American Health Information Management Association

INFORMATICS TOOLKIT

for Health Informatics and Information Management Professionals

AHIMA
American Health Information Management Association®

INFORMATICS TOOLKIT

for Health Informatics and Information Management Professionals

ISBN: 978-1-58426-620-4

AHIMA Product No.: ONB201317

AHIMA Staff:

Chelsea Brotherton, Assistant Editor

Anne Zender, Senior Director, Periodicals

For more information about AHIMA Press publications, including updates, visit ahima.org/education/press

American Health Information Management Association
233 N. Michigan Ave., 21st Fl.
Chicago, IL 60601

AHIMA.ORG

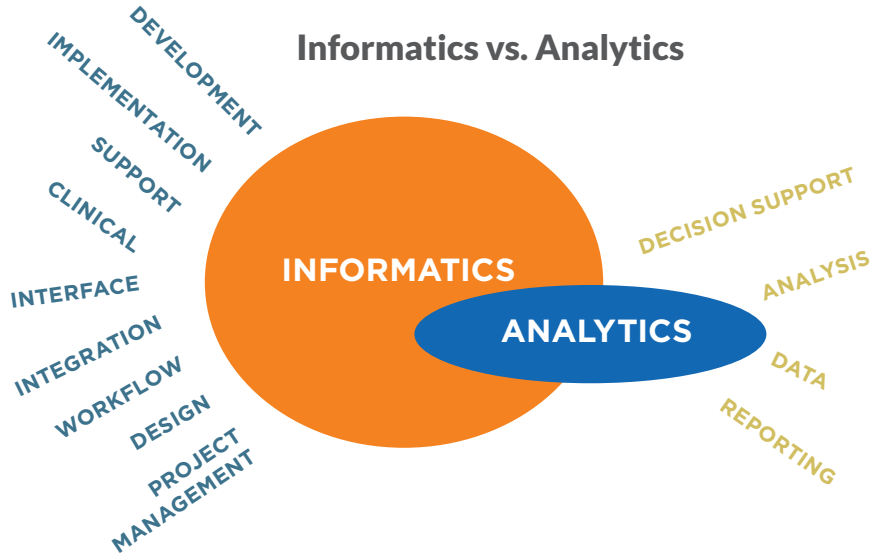
TABLE OF CONTENTS

FOREWORD

This toolkit provides a variety of resources and tools for health information and informatics management (HIIM) professionals who are developing or integrating an informatics initiative, whether they are managing an enterprise master patient index, overseeing or maintaining an electronic health record, initiating an informatics project, or working with other electronic forms of operational, health, or financial information. The toolkit provides tools for HIIM professionals to help them meet the increased demand for informatics in healthcare. It includes an appendix of resources and sample documents, along with suggestions for additional reading.

This toolkit will assist in preparing HIIM professionals who are looking to expand their career, direct their department, or assist in the overall organizational direction toward increasing informatics competencies. Additionally, this toolkit will help prepare other healthcare professionals to come under the wing of HIIM (e.g. pharmacy staff, nurses, etc.)

As health information technology advances and its utilization increases, the available data elements continue to expand. Healthcare has become a technology-rich environment. Various government initiatives have further increased the imperative for healthcare organizations to use their data to make informed business decisions, and a strong informatics program ensures that data is available when and where it is needed and that it is consistently trustworthy and actionable. The goal of this toolkit is to provide a resource that facilitates and enhances the skills of informatics roles in healthcare for health informatics and information management (HIIM) and other health information technology professionals at a foundational level.



This graphic depicts the broader field of informatics versus the more focused field of analytics. It also highlights many of the skills HIIM professionals have acquired in formal education and day-to-day practice.

