 0.0001

Table 71

MD Documented Indication Compared to White Sheet
ICD-9-CM Text Indication – Bowker's Statistics

	Bowker's	P
Scorer 1 vs 2	5.571	0.134
Scorer 1 vs 3	30.000	< 0.001
Scorer 2 vs 3	23.000	< 0.001

Table 72

MD Documented Indication Compared to ISDW
ICD-9-CM Text Indication – Kappa Statistics

	Kappa	LCL	UCL	P
Kappa 1	0.378	0.287	0.470	< 0.0001
Kappa 2	- 0.032	- 0.174	0.111	0.6688

Table 73

MD Documented Indication Compared to ISDW
ICD-9-CM Text Indication – Bowker's Statistics

	Bowker's	P
Scorer 1 vs 2	23.552	< 0.001
Scorer 1 vs 3	17.000	< 0.001
Scorer 2 vs 3	44.000	< 0.001

Table 74

Ward Clerk Free Text Indication Compared to White Sheet
ICD-9-CM Text Indication – Kappa Statistics

	Kappa	LCL	UCL	P
Kappa 1	0.632	0.542	0.722	< 0.0001
Kappa 2	0.387	0.241	0.533	< 0.0001

Table 75

Ward Clerk Free Text Indication Compared to
White Sheet ICD-9-CM Text Indication
– Bowker's Statistics

	Bowker's	P
Scorer 1 vs 2	2.333	0.506
Scorer 1 vs 3	20.174	< 0.001
Scorer 2 vs 3	17.211	< 0.001

Table 76

Ward Clerk Free Text Indication Compared to ISDW
ICD-9-CM Text Indication – Kappa Statistics

	Kappa	LCL	UCL	P
Kappa 1	0.590	0.499	0.682	< 0.0001
Kappa 2	0.067	- 0.100	0.234	0.2158

Table 77

Ward Clerk Free Text Indication Compared to ISDW
ICD-9-CM Text Indication – Bowker's Statistics

	Bowker's	P
Scorer 1 vs 2	10.333	0.016
Scorer 1 vs 3	19.000	< 0.001
Scorer 2 vs 3	31.000	< 0.001

Table 78

White Sheet ICD-9-CM Code Indication Compared to ISDW
ICD-9-CM Code Indication – Kappa Statistics

	Kappa	LCL	UCL	P
Kappa 1	0.981	0.886	1.000	< 0.0001
Kappa 2	0.968	0.834	1.000	< 0.0001

Table 79

White Sheet ICD-9-CM Code Indication Compared
to ISDW ICD-9-CM Code Indication
– Bowker's Statistics

	Bowker's	P
Scorer 1 vs 2	1.000	0.801
Scorer 1 vs 3	1.000	0.801
Scorer 2 vs 3	-	-

Table 80

White Sheet ICD-9-CM Text Indication Compared to ISDW ICD-9-CM Text Indication – Kappa Statistics

	Kappa	LCL	UCL	P
Kappa 1	0.545	0.453	0.638	< 0.0001
Kappa 2	0.292	0.152	0.431	< 0.0001

Table 81

White Sheet ICD-9-CM Text Indication Compared to ISDW ICD-9-CM Text Indication – Bowker's Statistics

	Bowker's	P
Scorer 1 vs 2	15.000	0.002
Scorer 1 vs 3	20.000	< 0.001
Scorer 2 vs 3	35.000	< 0.001

APPENDIX G

OUTPATIENT AGREEMENT AND SYMMETRY STATISTIC RESULTS

Table 82

Imaging Services Scheduler Documented Indication Compared to White Sheet ICD-9-CM Text Indication – Kappa Statistics

	Kappa	LCL	UCL	P
Kappa 1	0.590	0.433	0.747	< 0.0001
Kappa 2	0.590	0.433	0.747	< 0.0001

Table 83

Imaging Services Scheduler Documented Indication Compared to ISDW ICD-9-CM Text Indication – Kappa Statistics

	Kappa	LCL	UCL	P
Kappa 1	0.673	0.557	0.789	< 0.0001
Kappa 2	0.502	0.330	0.675	< 0.0001

Table 84

Imaging Services Scheduler Documented Indication Compared to ISDW ICD-9-CM Text Indication – Bowker's Statistics

	Bowker's	P
Scorer 1 vs 2	0.667	0.881
Scorer 1 vs 3	12.000	0.007
Scorer 2 vs 3	14.000	0.003