AUTHORS

Jeannine P. Cain, MSHI, RHIA, CPHI

Shannon H. Houser, RHIA, FAHIMA, PHD, MPH

Lesley Kadlec, MA, RHIA, CHDA

Raymound Mikaelian, RHIA

Barbara Ryznar, RPh

Clarice P. Smith, RHIA, CHP

Annemarie Wendicke, CHDA, MPH

ACKNOWLEDGEMENTS

Kathy Alberson, RHIA

Aurae Beidler, MHA, RHIA, CHPS

Patty Buttner, RHIA, CDIP, CHDA, CPHI, CCS

Jeannine Cain, MSHI, RHIA, CPHI

Lesley Clack, ScD

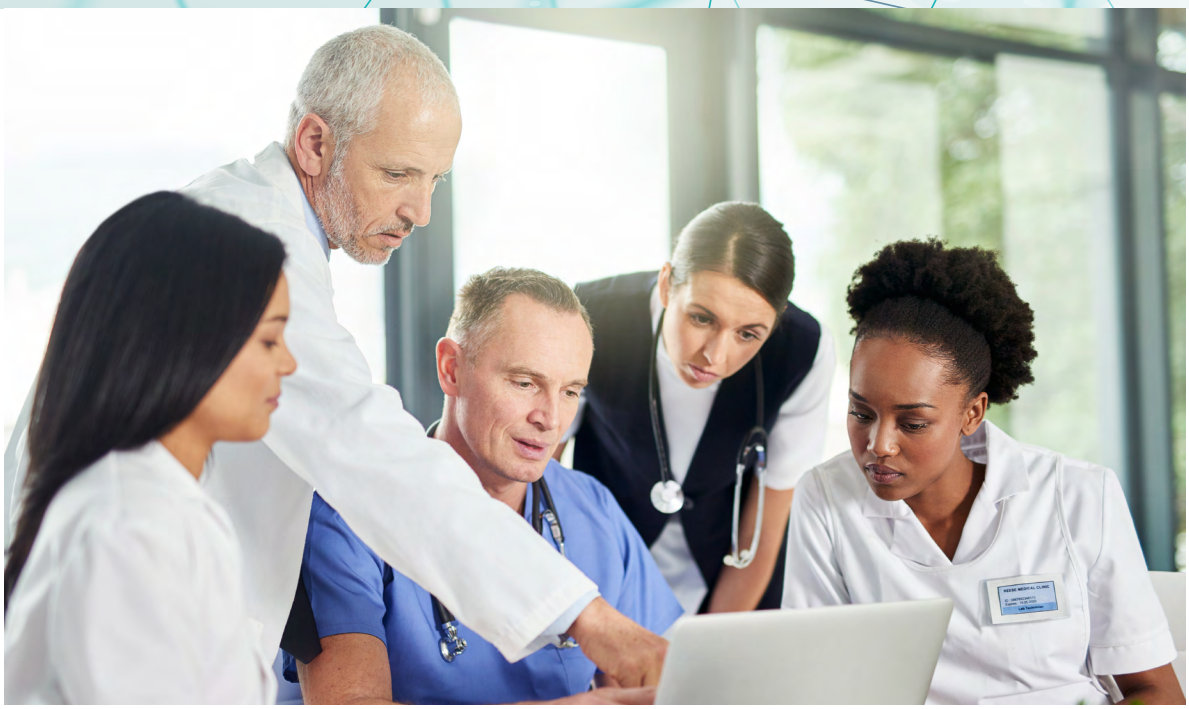
Angie Comfort, RHIA, CDIP, CCS, CCS-P

Melanie Endicott, MBA/HCM, RHIA, CDIP, CHDA, CCS, CCS-P, FAHIMA

Annessa Kirby

Margaret Stackhouse, RHIA

Jami Woebkenberg, MHIM, RHIA, CPHI



INTRODUCTION

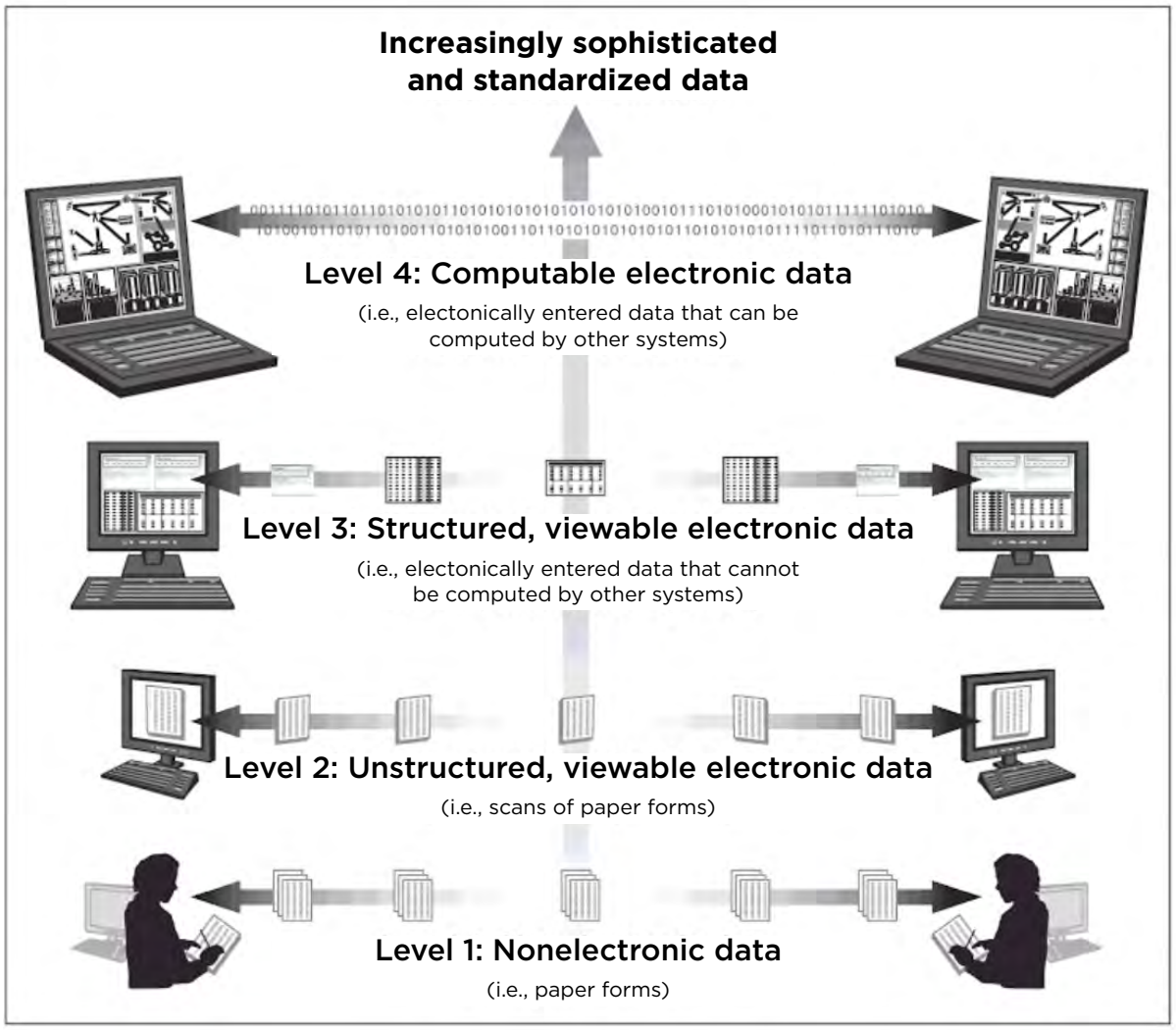
AHIMA recently defined health informatics as involving "a collaborative activity that involves people, processes, and technologies to produce and use trusted data for better decision making."[1] Health informatics professionals are essential through the continued use and evolution of electronic health records (EHRs), because they are the professionals who understand the value of every bit of data. Fundamentally, at this moment in time, computers cannot generate information; they are merely tools in helping healthcare professionals generate, modify, and share data. To a computer, data are merely a sequence of binary characters that can only be processed as such. However, to a health informaticist, data holds the transformative key to the future of healthcare delivery. Thus, the modern informatics professional must work diligently to ensure the governance of data through its lifecycle, to ensure that our information, knowledge, and wisdom have the integrity to positively transform healthcare delivery.

Definitions and Examples[2]

	Definition	Example
Data	Symbols or observations reflecting differences in the world.	11110 (Binary) 30 (Decimal)
Information	Meaningful data or facts from which conclusions can be drawn by humans or computers. Modern computers do not process information, they process data.	BMI = 30 A BMI of 30 indicates that the patient is obese.
Knowledge	Information that is justifiably considered to be true.	A patient with a BMI at, or more than, 30 is more at risk for diabetes and heart-related conditions.
Wisdom	The critical use of knowledge to make intelligent decisions and to work through situations of signal versus noise.	A BMI of 30 may be attributed to a range of factors, including, but not limited to, eating habits, activity level, genetics, and socioeconomic status.

The transformation of data should be viewed as a hierarchy, wherein each step is dependent on the last to support the overall structure. Furthermore, to support a data-rich environment, healthcare's data must reach greater levels of sophistication.

A diagram from the GAO represents a view of the evolution of data.[3]

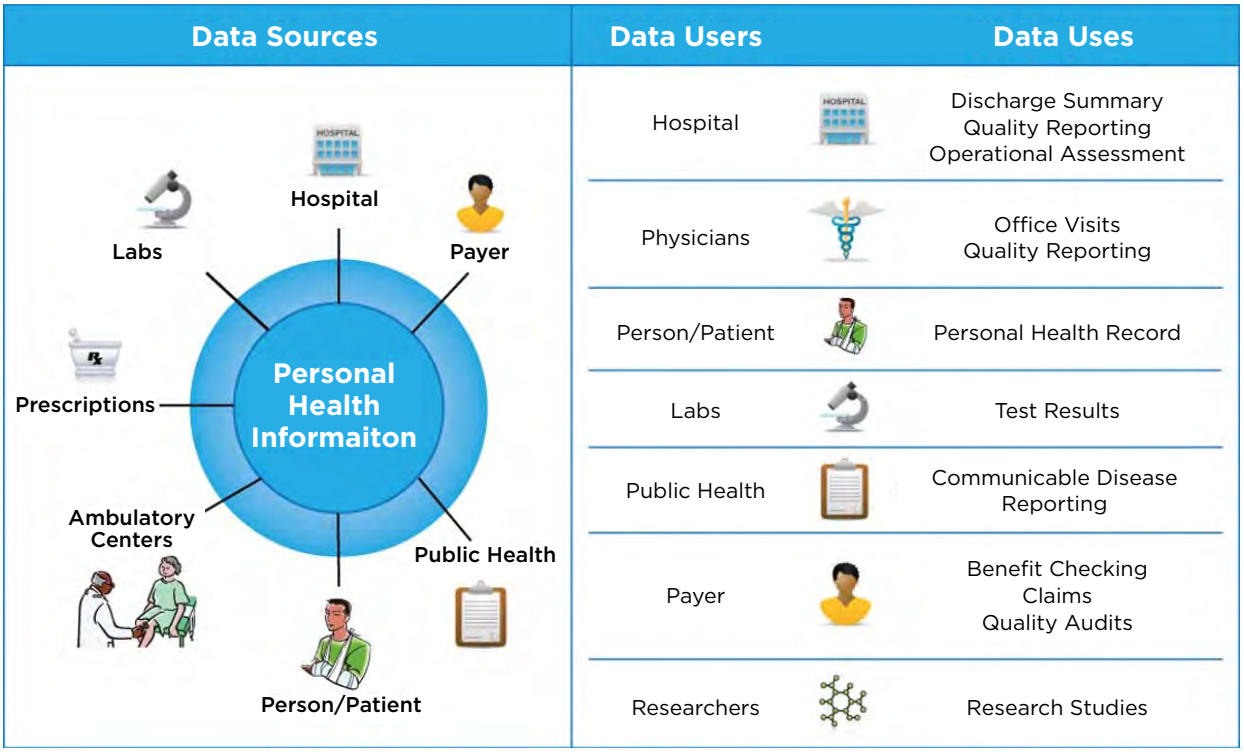


Source: GAO analysis based on data from the Center for Information Technology Leadership.

DATA ANALYSIS AND UTILIZATION

Our healthcare systems create massive amounts of data which, when utilized for decision making, enables improved care by helping clinicians make more informed decisions. As shown in the graphic below, payers may utilize health information to create/alter medical policies for payment/precertification requirements, to promote population health/wellness, and to monitor provider habits to help prevent unnecessary utilization or care that is inappropriate for a patient. It is based primarily on claims data, but it is used for more than just claims payment. Many payers also receive feeds from other sources to use in their analysis.

Data Sources, Users, and Uses



Source: National Committee on Vital and Health Statistics. "Health Data Stewardship: What, Why, Who, How: An NCVHS Primer." September 2009. http://bok.ahima.org/PdfView?oid=94786.

Data analysis is "the process of systematically applying statistical and/or logical techniques to describe and illustrate, condense and recap, and evaluate data."[4]

After the database for the data collection is created and data is obtained as well as entered, the data will need to be cleaned via spot-checking, eyeballing, and logic checks. Before the data can be cleaned, an individual has to be aware of the data, including how it is collected, what values it contains, and what the values mean in order to know how the data needs to be scrubbed or cleaned to yield the proper output.

After that, it needs to be decided on whether the dataset should be imported into a statistical program such as SAS or SPSS for data analysis purposes. Regardless of the software that is being used, the next step would be to perform analysis to find out what exactly the data tells us.

A hypothesis about the data should be developed and tested using a variety of statistical tests, as well as strategies to know how to apply the appropriate one to the specific data and evaluation question. A data analysis to answer evaluation questions should start with descriptive statistics, which include frequencies, descriptors, and crosstabs, or compare means depending on the level of measurement that is included in the dataset. A statistical program can help with exploring concepts such as ANOVA, Chi-Square, and Pearson correlation analysis and get the results very quickly.

The most important tool in the data analysis arsenal is regression analysis, which will allow users to examine multiple predictors of the outcome and be able to identify confounding variables. The quality of the regression model can be evaluated by interpreting regression coefficients and regression diagnostic plots.

The Pell Institute and Pathways to College Network give some examples:[5]

Level of Measurement	Description	Example
Nominal	• Data has no logical order • Data is basic classification data	Male or female
Ordinal	• Data has a logical order, but the differences between values are not constant	T-shirt size (small, medium, large)
Interval	• Data is continuous and has a logical order • Data has standardized differences between values, but no natural zero	Fahrenheit degrees
Ratio	• Data is continuous, ordered • Has standardized differences between values and a natural zero	Height, weight, age, length

After the level of measurement has been identified, a quantitative data analysis can be performed using the following procedures:

- Data tabulation (frequency distributions and percent distributions)
- Descriptive data (mean, median, min/max, mode)
- Crosstabs
- Moderate and advanced analytical methods (correlation, regression, ANOVA)

Limitations:

- Small sample sizes are not appropriate for all procedures.
- Mean can only be calculated from interval and ratio data.
- Median can only be calculated from ordinal, interval, and ratio data.

In contrast to quantitative analysis, qualitative data analysis involves the identification, examination, and interpretation of patterns and themes in textual data and determines how these patterns and themes help answer the research questions at hand.

Steps for a qualitative data analysis:

- Process and record data immediately
- Begin analyzing as data is being collected
- Data reduction
- Identifying meaningful patterns and themes
- Data display
- Conclusion drawing and verification

Appendix F contains sample business use requirements.

DATA REPORTING

Data reporting is the process of identifying data, validating data, building and designing data reports, and submitting reports. There are various reasons for data reporting, such as direct patient care, clinical research, administration, financial or reimbursement, or quality improvement. It is important to compile and generate accurate and valid data, because inaccurate and invalid data reports will provide false information or erroneous evidence for healthcare administrative and clinical decision making or misleading research outcomes. Data reporting includes major tasks such as identifying data sources, conducting data mining, extracting data from identified sources, standardizing data, building data and designing reports, and ensuring standardization of templates and reports.

IDENTIFY APPROPRIATE DATA SOURCES

The first step in data reporting is to identify appropriate data sources. Data selection should be based on the purpose and objectives of the reports or studies. It is essential to know when, where, what, who and how a report or study objectives are defined.

Determine if the data exists, or if data collection should be conducted from scratch.

RECOMMENDED STEPS FOR IDENTIFYING APPROPRIATE DATA SOURCES

- Data selection should be based on the purpose/objectives of the reports.
  - Know when, where, what, who, and how report objectives are defined.
- Does the data exist, or should data collection be conducted from scratch?
  - Use existing data from current research literature, national, state, local data sources:
    - Search www.pubmed.gov or www.scholar.google.com
    - Public use data, which is available at national, state, and local level
    - Internal agency/organization existing data for financial analysis, administration data for quality improvement, clinical patient care, etc.
    - Benchmark data for quality improvement, which will allow the comparison of organizational data to other organizations.
    - Verify the data source to check reliability, validity, credibility of the data and publication sources.
    - Understand existing data sources, including restrictions, definition of the variables, reporting period, reporting frequency, and terms of usage of the existing data source.
- Collect new data and determine:
  - Population of interest
  - Time frame
  - Methods/tools use for data collection
  - Variables, parameters, analysis
  - Report type
  - Always check each step of data collection to match the purpose/objectives.

A good primer by NEDARC can be found at http://www.nedarc.org/tutorials/collectingdata/identifydatasources/nationaldataresources.html.

DATA EXTRACTION AND DATA MINING

Data extraction is the process of retrieving and importing largely unstructured or different data sources for further data processing. Unstructured data, such as narrative progress notes, audio voice dictations, and text messages account for the majority of healthcare data. A data warehouse requires three different methods and processes for utilizing the incoming data, such as extraction, transformation, and loading.

Data mining, also called data discovery, is the process of converting data to information and knowledge by using combined skills from computer science, statistics, and machine learning.

Major elements in data mining include:[6]

- Extracting, transforming, and loading transaction data into the data warehouse system
- Storing and managing the data in a multidimensional database system
- Providing data access to business analysts and information technology professionals
- Analyzing the data by application software
- Presenting the data in a useful format, such as a graph or table

Some key techniques for data mining include:[7]

- Identifying association or relationships between data items
- Classifying data into groups
- Clustering data based on logical relationships
- Using sequential patterns to anticipate trends
- Building decision trees to support data selection and the decision-making process

DIFFERENT LEVELS OF ANALYSIS

Analytics encompass a broad scope of tools, which vary in complexity and purpose. The list below provides insight into various forms of analytics, from most to least complex. As healthcare begins to shift its analytical focus towards big data, tools such as artificial intelligence and neural networks will have greater prominence. Analytics at this level provide a greater understanding from exponentially larger data sets.

- Artificial neural networks: Non-linear predictive models that learn through training and resemble biological neural networks in structure
- Genetic algorithms: Optimization techniques that use processes such as genetic combination, mutation, and natural selection in a design based on the concepts of natural evolution
- Decision trees: Tree-shaped structures that represent sets of decisions. These decisions generate rules for the classification of a dataset. Specific decision tree methods include Classification and Regression Trees (CART) and Chi Square Automatic Interaction Detection (CHAID). CART and CHAID are decision tree techniques used for classification of a data set. They provide a set of rules that you can apply to a new (unclassified) dataset to predict which records will have a given outcome. CART segments a dataset by creating two-way splits while CHAID segments using chi square tests to create multi-way splits. CART typically requires less data preparation than CHAID.
- Nearest neighbor method: A technique that classifies each record in a dataset based on a combination of the classes of the k record(s) most similar to it in a historical dataset (where k=1); sometimes called the k-nearest neighbor technique.
- Rule induction: The extraction of useful if-then rules from data based on statistical significance
- Data visualization: The visual interpretation of complex relationships in multidimensional data. Graphics tools are used to illustrate data relationships.

Key techniques:

- Building association or relation
- Classification
- Clustering
- Prediction
- Sequential patterns
- Decision trees
- Combinations
- Long-term (memory) processing

STANDARDIZING DATA

Data standardization is the process of formatting data to a consistent standard to ensure high-quality data. Healthcare data identified and retrieved from different sources are often varied across organizations, with different formats, values, definitions, computer systems, or software. It is important to standardize data (using standard vocabulary or building common data model) to enable high-data quality in reports and reduce bad consequences and negative effects from poor quality of data.

Data standardization can be achieved by:[8]

- Understanding and cleaning data
- Knowing the data entry points
- Choosing data standards
- Defining the normalization matrix

BUILDING DATA AND DESIGNING REPORTS

Building data and designing reports are based on the purpose and objectives to be communicated and the audience that uses the report. Presentation design is critical in making a clear and meaningful report of the data and ensuring standardization of templates and reports.

Things to consider in report design:

- Select the reporting tool that you will utilize to design the report.
- Understand the report purpose, functions, layout, and interaction with report design surface, working with report parameters, expression, and naming conventions.
- Use queries variables and filter function to allow users to select data for the report
- Use graph, map and tabular data to enhance visualization of the report.
- Use visualization tools appropriate to the data types for easy and effective communication of data results (see figure below).
- Use different background color to differentiate multiple data items.
- Build complex layout in stages.
- Ensure standardization of templates and report.

DATA QUALITY MANAGEMENT, PRIVACY, AND SECURITY

DATA QUALITY MANAGEMENT

To garner the greatest potential from data used by an organization, data quality must be managed and maintained. AHIMA defines data quality management (DQM) as "the business processes that ensure the integrity of an organization's data during collection, application (including aggregation), warehousing, and analysis."[10] Furthermore, while it may be assumed that an information governance (IG) program is dependent on a DQM program, both are dependent on each other. A DQM program cannot effectively be maintained without the support of an IG program, and an IG program will not be sustainable without proper DQM protocols in place.

The concept of data quality management is comprehensively described within AHIMA's "Data Quality Management Model" Practice Brief. An excerpt from this practice brief can be found below.