Table 85

White Sheet ICD-9-CM Code Indication Compared to ISDW ICD-9-CM Code Indication – Kappa Statistics

	Kappa	LCL	UCL	P
Kappa 1	1.000	0.874	1.000	< 0.0001
Kappa 2	1.000	0.837	1.000	< 0.0001

Table 86

White Sheet ICD-9-CM Text Indication Compared to ISDW ICD-9-CM Text Indication – Kappa Statistics

	Kappa	LCL	UCL	P
Kappa 1	0.663	0.547	0.779	< 0.0001
Kappa 2	0.440	0.268	0.613	< 0.0001

Table 87

White Sheet ICD-9-CM Text Indication Compared to ISDW ICD-9-CM Text Indication – Bowker's Statistics

	Bowker's	P
Scorer 1 vs 2	0.667	0.881
Scorer 1 vs 3	12.000	0.007
Scorer 2 vs 3	8.333	0.040

Table 88

White Sheet ICD-9-CM Code Indication Compared to Admit Diagnosis ICD-9-CM Code Indication – Kappa Statistics

	Kappa	LCL	UCL	P
Kappa 1	1.000	0.887	1.000	< 0.0001
Kappa 2	1.000	0.820	1.000	< 0.0001

REFERENCES

1. Reep J, Lohman P. Market forces and the environment. In: Ball MJ, Douglas JV, eds. Performance improvement through information management. New York: Springer-Verlag, Inc.; 1999:3-20.

2. Haughom JL. Transforming U.S. health care: the arduous road to value. Top Health Inf Manage. 2000;20(3):1-10.

3. Millenson ML. Demanding medical excellence. Doctors and accountability in the information age. Chicago: The University of Chicago Press, 1997.

4. Drazen EL, Metzger JB. Strategies for integrated health care: emerging practices in information management and cross-continuum care. San Francisco: Jossey-Bass, 1999.

5. Herzlinger RE. Market-driven health care: who wins, who loses in the transformation of america's largest service industry. Reading: Addison-Wesley, 1997.

6. Shortell SM, Gillies RR, Anderson DA, Erickson KM, Mitchell JB. Remaking health care in America. San Francisco: Jossey-Bass, Inc., 1996.

7. Gabler JM. Healthcare industry trends: the five-year outlook. Gartner Group Report; 2000.

8. Davies D, Girzadas J, Resnick L, Weitzman L, Pallet G. The emergence of the e-health consumer. Deloitte Consulting and Deloitte and Touche Report; 1999.

9. Noffsinger R, Chin S. Improving the delivery of care and reducing healthcare costs with the digitization of information. J Healthc Inf Manag. 2000;14(2):23-30.

10. Leape LL, Brennan TA, Laird NM, et al. The nature of adverse events in hospitalized patients. Results of the Harvard medical practice study II. N Engl J Med. 1991;324(6):377-384.

11. Leape LL, Lawthers AG, Brennan TA, Johnson WG. Preventing medical injury. QRB Qual Rev Bull. 1993;19(5):144-149.

12. Leape LL. Error in medicine. JAMA. 1994;272(23):1851-1857.

13. Kohn LT, Corrigan JM, Donaldson MS. To err is human: building a safer health system. Washington, DC: National Academy Press, 2000.

14. Wennberg JE, Gittelsohn A. Small area variations in health care delivery. Science. 1973;182(117):1102-1108.

15. Wennberg JE, Gittlesohn A. Variations in medical care among small areas. Sci Am. 1982;246(4):120-134.

16. Wennberg JE, McPherson K, Caper P. Will payment based on diagnosis-related groups control hospital costs. N Engl J Med. 1984;311(5):295-300.

17. Wennberg JE. Population illness rates do not explain population hospitalization rates. A comment on Mark Blumberg's thesis that morbidity adjusters are needed to interpret small area variations. Med Care. 1987;25(4):354-359.

18. Lee P. The role of public health in population health. Physician Exec. 1997;23(7):25-29.

19. Boren SA, Balas EA. Evidence-based quality measurement. J Ambulatory Care Manage. 1999;22(3):17-23.

20. Muir Gray JAM. Evidence-based healthcare: how to make health policy and management decisions. New York: Churchill Livingstone, 1997.