OVERVIEW OF THE DATA QUALITY MODEL[11]

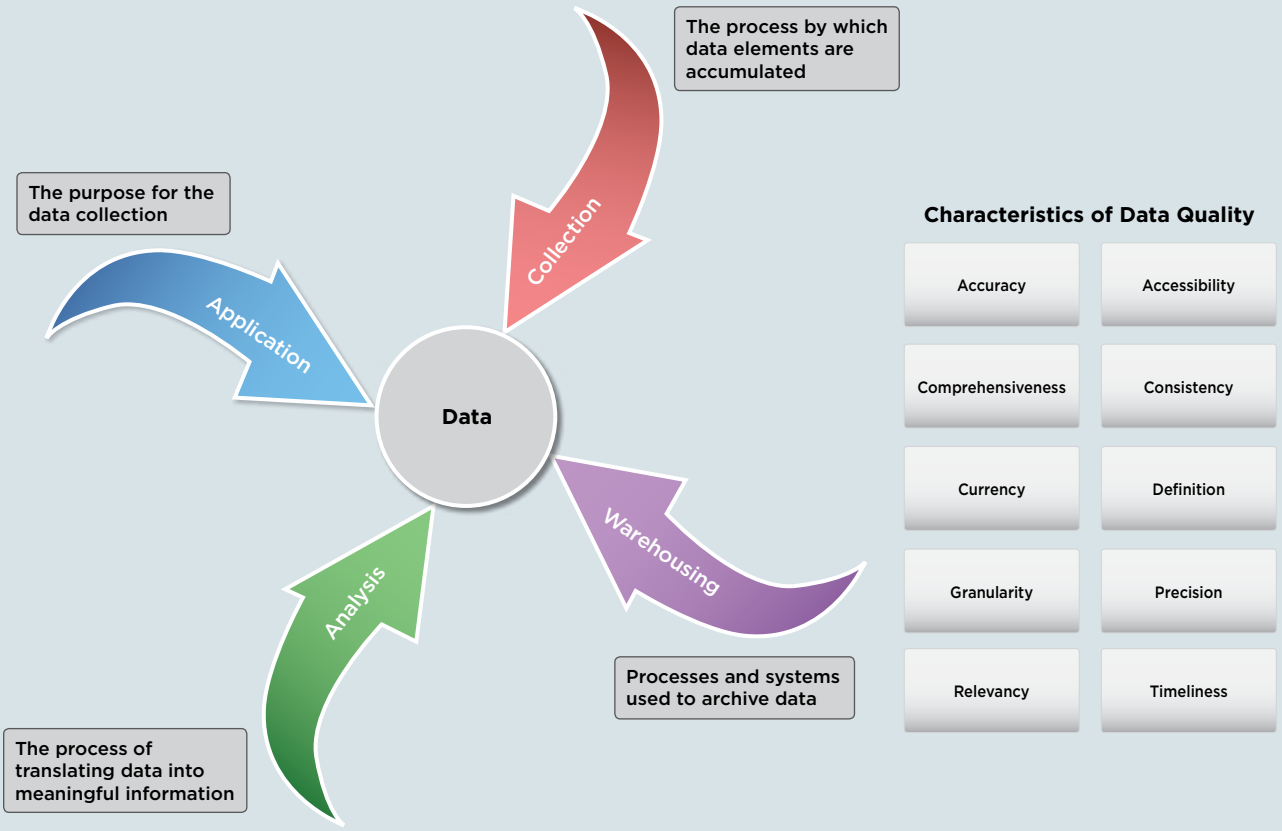
The DQM model was originally developed to illustrate the different data quality challenges that healthcare professionals face. Similar to AHIMA's IGPHC™, this model is generic and adaptable to any care setting and for any application. The tool expands beyond the EHR to include data quality across the healthcare continuum. It is a tool or a model for all healthcare professionals to assist in the transition to enterprise-wide DQM roles. The tool can be applied to all data in the organization, and expands to encompass both clinical and non-clinical areas.

As demonstrated in the table below, data quality management functions include:

- **Application:** The purpose of the data collection
- **Collection:** The processes by which data elements are accumulated
- **Warehousing:** Processes and systems used to archive data
- **Analysis:** The process of translating data into meaningful information

Characteristics of data quality include:

- **Data Accuracy:** The extent to which the data are free of identifiable errors
- **Data Accessibility:** The level of ease and efficiency at which data are legally obtainable, within a well protected and controlled environment
- **Data Comprehensiveness:** The extent to which all required data within the entire scope are collected, documenting intended exclusions
- **Data Consistency:** The extent to which the healthcare data are reliable, identical, and reproducible by different users across applications
- **Data Currency:** The extent to which data are up-to-date; a datum value is up-to-date if it is current for a specific point in time, and it is outdated if it was current at a preceding time but incorrect at a later time
- **Data Definition:** The specific meaning of a healthcare-related data element
- **Data Granularity:** The level of detail at which the attributes and characteristics of data quality in healthcare data are defined
- **Data Precision:** The degree to which measures support their purpose and/or the closeness of two or more measures to each other
- **Data Relevancy:** The extent to which healthcare-related data are useful for the purposes for which they were collected
- **Data Timeliness:** The availability of up-to-date data within the useful, operative, or indicated time



Source: AHIMA. *Pocket Glossary of Health Information Management and Technology,* Third Edition. Chicago, IL: AHIMA Press, 2012.

PRIVACY AND SECURITY

Ensuring the privacy and security of a patient's health information throughout the lifecycle of data is paramount. Privacy and security are not only threatened by ever-increasing cybersecurity threats, they are impacted by day-to-day actions of a covered entity's workforce. This section will review some content as it relates to the Health Insurance Portability and Accountability Act (HIPAA) and the Health Information Technology for Economic and Clinical Health (HITECH) Act. However, there are numerous other federal and state regulations that may impact a covered entity with considerations to preemption. Furthermore, special provisions must be made for health information that is considered to be especially sensitive as defined by the federal and state laws, such as substance abuse, mental health, HIV, and health information of minors. An organization's privacy officer, security officer, and/or compliance officer should be a resource for additional information and/or education regarding the privacy and security regulations impacting an organization.

- **Privacy:** "As it relates to one's health information is the right of a patient to control disclosure of that information."[12]
- **Security:** "As it relates to one's health information, refers to measures that control access and protect information from unauthorized, disclosure, alteration, destruction, or loss while maintain availability for those who need it."[13]

From an informatics perspective, it should be understood that protected health information (PHI) or electronic PHI (ePHI) that's created, maintained, or transmitted within an EMR/EHR may have downstream privacy and security implications (HIPAA's 18 PHI identifiers are provided in Appendix H). Accordingly, special attention must be given when creating new content that will be stored within an EMR/EHR.

Examples:

- Consider the creation of a lab summary sheet, where an HIV lab result, STI, or pregnancy test results automatically. This lab result may be disclosed inappropriately, which will result in a breach of the patient's privacy.
- Consider deficiency letters that may be sent to providers outside of an organizations local area network or enterprise e-mail domain. If deficiency letters contain elements of PHI, they will need to be sent as encrypted emails to protect the inappropriate disclosure of ePHI.

Furthermore, as defined by HIPAA's Security Rule, considerations must be placed on administrative safeguards, physical safeguards, technical safeguards, and organizational policies and procedures and documentation requirements. These are further discussed in Appendix I.

Finally, AHIMA provides an analysis of the 2013 modifications made to HITECH.[14] From a technical perspective, HITECH created additional requirements and restrictions that support the privacy and security of a patient's health information.

Examples:

- Patients can now limit the disclosure of information to health plans, if the healthcare services were paid out of pocket. From a technical standpoint, does your organization have safeguards in place to identify when a patient pays for services rendered out of pocket?
  - » Also, with this in mind, are there any data feeds that may automatically send subscriber data back to the health plan? From an information governance perspective, it would be prudent to inventory outbound data feeds and ensure proper safeguards are in place to prevent inappropriate disclosures of health information.
- HITECH requires that covered entities provide an accounting of disclosures for up to six years prior to the date of the request.[15] With this in mind, a covered entity must ensure that its technical systems are capturing the necessary information as defined by HITECH.

MANAGEMENT OF HEALTH INFORMATION SYSTEMS AND PROCESSES

The implementation and management of health information systems and processes are ongoing in the healthcare environment. As an HIM professional, enjoy the challenge and the opportunity the project brings to you, the department, and the organization. The HIM professional should be involved in the selection and implementation of any health information system that creates clinical data that becomes part of the legal health record or the designated record set. The clinical and information services departments need the direction and leadership of the HIM professional to make informed decisions on what constitutes a legal document which will meet the needs of the user as well as the myriad state and federal regulations. The HIM professional must be proactive in participating in the process and should help drive the project where appropriate.

IDENTIFY WORK PROCESSES

When an upgrade to a current system or the implementation of a new system is being planned, all work processes being impacted should be evaluated. A detailed "current state" analysis should be performed to assess how the process works in the current environment. Once this is complete, a detailed future state analysis should be performed to determine how the system will work once the change is implemented. Once these two analyses have been performed, complete a gap analysis of the change between the two processes.

Careful consideration should be given to process improvement. How can the system improve the current process? When converting from one system to another, users may want to re-create processes they are familiar with in the new system. This is particularly true of paper forms. Challenge the users to think of the possibilities.

Ask questions such as:

- What would be the ideal workflow for this process?
- Why is this process being performed?
- Does it add value?
  - If not, why is it being performed?
- Is there a better way of managing this process?
- Can manual processes be replaced with electronic processes?
- What is the impact to other departments, areas, users, etc.?
  - For example, an organization's decision to eliminate or reduce traditional dictation and transcription services to reduce cost may impact the physician's productivity. The physician must use a template for documentation or a voice to text product that requires front-end editing and is more time consuming for the physician, possibly reducing the physician's productivity and impacting revenue.
- Does the upgrade reduce functionality, response time or other attributes of the system?
  - If so, are there steps to mitigate the impact?
  - For example, when systems are moved to a web-based service, response time can diminish.

The following table can be used to perform a gap analysis. An Excel spreadsheet could also be used to perform the same function. There are various options online for tools to perform a gap analysis.

Sample:

Process			
Current State Process	Future State Process	Gap	Risk

Gap analysis is defined as:

> *A technique that businesses use to determine what steps need to be taken in order to move from its current state to its desired, future state. Also called need-gap analysis, needs analysis, and needs assessment. Gap analysis consists of (1) listing of characteristic factors (such as attributes, competencies, performance levels) of the present situation ("what is"), (2) listing factors needed to achieve future objectives ("what should be"), and then (3) highlighting the gaps that exist and need to be filled. Gap analysis forces a company to reflect on who it is and ask who they want to be in the future.*[16]

MANAGEMENT OF PROCESSES

An important step in implementing any system or making process changes is how to measure success and manage the change. The process change should include a method of evaluating the process and quality checking staff who perform the function, particularly during the implementation stage.

One item that is important to monitor for is the "exceptions" or items that do not fit within a well-defined process so these can be corrected and that data can be used instead of discarded because it does not fall within a particular bucket.

Daily operations management in the HIM department is best handled through management by numbers. The leadership team in the department should review key performance indicators (KPIs) on a routine basis such as volume reports of admissions, discharges, outpatient procedures, ancillary volumes, volume of transcription, unbilled accounts by type, number of charts to code, duplicate numbers created and by whom, chart deficiencies, outstanding release of information requests, revenue, unanswered queries, days in accounts receivable, etc. The type of KPIs will depend on the type of organization in which you work. With the review of KPIs, the leadership team can determine staffing needs and make adjustments as needed to manage the workload.

IMPLEMENT WORK PROCESSES

Implementing work processes requires good planning. The steps are varied, depending on the process, from a very detailed project plan for implementing a system to a simple plan for implementing a minor process change. See the "Project Management" section of this toolkit.

DOCUMENT CLINICAL WORKFLOW

Workflows for all changes should be created before proceeding to build. There are numerous products on the market for the development of workflow diagrams, some of which are free. (As always, use caution when downloading free products from the Internet.)

A workflow diagram is defined as graphical depiction of steps taken, time spent, distance traveled, and other aspects of the way a particular piece of work is done.[17]

EXAMINE WORKFLOW COMPATIBILITY WITH THE EHR SYSTEM

The implementation of an EHR system, whether new or a replacement, does create compatibility issues at times. Depending on the work process, the programming of the system may be modified to adjust to the workflow. (Or the workflow must be modified to adjust to the system programming, which is often the case.) Last but not least and the least preferable of the solutions is that a "workaround" must be developed. Workarounds typically involve manual processes that compensate for the lack of system functionality.

WORK WITH THE ORGANIZATIONAL LEADERSHIP TO IMPROVE INFORMATION PROCESSES

When systems are implemented or processes are changed that impact users beyond the HIM department, it is important to obtain support and buy-in from organizational leaders as well as key stakeholders, particularly leaders whose users will be impacted by the change. In the patient care environment, support should be obtained from the medical staff leadership. An executive sponsor must be on board with change in processes from the very beginning.

An executive sponsor should be assigned to the project at inception.

A champion should be selected early in the process.[18] In Six Sigma, a project champion is an individual who has the authority to use resources within or outside an organization for completion of a given project. A project champion is chosen by the management so as to ensure supervision of a specific project from its initiation to its execution.

BUILD INTERFACES WITH OTHER APPLICATIONS INTERNALLY AND EXTERNALLY

Interfaces share required data from one system to another. There can be inbound messages that go into a system and outbound message that go out of the system to another system. Interfaces are a necessity for all systems with the rare exception of standalone systems that serve one function typically and do not warrant the expense of an interface. Interfaces can receive information from other systems both internal and external to the organization, using standardized computer language (such as HL7) and can send the information outbound as well.

Always consider safeguards in the process for when something changes, i.e., an update that may alter current values. You also have to monitor for failures and distinctions from the source data. Metadata timestamps and values may vary in the original source.

DATABASE MANAGEMENT

DATABASE ADMINISTRATORS

According to the Bureau of Labor Statistics, "Database administrators (DBAs) use specialized software to store and organize data, such as financial information and customer shipping records. They make sure that data are available to users and are secure from unauthorized access."[19] DBAs should be the first point of contact for issues related to the content listed below. Furthermore, while larger healthcare systems may have a dedicated database administrator; smaller organizations or systems may find that their DBA is an IT professional who has several responsibilities.

ANALYZING DATABASES TO ENSURE ACCURACY AND COMPLETENESS

Accurate and reliable data is essential to the multitude of healthcare initiatives currently under way. (For a more in depth discussion of the characteristics of data quality, see the AHIMA Data Quality Management Model). A data dictionary is a tool that healthcare organizations often utilize to assist in ensuring data accuracy. According to the International Organization for Standardization:

> The increased use of data processing and electronic data interchange heavily relies on accurate, reliable, controllable, and verifiable data. One of the prerequisites for a correct and proper use and interpretation of data is that both users and owners of data have a common understanding of the meaning and descriptive characteristics (e.g., representation) of that data. To guarantee this shared view, a number of basic attributes have to be defined.[20]

Healthcare data may be maintained in many different databases and as a result may have inconsistent quality. Naming conventions that are not the same throughout each database, inconsistent definitions, varying field lengths for the same data element, and variation in the element values all can lead to problems or errors, including reduced data quality and misuse of data when writing reports. The AHIMA Health Data Analysis Toolkit lists the following examples of inconsistencies in data that may be seen throughout a healthcare organization:[21]

1. Inconsistent naming conventions
   - The date of the patient's admission is referred to as the *date of admission* in the patient management (PM) system; *admit date* in the fetal monitoring system, and *admission date* in the cardiology database.
   - The unique patient identifier is referred to as a *medical record number* in the PM system, *patient record identifier* in the operating room system, and *"A" number* (a moniker leftover from a legacy system from 25 years ago) and *enterprise master patient identifier* in the catheterization laboratory system.
2. Inconsistent definitions
   - Admission, discharge, transfer (ADT) system: *date of admission* is the date on which an inpatient or day surgery case admission occurs; in the trauma registry system, *date of admission* is the date on which the trauma patient enters the operating room.
   - The *pediatric age* is defined as age less than or equal to 13 in the PM system, whereas the pediatric disease registry defines a *pediatric age* as less than the age of 18.
   - In the bed board system, a *nursing unit* may be defined as 5W or 5 West. Within the scheduling system, *unique locations* are defined as *short procedure unit* or *SPU,* such as *X-ray* or *radiology,* for example.
3. Varying field length for same data element
   - The field length for a patient's last name is 26 in the PM system, whereas the field length for a patient's last name is 15 in the cancer registry system.
   - The medical record number in the PM system is 16 characters long, whereas the cancer registry system maintains a length of 13 characters for the medical record number.
4. Varied element values: In one database the value could be varchar (variable character field) and in another database it could be an integer.
   - The patient's sex is captured as *M, F,* or *U* in the ADT system, whereas the patient's sex is captured as *Male, Female,* or *Other* in the peripheral vascular laboratory database.