21. Kasper JF, Mulley AG, Wennberg JE. Developing shared decision-making programs to improve the quality of health care. QRB Qual Rev Bull. 1992;18(6):183-190.

22. Lenert LA, Cher DJ. Use of meta-analytic results to facilitate shared decision making. J Am Med Inform Assoc. 1999;6(5):412-419.

23. Jadad AR, Haynes RB, Hunt D, Browman GP. The internet and evidence-based decision-making: a needed synergy for efficient knowledge management in health care. CMAJ. 2000;162(3):362-365.

24. Hoehn BJ, Perreault L. Care delivery and care management. In: Ball MJ, Douglas JV, eds. Performance improvement through information management. New York: Springer-Verlag, Inc.; 1999:98-109.

25. Luke RD, Shukla RK. Market structure. In: Ball MJ, Douglas JV, eds. Performance improvement through information management. New York: Springer-Verlag, Inc.; 1999:21-42.

26. Coile RC. Challenges for physician executives in the millennium marketplace. Physician Exec. 1999;25(1):8-13.

27. Reep J, Lohman P. Implications and direction. In: Ball MJ, Douglas JV, eds. Performance improvement through information management. New York: Springer-Verlag, Inc.; 1999:173-188.

28. Rose JS. Medicine and the information age. Tampa: American College of Physician Executives, 1998.

29. Kramer RN, Douglas JV. Planning for performance. In: Ball MJ, Douglas JV, eds. Performance improvement through information management. New York: Springer-Verlag, Inc.; 1999:65-74.

30. Luke RD, Shukla RK. Strategy analysis and implications for information systems. In: Ball MJ, Douglas JV, eds. Performance improvement through information management. New York: Springer-Verlag, Inc.; 1999:43-62.

31. English LP. Improving data warehouse and business information quality: methods for reducing costs and increasing profits. New York: John Wiley & Sons, Inc, 1999.

32. Huang KT, Lee YW, Wang RY. Quality information and knowledge. Upper Saddle River: Prentice Hall PTR, 1999.

33. Strong DM, Lee YW, Wang RY. 10 Potholes in the road to information quality. IEEE Computer. 1997;30(8):38-46.

34. Redman TC. Data quality for the information age. Norwood: Artech House, Inc., 1996.

35. Inmon WH, Imhoff C, Sousa R. Corporate information factory. New York: Wiley Computer Publishing, 1998.

36. Braun P, Caper P. Information needs in a changing health care system: capitation and the need for a population-oriented view. J Ambulatory Care Manage. 1999;22(3):1-10.

37. DesHarnais S, Marshall B, Dulski J. Information management in the age of managed competition. Jt Comm J Qual Improv. 1994;20(11):631-638.

38. Ricord LC. Meeting the joint commission's information management planning standard: a primer and example. Top Health Inf Manage. 1995;15(4):44-54.

39. Levinson D. Information management in clinical practice. J Fam Pract. 1978;7(4):799-805.

40. Krouse M. Personal Communication. First Consulting Group. Seattle: 1999.

41. McDonald CJ. Protocol-based computer reminders, the quality of care and the non-perfectibility of man. N Engl J Med. 1976;295(24):1351-1355.

42. Strong DM, Lee YW, Wang RY. Beyond accuracy: how organizations are redefining data quality. Total Data Quality Management Research Program, MIT Sloan School of Management. http://web.mit.edu/tdqm/www/papers/94/94-07.html, 1994

43. English LP. Data quality in the data warehouse: a key critical success factor. Information Impact. 1996;4(1):1-5.

44. Leonard K, Tan J, Pink G. Designing health care information systems for integrated delivery systems: where we are and where we need to be. Top Health Inf Manage. 1998;19(1):19-30.

45. Scott TJ. Creating a new view of information processes. Advance for Health Information Executives. 1999;3(3):71-76.

46. Bysinger B. The key to integrated health care. Advance for Health Information Executives. 1999;3(6):37-40.

47. Glaser JP. Developing a clinical information system: the role of the chief information officer. Jt Comm J Qual Improv. 1994;20(11):614-621.

48. Dalto JD, Jacobson J, Haug PJ. Information management plan for urban central region hospitals. Intermountain Health Care; 1997

49. Mohlenbrock WC. Physicians reestablishing clinical autonomy. Physician Exec. 1998;24(1):26-29.

50. van Amerongen D. The next stage of managed care. Physician Exec. 1999;25(1):28-32.