DEFINING HEALTHCARE-RELATED DATA STRUCTURES

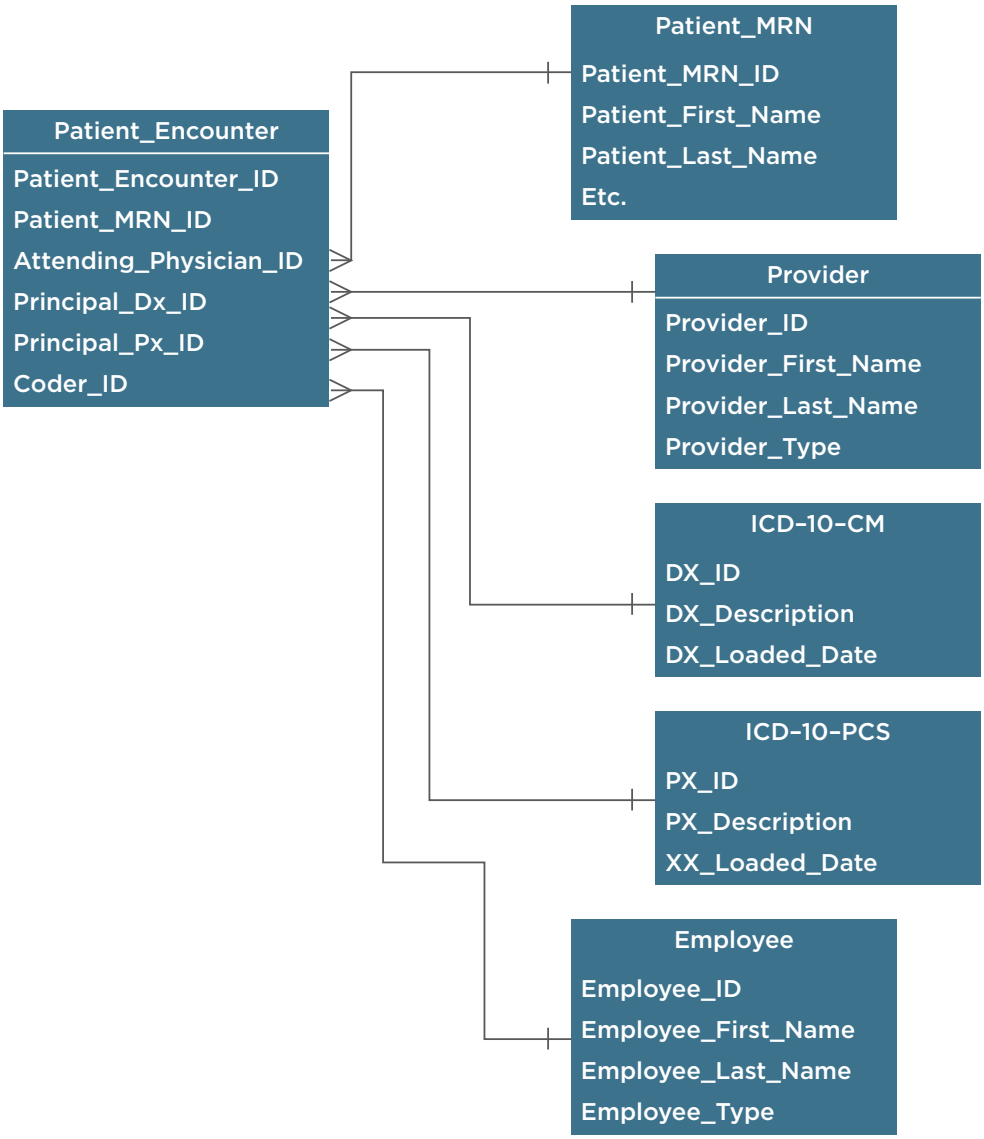
When considering healthcare-related data structures, there are several important concepts to learn about.

- Data structure: "A framework for organizing, storing, and managing data. Data structures consist of files or tables that interact in various ways"[22]

Relational database: "In a relational database, data is organized in a structure of related tables. A table consists of rows and columns, much like a spreadsheet. Each row is simply a record of data on some person (e.g., a patient, a physician), place (e.g., an operating room, the emergency department), or thing (e.g., devices, supplies, costs, charges) within that table."[23]

- A data model is "a representation of the data to be stored in a database and the relationships between the tables and data fields."[24]
  - » An entity relationship diagram, as illustrated below, is a form of data modeling.
- Diagrams to know:
  - » Entity relationship diagram (ERD): "A type of flowchart that illustrates how "entities" such as people, objects, or concepts relate to each other within a system."[25] Below we see a basic ERD of a relational database, with focus around the "Patient_Encounter" table.
    1. The "Patient_Encounter" table has data that is specific to that table, indicating that it is created within that data table.
       a. "Patient_Encounter_ID" would be a value in an EMR/EHR that equates to an encounter/account number.

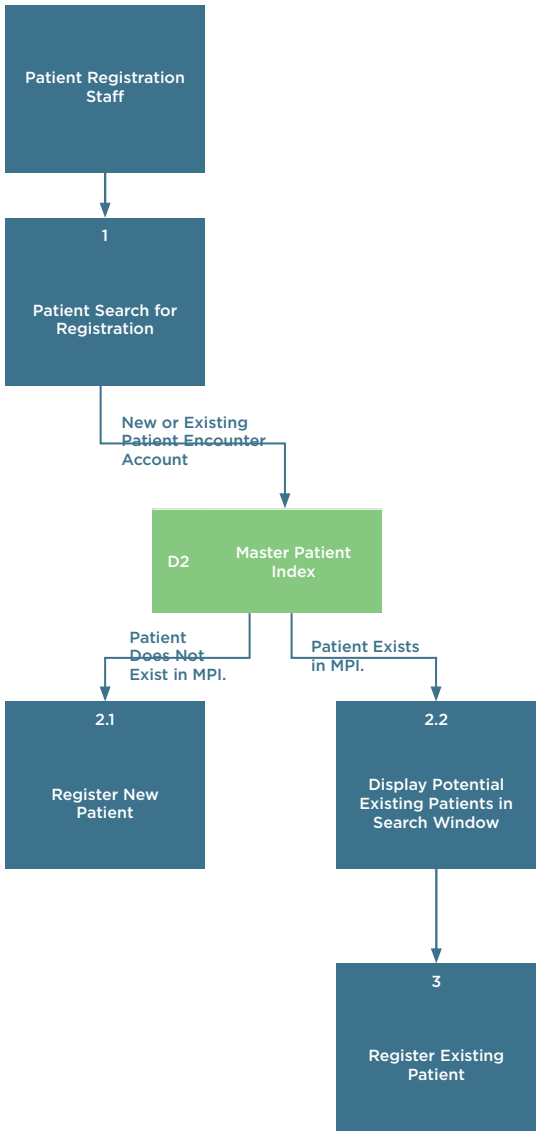
2. The "Patient_MRN_ID" in the "Patient_Encounter" table actually pulls from the "Patient_MRN" table.
   a. Within an EMR/EHR there is normally patient-level data and account level data. Based on this diagram we see that this is possible because the patient-level data is stored within its own data table, and the same is true of the encounter level information. This is important to remember when considering downstream impacts of data flow.



- Data Flow Diagram (DFD): "Maps out the flow of information for any process or system."[26] Through a DFD, we can map user and system interactions through the input and output of data. DFDs are beneficial tools as they can help trace the creation of data, how the data are used within the system, and what is the end result of the data. For example, below is a basic DFD tracking the patient registration process and how the system works in preventing the creation of a duplicate patient.

  1. Patient registration staff start this process.

2. They first conduct a patient search during their registration process.
3. The searched criteria is queried within the MPI database.
    a. If the patient does not exist in the system, the registration staff will be allowed to register a new patient.
    b. If there are results from the queried parameters, the system will display the similar results in the search window, and the registration staff will register the existing patient.



DEVELOPING TERMINOLOGIES AND TAXONOMIES OF HEALTH INFORMATION SYSTEMS

A map from a clinical reference terminology to a classification system is used to assist in automating the classification process in EHRs. The process of mapping is used to link one clinical classification, nomenclature, or terminology to another. Each map from a source classification, nomenclature, or terminology to a target requires a distinct function. There may be more than one purpose for mapping the individual sources, which then results in the need for additional maps.

In creating the maps between different schemes, is it important to understand the purpose of each mapping scheme and how the map from one classification, nomenclature, or terminology to the other will be used. This will be defined within the use case for the mapping. Since each scheme has its own function, the map becomes a way to link the source and target schemes. The selection of the schemes for mapping will depend on the definition as explained in the use case and what is ultimately necessary to ensure the intended outcome.

A reference terminology is ontology of concepts and the relationships that link them. A reference terminology allows concepts to be defined in a computer-readable format. By creating computer-readable definitions, a reference terminology supports reproducible transmission of patient data between information systems. It supports consistent and understandable recording of clinical events, and it is therefore a central feature for the function of EHRs. In addition, the computer-readable semantic network that is intrinsic in a reference terminology is a key component in enabling reliable and accurate data analysis.