51. Brailer DJ. Management of knowledge in the modern health care delivery system. Jt Comm J Qual Improv. 1999;25(1):6-19.

52. Sisk J. How are health care organizations using clinical guidelines? Health Aff. 1998;17(5):91-109.

53. Casanova J. Tools for the task: the role of clinical guidelines. Tampa: American College of Physician Executives, 1997.

54. Kindig DA. Managing population health. Physician Exec. 1997;23(7):34-39.

55. Rose E. TQM in health care: mistaken identity? Physician Exec. 1997;23(7):44-46.

56. Darby M. 12 ways to get physician buy-in to practice guidelines. Qual Lett Healthc Lead. 1998;10(3):2-8.

57. Glaser JP. The challenge of integrating clinical information systems in an integrated delivery system. Top Health Inf Manage. 1998;19(1):72-77.

58. Perreault LE, Metzger JB. A pragmatic framework for understanding clinical decision support. J Healthc Inf Manag. 1999;13(2):5-21.

59. Overhage JM, Tierney WM, McDonald CJ. Clinical decision support in ambulatory care: tools, trials, and tribulations. J Healthc Inf Manag. 1999;13(2):67-79.

60. Kuperman GJ, Sittig DF, Shabot MM. Clinical decision support for hospital and critical care. J Healthc Inf Manag. 1999;13(2):81-96.

61. Geisler MA, Will D. Implementing enterprise wide databases: a challenge that can be overcome. Top Health Inf Manage. 1998;19(1):11-18.

62. Henchey P. Maximizing ROI in data warehouses. Advance for Health Information Executives. 1998;2(1):60-70.

63. Benko C, DePriest T, Girzadas J. Data warehousing, the internet and today's health care CIO. Advance for Health Information Executives. 1998;2(11):35-42.

64. Essex D. Life line: consumer informatics has gone beyond patient education on the Web. Healthc Inform. 1999;16(2):119-129.

65. Ferguson T. Healthcare in cyberspace: patients lead a revolution. The Futurist. 1997;November-December:29-33.

66. Tetzlaff L. Consumer informatics in chronic illness. J Am Med Inform Assoc. 1997;4(4):285-300.

67. Nicholson L. The internet and healthcare. Chicago: Health Administration Press, 1999.

68. Goedert J. Preparing for the looming EDI standards mandates. Health Data Manag. 1998;6(9):68-75.

69. Goedert J. Proposed data security rules a wake-up call for CIOs. Health Data Manag. 1998;6(9):12-14.

70. Jones RL, Mancino T. HIPAA issues for health care IT management. Advance for Health Information Executives. 1998;2(9):64-70.

71. Goedert J. Data security: the time to begin is now. Health Data Manag. 1999;7(2):108-118.

72. Olson LA, Peters SG, Stewart JB. Security and confidentiality in an electronic medical record. HIMSS News. 1999;March:9-14.

73. Dorodny VS. Security and confidentiality of health information systems: implications for physicians. Physician Exec. 1998;24(3):32-37.

74. Nelson LD. Data, data everywhere. Crit Care Med. 1997;25(8):1265.

75. Ruffin MG. Standardizing medical data. Physician Exec. 1997;23(7):61-64.

76. Chute CG, Cohn SP, Campbell JR. A framework for comprehensive health terminology systems in the United States: development guidelines, criteria for selection, and public policy implications. J Am Med Inform Assoc. 1998;5(6):503-510.

77. Pringle M, Ward P, Chilvers C. Assessment of the completeness and accuracy of computer medical records in four practices committed to recording data on computer. Br J Gen Pract. 1995;45(399):537-541.

78. Einbinder JS, Rury C, Safran C. Outcomes research using the electronic patient record: Beth Israel Hospital's experience with anticoagulation. In: Gardner R, ed. Proceedings on the Nineteenth Annual Symposium on Computer Applications in Medical Care. New Orleans: Hanley & Belfus, Inc.; 1995:819-823.

79. Chun AT, Davidson BN. Implementing the information quality survey: a case study - baseline vs follow-up results of an IQ survey for the Cedars-Sinai Health System. In: Lee YW, Tayi GK, eds. MIT Conference on Information Quality. Boston: MIT; 1999:266-284.

80. Graham G. Personal Communication. Veteran's Administration Health System. Washington, DC: 2000

81. Hallock T. Personal Communication. St. Mary's Hospital. Milwaukee: 2000.

82. Houtsma JB, Leonard TE. Establishing an operational data quality practice: the data quality for Army Medical Department success experience - a first annual review. In: Lee YW, Tayi GK, eds. MIT Conference on Information Quality. Boston: MIT; 1999:220-239.

83. Kosar D. Developing a framework to manage data quality in healthcare. In: Lee YW, Tayi GK, eds. MIT Conference on Information Quality. Boston: MIT; 1999:58-76.

84. Lumpkin JR, Cohn SP, Blair JS, et al. Uniform data standards for patient medical record information. National Committee on Vital and Health Statistics Report; 2000

85. Brennan PF, Stead WW. Assessing data quality: from concordance, through correctness and completeness, to valid manipulatable representations. J Am Med Inform Assoc. 2000;7(1):106-107.

86. Deming EW. Out of the crisis. Cambridge: Center for Advanced Engineering Study, MIT, 1986.

87. Juran JM. Juran on leadership for quality: an executive handbook. New York: The Free Press, 1989.

88. Juran J, Gryna F. Quality planning and analysis. New York: McGraw-Hill, 1980.

89. Gillem TR. Deming's 14 points and hospital quality: responding to the consumer's demand for the best value health care. J Nurs Qual Assur. 1988;2(3):70-80.

90. Prahalad CK, Krishnan MS. The new meaning of quality in the information age. Harv Bus Rev. 1999;September-October:109-118.

91. Strong DM, Lee YW, Wang RY. Data quality in context. Comm. ACM. 1997;40(5):103-110.

92. Lee YW. Why is "know why" knowledge useful for solving information quality problems? In: Americas Conference on Information Systems, The Association for Information Systems (AIS). Phoenix: The Association for Information Systems; 1996:200-202.

93. Orman L, Storey V, Wang RY. Systems approaches to improving data quality. Total Data Quality Management Research Program, MIT Sloan School of Management. http://web.mit.edu/tdqm/www/papers/94/94-05.html, 1994

94. Wang RY, Strong DM, Guarascio LM. An empirical investigation of data quality dimensions: a data consumer's perspective. Total Data Quality Management Research Program, MIT Sloan School of Management. http://web.mit.edu/tdqm/www/papers/94/94-01.html, 1993

95. Wang RY, Reddy MP, Kon HB. Toward quality data: an attribute-based approach. Decision Support Systems. 1995;13:349-372.

96. Wang RY, Storey V, Firth C. A framework for analysis of data quality research. IEEE Trans Knowledge and Data Eng. 1995;7(4):623-640.

97. Wang RY, Madnick S. Introduction to the TDQM Research Program. Total Data Quality Management Research Program, MIT Sloan School of Management. http://web.mit.edu/tdqm/www/intro.html, 1996

98. Wang RY, Strong DM. Beyond accuracy: what data quality means to data consumers. J Management Information Systems. 1996;12:5-34.

99. Demlo LK, Campbell PM, Brown SS. Reliability of information abstracted from patient's medical records. Med Care. 1978;16(12):995-1005.

100. Fetterolf DE. The use of imperfect data in managed care organizations. Physician Exec. 1998;24(3):42-46.

101. Komaroff AL. The variability and inaccuracy of medical data. Proc IEEE; 1979:1196-1207.

102. DeSimone JP, Hudson N, Konen JC. Does encounter data really reflect what is seen in the office? J Fam Pract. 1976;3(6):666.

103. Koran LM. The reliability of clinical methods, data and judgments (first of two parts). N Engl J Med. 1975;293(14):642-646.

104. Koran LM. The reliability of clinical methods, data, and judgments (second of two parts). N Engl J Med. 1975;293(14):695-701.

105. Romm FJ. The validity of the medical record. Med Care. 1981;19(3):310-315.

106. Parkin DM, Muir CS. Cancer incidence in five continents: comparability and quality of data. IARC Scientific Publications. 1992;120:45-173.

107. Barrie JL, Marsh DR. Quality of data in the Manchester Orthopaedic Database. Br Med J. 1992;304(6820):159-162.

108. Ricketts D, Newey M, Patterson M, Hitchin D, Fowler S. Markers of data quality in computer audit: the Manchester Orthopaedic Database. Ann R Coll Surg Engl. 1993;75(6):393-396.