The authors of "Essentials for Mapping from a Clinical Terminology" write that using a standard terminology is important for a wide variety of reasons.[27] The terminology is the link between recorded data and the intended goals of an EHR, which may include:

- Ensuring access to complete and legible clinical data with links to medical knowledge for real-time clinical decision support
- Interoperability that permits many different sites and providers to send and receive medical data in an understandable and usable manner, thereby speeding care delivery and reducing duplicate testing and duplicate prescribing
- Computer manipulation of standardized data used to find information relevant to individual patients for the purpose of producing automatic reminders or practitioner alerts
- Capability to form queries to look at standards of care for benchmarking, measuring and interpreting effectiveness, and quality improvement

ENSURING DATA MEETS INDUSTRY OR REGULATORY STANDARDS

Standards help establish the definitions for data elements in an EHR. They specify the type of data to be collected in each data field as well as the attributes and values for each field, all of which are captured in the organization's data dictionaries.

Data quality and data consistency are essential to ensuring patient safety, for communicating the delivery of healthcare services, for coordination of care, and for meeting reporting requirements. Evaluating the quality and consistency of data requires data standards.

For a list of structure and content standards, functional EHR standards, technical/interoperability standards, and standards development organizations, refer to Appendix A of the AHIMA Data Standards, Data Quality, and Interoperability Practice Brief.

DEVELOPING QUERIES FOR RECURRING DATA EXTRACTION FROM DATABASE

The rationale behind creating a database is typically to store data that you need to access for business purposes. While creating the actual database and its data entry forms and reports is a necessary first step, the day-to-day work on the database generally involves getting business information out of it. Within your database, you'll have different tables containing the data you want to work with.

To retrieve data from a database, you use queries. Structured query language, or SQL (pronounced "sequel"), is a flexible language that is often used to extract information from a relational database using queries. If the data in the database matches the conditions outlined in the SQL query, then the information will be retrieved and available for analysis.

It is very important for the SQL statement to be written correctly, otherwise the output could be wrong. A SQL statement can be written in different ways causing you to miss data, alter data, or create too much data.

Microsoft Excel, a proprietary spreadsheet application, may be used to create graphs, spreadsheets, and pivot tables. Tables and graphs are useful for depicting data, and Microsoft Excel can create a table or a variety of graphs using data entered into a spreadsheet.

Additionally, specialized software packages such as Statistical Analysis System (SAS) and Statistical Package for the Social Sciences (SPSS) make analyzing even the largest databases possible. The preference for one or the other of these statistical software packages may be based on company policy, education, and training of the individual using the package, or analyst preference.

SCRUBBING DATA

The issue of dealing with duplicate data has historically plagued data warehouse initiatives. Such situations require data scrubbing, also known as data cleansing, as well as record cleanup and record linkage processes. It is important to always keep a copy of the original data in case you have to review missing data or numbers that may seem incorrect.

Data scrubbing or cleansing involves the detection, removal, and correction of incorrect, incomplete, or poorly formatted data in health information technology systems. It deals with the detection and elimination of duplicate or redundant information, mistakes during data entry, invalid values, or non-populated values within each information system.

HEALTH INFORMATICS TRAINING

Informatics training should consist of a comprehensive overview of the fundamental uses of health information systems and appropriate clinical documentation standards in healthcare, so that the learner acquires a clear understanding of health informatics and its contribution to administrative and clinical best practice. The material covered should provide an entry-level foundation that covers the basics of efficient us of IT in healthcare and serves as excellent preparation for subsequent system-specific training by the employer.

In general, the fundamental knowledge and skills that are taught during informatics training should focus on technical, business and health care topics, such as database modeling and design, project management and the US healthcare system. While a background in the healthcare field or information technology/computer science is not required, it is often considered beneficial for the learner as he/she is familiar with the work setting and its challenges.

TRAIN CLINICIANS ON DOCUMENTATION

Training clinicians on documentation should begin with an introductory survey of fundamentals of health information technology and how information technology enables patient care and how information technology is used by healthcare providers and caregivers (e.g., electronic health records, health information exchange). This will also allow the informatics professional to understand the impact that health information technology can have on the quality of care and patient safety, big data and predictive analytics, clinical decision support and knowledge management, regulatory issues, consumerism and technology, systems integration, and virtual health.

After the foundations of health informatics are reviewed, the fundamentals of database design, management and implementation need to be covered. Specific topics to study should include the principles and methodologies of database design, database application development, normalization, referential integrity, security, relational database models, and database languages.

TRAIN HELP DESK PERSONNEL ON DELEGATED TASKS

Both the HIPAA Privacy Rule and the HIPAA Security Rule have training requirements. The HIPAA Privacy Rule training requirement is at 45 CFR § 164.530(b)(1). The HIPAA Security Rule training requirement is an administrative safeguard at 45 CFR § 164.308(a)(5). A sample can be found here.

The help desk or service desk is the resource for computer users in dealing with hardware and software access and utilization problems. The service may be centralized within a department, decentralized with one help desk to serve all departments, or virtual with online or telephone access using either internal staffing or external contracted providers.[28]

The management of the help desk involves both the aspects of service support and the delivery of the service; therefore, the training must address both dimensions. The delivery of the service follows established administrative policy and knowledge of the procedures becomes paramount to the training. For service support, incident management provides a methodology to stratify calls according to the level of support required as described below: [29]

- The first-level incident involves the initiation of the trouble ticket and handling of low level problems. A tip sheet of service level management data can be utilized for quick answers to common computer questions or indicate circumstances requiring the need for a higher level of support.
- Second-level incidents require expertise beyond the scope of first-level problems and are channeled to appropriate IT technicians for resolution.
- Third-level incidents are issues which require system analysis and determination of the root cause of the problem.

While first-level incident management serves to stratify calls, the trouble ticket initiated becomes an intricate part of both support tracking and quality assurance in the performance of help desk services in documenting the time spent for incident response and resolution. Technician and service performance can be tracked electronically for data analysis of response times in the incident resolution process for ongoing quality control for service management.

In addition to orientation training, ongoing training is needed for change management and for system and software upgrades from new releases. Training for disaster response is also a component in the event of service disruption to assure effective and efficient recovery of services.

Training in soft skills such as communication, courtesy, and patience as well as organizational skills should be included in orientation and in-service staff development programs. Confidence in performing the role of the help desk technician requires adequate background training in health information technology (HIT), and evidence of this level of expertise should be a prerequisite for the help desk position.

IDENTIFY DOCUMENTATION ERROR TRENDS/DEFICIENCIES

In order to be able to identify deficiencies and initiate quality improvement activities, the healthcare informatics professional should be familiar with effective project management strategies that can be applied to healthcare operations as well as other industries, which include team development, communication, stakeholder management, and organization dynamics.

A solid understanding of data management and data analytics will enable informatics professionals to make use of appropriate analytics, know its limitations to identify trends/errors/deficiencies and help with quality improvement. By working through examples and interpreting performance-based organizational issues while also concurrently identifying solutions for these same performance-based organizational issues, the health informatics professional will be able to identify best practices for engaging, implementing, and sustaining organizational change.

TRAIN USERS ON EFFICIENT USE OF HEALTH INFORMATION SYSTEMS

Before users attain specialized knowledge about specific systems, it is important to obtain a general understanding of healthcare enterprise operations, which include the entire information technology needs of every part of the hospital organization and management, including patient access services, ambulatory care, clinical practice and organization, nursing services, managing facilities and resources, personnel and staffing, and finance.

Users should be familiar with HIT standards as well as system interoperability. By learning more about standards, health information models, the IHE Initiative, HL7, DICOM, CCOW, CorbaMED, as well as the role of nonmedical standards in medical informatics (HTTP, XML, etc.), users will be more likely to make efficient use of health information systems.

While the efficient use of health information systems is significant, up-to-date knowledge about federal regulations/policies and their impact on healthcare informatics is required to ensure compliance. Since healthcare is highly regulated, any regulatory changes can have an effect on quality of care, payment methodologies, care delivery systems, and information management, privacy and security, fraud and abuse, confidentiality, antitrust law, intellectual property, the Joint Commission, disclosure, and compliance programs.

This will allow informatics professionals to have the knowledge and analytic tools needed to spot key issues, thereby better protecting the employer and ultimately the patients.

For foreign healthcare informatics professionals, it might also be useful to learn more about the key components of the American healthcare system that shape healthcare delivery in the United States.

For health information systems (HIS), a comprehensive training program can provide the necessary skills for the workforce in achieving the work goals and objectives of the organization. For the workforce, effective training provides the employee with the knowledge and tools in performing task completion according to established policy. For the organization, training and human resource development aligned with the strategic plan benefit the organization by ensuring maintenance of a workforce possessing the skills for current and future organizational needs. However, training challenges can arise from the diversity of a multilevel, multicultural, and multigenerational workforce to tailor appropriate training for the intended audience.