109. Neaton JD, Duchene AG, Svendsen KH, Wentworth D. An examination of the efficiency of some quality assurance methods commonly employed in clinical trials. Stat Med. 1990;9(1-2):115-123; discussion 124.

110. Pollock BH. Quality assurance for interventions in clinical trials: multicenter data monitoring, data management, and analysis. Cancer. 1994;74(9 Suppl):2647-2652.

111. Vantongelen K, Rotmensz N, van der Schueren E. Quality control of validity of data collected in clinical trials. EORTC Study Group on Data Management (SDGM). European Journal of Cancer & Clinical Oncology. 1989;25(8):1241-1247.

112. Horbar JD, Leahy KA. An assessment of data quality in the Vermont-Oxford Trials Network database. Control Clin Trials. 1995;16(1):51-61.

113. Romano PS, Luft HS. Getting the most out of messy data: problems and approaches for dealing with large administrative data sets. In: Grady ML, ed. Medical effectiveness research data methods. Rockville: Agency for Health Care Policy and Research; 1992:57-75.

114. Turner BJ, Markson LE, Houchens R, Fanning T. Research on outcomes of care using a medicaid AIDS data base: overcoming the curses of imperfect data. In: Grady ML, ed. Medical effectiveness research data methods. Pub No. 92-0056 ed. Rockville: Agency for Health Care Policy and Research; 1992:37-52.

115. Reidel DC, Brown F, Charles G, et al. Institute of Medicine report of a study: reliability national hospital discharge survey data. In. Washington, DC: National Academy of Sciences; 1980.

116. Colin C, Ecochard R, Delahaye F, et al. Data quality in a DRG-based information system. Int J Qual Health Care. 1994;6(3):275-280.

117. Hogan WR, Wagner MM. Accuracy of data in computer-based patient records. J Am Med Inform Assoc. 1997;4(5):342-355.

118. Dambro MR, Weiss BD. Assessing the quality of data entry in a computerized medical records system. J Med Syst. 1988;12(3):181-187.

119. Wyatt J. Acquisition and use of clinical data for audit and research. J Eval Clin Pract. 1995;1(1):15-27.

120. Stein HD, Nadkarni P, Erdos J, Miller PL. Exploring the degree of concordance of coded and textual data in answering clinical queries from a clinical data repository. J Am Med Inform Assoc. 2000;7(1):42-54.

121. Jelovsek F, Hammond W. Formal error rate in a computerized obstetric medical record. Methods Inf Med. 1978;17(3):151-157.

122. Yarnall KS, Michener JL, Broadhead WE, Hammond WE, Tse CK. Computer-prompted diagnostic codes. J Fam Pract. 1995;40(3):257-262.

123. Wagner MM, Hogan WR. The accuracy of medication data in an outpatient electronic medical record. J Am Med Inform Assoc. 1996;3(3):234-244.

124. Jones R, Hedley A. A computer in the diabetic clinic: completeness of data in a clinical information system for diabetes. Practical Diabetes. 1986;3:295-296.

125. Hohnloser JH, Fischer MR, Konig A, Emmerich B. Data quality in computerized patient record: analysis of a haematology biopsy report database. Int J Clin Monit Comput. 1994;11(4):233-240.

126. Edsall DW, Deshane P, Giles C, Dick D, Sloan B, Farrow J. Computerized patient anesthesia records: less time and better quality than manually produced anesthesia records. J Clin Anesth. 1993;5(4):275-283.

127. Bryant G. Aspects of data quality in the new millennium. Top Health Inf Manage. 1998;18(4):81-88.

128. Wilton R, Pennisi AJ. Evaluating the accuracy of transcribed computer-stored immunization data. Pediatrics. 1994;94(6 Pt 1):902-906.

129. Aronsky D, Haug PJ. Assessing the quality of clinical data in a computer-based record for calculating the pneumonia severity index. J Am Med Inform Assoc. 2000;7(1):55-65.

130. McDonald CJ, Hui SL. The analysis of humongous databases: problems and promises. Stat Med. 1991;10(4):511-518.

131. Beckman MO, Ringertz HG. Completeness and validity of a radiologic-epidemiologic database in Sweden. Acad Radiol. 1995;2(8):717-720.

132. Lyons C, Gumpert R. Medical audit data: counting is not enough. BMJ. 1990;300(6739):1563-1566.

133. Steinberg EP, Whittle J, Anderson GF. Impact of claims data research on clinical practice. Int J Technol Assess Health Care. 1990;6(2):282-287.

134. Roos LL, Sharp SM, Cohen MM. Comparing clinical information with claims data: some similarities and differences. J Clin Epidemiol. 1991;44(9):881-888.

135. Johnson AN, Appel GL. DRGs and hospital case records: implications for Medicare case mix accuracy. Inquiry. 1984;21(2):128-134.

136. Bentsen BG. The accuracy of recording patient problems in family practice. J. Med Educ. 1976;51(4):311-316.

137. Doremus HD, Michenzi EM. Data quality. An illustration of its potential impact upon a diagnosis-related group's case mix index and reimbursement. Med Care. 1983;21(10):1001-1011.

138. Aas IH. Quality of hospital data and DRGs. Scand J Soc Med. 1988;16(4):223-226.

139. Cassidy B, Fenton S, Fletcher D, Koch D, Stewart M, Watzlaf V. Practice Brief. J AHIMA. 1998;69(6):suppl 7.

140. Wyatt JC. Clinical data systems. Part I. Data and medical records. Lancet. 1994;344(8936):1543-1547.

141. Hogan WR, Wagner MM. Using belief networks to enhance sharing of medical knowledge between sites with variations in data accuracy. In: Gardner R, ed. Proceedings of the Nineteenth Annual Symposium on Computer Applications in Medical Care. New Orleans: Hanley & Belfus, Inc.; 1995:218-222.

142. Iezzoni LI. Risk adjustment for medical outcomes studies. In: Grady ML, ed. Medical effectiveness research data methods. Pub No. 92-0056 ed. Rockville: Agency for Health Care Policy and Research; 1992:83-97.

143. Fitzmaurice JM. Keynote address: what do we do when the data for medical effectiveness research are not perfect? In: Grady ML, ed. Medical effectiveness research data methods. Pub No. 92-0056 ed. Rockville: Agency for Health Care Policy and Research; 1992:5-7.

144. Gatsonis CA. Risk adjustment for medical outcome studies: are we opening Pandora's Box. In: Grady ML, ed. Medical effectiveness research data

methods. Pub No. 92-0056 ed. Rockville: Agency for Health Care Policy and Research; 1992:99-102.

145. Rice CA, Godkin MA, Catlin RJO. Methodological, technical, and ethical Issues and of a computerized data system. J Fam Pract. 1980;10(6):1061-1067.

146. Newhouse JP. Medical effectiveness research methods: summary and reactions. In: Grady ML, ed. Medical effectiveness research data methods. Pub No. 92-0056 ed. Rockville: Agency for Health Care Policy and Research; 1992:177-181.

147. Fortinsky RH, Gutman JD. A two-phase study of the reliability of computerized morbidity data. J Fam Pract. 1981;13(2):229-235.

148. Witt G. The role of meta data in data quality. Journal of Data Warehousing. 1998;3(4):2-10.

149. Lloyd SS, Rissing JP. Physician and coding errors in patient records. JAMA. 1985;254(10):1330-1336.

150. Weissman C. Can hospital discharge diagnoses be used for intensive care unit administrative and quality management functions? Crit Care Med. 1997;25(8):1320-1323.

151. Hsia DC, Krushat WM, Fagan AB, et al. Accuracy of diagnostic coding for Medicare patients under the prospective payment system. N Engl J Med. 1988;318(6):352-355.

152. Laudon KC. Data quality and due process in large interorganizational record systems. Commun ACM. 1986;29(1):4-11.

153. Chin HL, Krall MA, Lester S. ICD9 and SNOMED: adapting clinical coding systems for use in the computer-based patient record. In: Masys DR, ed. AMIA Annual Fall Symposium. Nashville: Hanley & Belfus, Inc.; 1997:849.

154. Fletcher DM. No fool's gold: guarantee riches from your data mine. Healthc Inform. 1997;14(10):115-118.

155. Musen M. The strained quality of medical data. Methods Inf Med. 1989;28(3):123-125.

156. Bates DW, Leape LL, Cullen DJ, et al. Effect of computerized physician order entry and a team intervention on prevention of serious medication errors. JAMA. 1998;280(15):1311-1316.

157. Raschke RA, Gollihare B, Wunderlich TA, et al. A computer alert system to prevent injury from adverse drug events: development and evaluation in a community teaching hospital. JAMA. 1998;280(15):1317-1320.

158. Bates DW, Cullen DJ, Laird N, et al. Incidence of adverse drug events and potential adverse drug events: implications for prevention. JAMA. 1998;274(1):29-34.

159. Classen DC, Pestotnik SL, Evans RS, Lloyd JF, Burke JP. Adverse drug events in hospitalized patients: excess length of stay, extra costs, and attributable mortality. JAMA. 1997;277(4):301-306.

160. Evans RS, Pestotnik SL, Class DC, Horn SD, Bass SB, Burke JP. Preventing adverse drug events in hospitalized patients. Ann Pharmacother. 1994;28(4):253-257.

161. Friedman CP, Wyatt JC. Evaluation methods in medical informatics. New York: Springer, 1997.

162. House ER. Evaluating with validity. Beverly Hills: Sage, 1980.

163. Kaplan B, Maxwell JA. Qualitative research methods for evaluating computer information systems. In: Anderson JG, Aydin CE, Jay SJ, eds. Evaluating health care information systems: methods and applications. Thousand Oaks: Sage Publications, Inc.; 1994:45-68.

164. Doubilet P, Herman PG. Interpretation of radiographs: effect of clinical history. AJR. 1981;137(5):1055-1058.

165. Rhea JT, Potsaid MS, DeLuca SA. Errors of interpretation as elicited by a quality audit of an emergency radiology facility. Radiology. 1979;132(2):277-280.

166. McNeil BJ, Hanley JA, Funkenstein HH, Wallman J. Paired receiver operating characteristic curves and the effect of history on radiographic interpretation: CT of the head as a case study. Radiology. 1983;149(1):75-77.

167. Renfrew DL, Franken EA, Berbaum KS, Weigelt FH, Abu-Yousef MM. Error in radiology: classification and lessons in 182 cases presented at a problem case conference. Radiology. 1992;183(1):145-150.

168. Schuster DM, Gale ME. The malady of incomplete, inadequate, and inaccurate radiology requisition histories: a computerized treatment. AJR. 1996;167(4):855-859.

169. Markert DJ, Haney PJ, Allman RM. Effect of computerized requisition of radiology examinations on the transmission of clinical information. Acad Radiol. 1997;4(2):154-156.

170. Gardner RM, Pryor TA, Warner H. The HELP hospital information system: update 1998. Int J Med Inf. 1999;54(3):169-182.

171. Evans RS. The HELP system: a review of clinical applications in infectious diseases and antibiotic use. MD Comput. 1991;8(5):282-288.

172. Haug PJ, Farrell M, Frear J, Blatter D, Frederick PR. Developing a radiology data base for quality assurance. J Digit Imaging. 1997;10(3, Suppl 1):103-108.

173. Grandia LD, Pryor TA, Willson DF, et al. Building a computer-based patient record system in an evolving integrated health system. In: Steen EB, ed. First Annual Nicholas E. Davies CPR Recognition Symposium. Washington, DC: Computer-based Patient Record Institute, Inc.; 1995:5-33.

174. Pryor TA. The HELP medical record system. MD Comput. 1988;5(5):22-33.

175. Kuperman GJ, Gardner RM, Pryor TA. HELP: a dynamic hospital information system. New York: Springer-Verlag, 1991.

176. Kuperman GJ, Gardner RM. The HELP system: a snapshot in time. Department of Biophysics, LDS Hospital; 1988

177. Usner D. Personal Communication. Oregon State University, Department of Statistics. Corvallis: 1999

178. Fisher LD, Van Belle G. Categorical data: contingency tables, Section 7.4 The measurement of agreement: kappa. In. Biostatistics: a methodology for the health sciences. New York: John Wiley & Sons, Inc.; 1993:256-259.

179. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-174.

180. Fleiss J. Measuring nominal scale agreement among many raters. Psychol Bull. 1971;76(5):378-382.

181. Light RJ. Measures of response agreement for qualitative data: some generalizations and alternatives. Psychol Bull. 1971;76(5):365-377.

182. Everitt B. The cambridge dictionary of statistics. Cambridge, United Kingdom: Cambridge University Press, 1998.

183. Nightingale F. Notes on a Hospital, 1873.

184. Redman TC. Data as a resource: properties, implications, and prescriptions. Sloan Manage Rev. 1998;40(1):89-101.

185. Redman TC. Opinion: improve data quality for competitive advantage. Sloan Manage Rev. 1995;36(2):99-107.