PLANNING A TRAINING PROGRAM

Achieving an effective training program requires planning. Using project management skills to develop a training plan can make training more effective. Typical steps that may be used in developing a training plan are:[30]

1. Conduct training needs analysis to determine the type, extent, and method(s) of instruction. Aspects to consider involving the target audience is the level of computer fluency, time scheduling, computer hardware available for training, and group size. Will it require a single or series of training events? Will it require classroom or can self-study with computer training modules be used asynchronously?

2. Prepare the plan.
    a. Develop a budget and set a schedule for the training event.
    b. Prepare the training event.
    c. Develop the deliverables such as handouts, examinations, or certificate of completion.
3. Determine location and method and test the plan.
    a. One-on-one
    b. Online, self-study
    c. Classroom
    d. Test the training plan.
4. Conduct training.
5. Control training.
    a. Measure effectiveness using performance assessments and user surveys.
    b. Make changes.
    c. Evaluate training performance (quality assurance).
    d. Documentation for participation as the final deliverable.

TYPES OF TRAINING

HIS training is multifaceted starting with new employee orientation structured to the organization's policies, departmental functions and policies, and individual training dimensions. The organization's new employee orientation can be accomplished in a group setting with a checklist of topics to include policies related to computer usage. This serves to introduce the employee to the basics of the administrative safeguards, physical, and technical safeguards in place and the expectations of the employee related HIS use. These can include discussion of general policies regarding the following:

- Use of personal devices for work (BYOD/Bring Your Own Device)
- Use of facility devices
- Password policies
- Privacy and security
- Social media use restrictions

Departmental orientation for new employees requires tailoring to the role-specific tasks aligned with the further orientation to specific policy and procedures for compliance to performing tasks and role-specific access to the computer system and other equipment within the department. It also introduces the employee to team members and an understanding of departmental functions. The individual training is the most specific because it involves assessment of the individual's expertise, age, disability, diversity in determining a learning methodology in approaching the training necessary for responsibility in performing specific tasks in the workflow and expected performance levels.

On-the-job training is a method of introducing an employee to new tasks using a trained mentor for one-to-one guidance in teaching and appropriate training for software applications. This mentoring approach allows for assessment of the employee's progress and performance. On-the-job training may involve techniques of job rotation or cross-training in achieving the variety of skills for the job and may also use computer training modules for self-study modules or simulations.[31]

Staff development training within the organization can be accomplished through in-service training initiatives to maintain and update skills including both soft skills such as communication and customer service and in response to the introduction of new technology or for regulatory compliance. Continuing education by

professionals, either internally or externally, is an additional role-specific training dimension and may be mandatory for credential and licensing requirements. The use of computer-based training in either synchronous in classroom or webinar environments or asynchronous format such as online courses are common modalities for the delivery of in-service and continuing education instruction.[32]

Informatics training should consist of a comprehensive overview of the fundamental uses of health information systems and appropriate clinical documentation standards in healthcare, so that the learner acquires a clear understanding of health informatics and its contribution to administrative and clinical best practice. The material covered should provide an entry-level foundation that covers the basics of efficient use of IT in healthcare and serves as excellent preparation for subsequent system-specific training by the employer.

PROJECT MANAGEMENT

A project has been defined as "series of inter-related and sequenced activities, managed by a single individual, designed and organized to accomplish a specific goal, within a limited timeframe, frequently with specific budgetary requirements."[33]

Project management is essential to the realization of the healthcare entity's business and patient care goals because projects are a means by which strategy is implemented and maintained.

Project plans are used to divide the project into phases. The phases are called the "project life cycle." The project life cycle defines the beginning and end of a project and it includes the following phases of each project:

- Initiation
- Planning
- Execution
- Closeout

There are key players to be assigned to every project, and the project plan should identify roles for each of these key players:

Project sponsor

- Ensures the project meets the business needs
- Provides funding
- Approves key deliverables
- Assists in issues resolution and change management

Project manager

- Assembles the project team
- Assigns tasks and activities
- Monitors progress, risk, issues
- Manages the scope
- Delivers on time and within budget
- Ensures project documentation is prepared

Project stakeholders

- Individuals or organizations that are actively involved in the project or whose interests may be positively or negatively affected as a result of the project
- May exert influence over the project and its results
- Includes subject matter experts and business liaisons

Project leader

- Responsible for a subproject
- Ensures technical aspects are delivered
- Monitors progress, risk, issues
- Manages the scope and the change management process
- Delivers on time and within budget
- Communicates status to project manager

Project team members

- Complete assigned tasks on time
- Work with other team members
- Monitor progress on their tasks
- Resolve issues
- Monitor changes and risks
- Advise others of potential problems

SAMPLE PROJECT PLAN

The AHIMA Body of Knowledge contains a sample project plan that can be used as a template. The sample project plan is available here: http://bok.ahima.org/PdfView?oid=301384.

SAMPLE PROJECT PLAN SCOPE AND OBJECTIVES STATEMENT

Project Objective

To construct a high-quality, custom office building within five months at cost not to exceed $150,000

Scope of Project

In scope: All tools and materials necessary to build office building to include heating, air conditioning, lighting, flooring, windows, electrical, plumbing

Out of scope: Building furnishings, such as desks, computers, equipment

Deliverables

- 2,200-square-foot finished office space
- High-efficiency heating and air conditioning with programmable thermostat
  - » Ergonomically designed cubicles to accommodate 20 clerical staff

Milestones

1. Permits approved—Date
2. Foundation poured—Date
3. Framing, sheathing, plumbing, electrical, and mechanical complete–Date
4. Final inspection—Date

Technical Requirements

1. Must meet local building codes
2. Must pass NFRC energy ratings
3. Insulation must meet commercial building standards

Limits and Exclusions

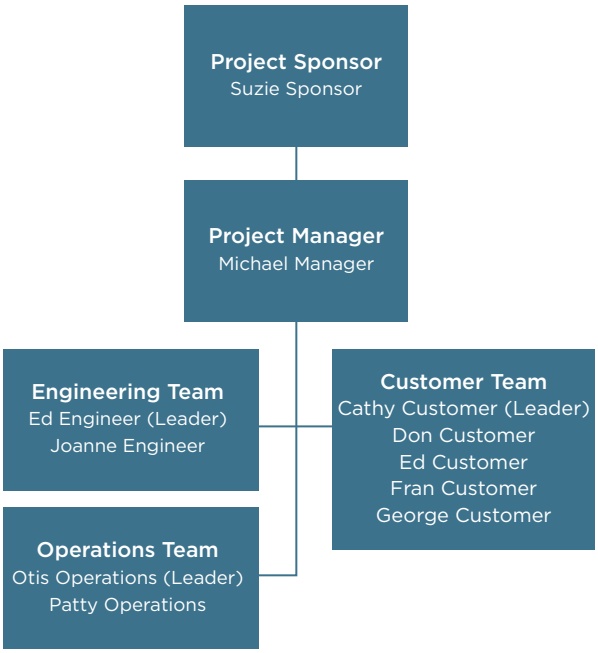
1. The space will be built to the specifications and design of the original blueprints to be provided by the purchaser.
2. Contractor reserves the right to contract out services.
3. Contractor is responsible for subcontracted work.
4. Site work limited to Monday through Friday, 8 a.m. to 6 p.m.

CUSTOMER REVIEW & SIGN OFF

Dr. John Doe and Dr. Jane Doe

ORGANIZATION AND RESOURCES

The graphic below outlines some of the key roles in a project plan. The various teams that are assigned to work on the project report to the project manager, who has overall accountability to the project sponsor. These roles and the responsibilities for each are often outlined in the project charter.



Source: AHIMA Information Governance Toolkit 1.0, 2015

NOTES

1. AHIMA. "Informatics and Data Analytics." http://www.ahima.org/topics/ida.
2. Hoyt, Robert E., and Ann Yoshihashi, eds. *Health Informatics: Practical Guide for Healthcare and Information Technology Professionals*, 6th ed. Pensacola, FL: Informatics Education, 2014.
3. Willemssen, Joel C. "Electronic Health Records: DOD and VA Interoperability Efforts Are Ongoing; Program Office Needs to Implement Recommended Improvements." US Government Accountability Office 10–332, January 2010. http://www.gao.gov/products/GAO-10-332.
4. Faculty Development and Instructional Design Center, Northern Illinois University. "Responsible Conduct in Data Management." 2005. https://ori.hhs.gov/education/products/n_illinois_u/datamanagement/datopic.html.
5. The Pell Institute and Pathways to College Network. "Evaluation Toolkit." http://toolkit.pellinstitute.org/evaluation-guide/analyze/analyze-quantitative-data.
6. Thorat, Rahul. "Data Mining." http://rahulthorat.weebly.com/data-mining.html.
7. Brown, Martin. "Data Mining Techniques." IBM DeveloperWorks, December 11, 2012. https://www.ibm.com/developerworks/library/ba-data-mining-techniques.
8. Ibid.
9. Ringlead, Inc. "The 4 Steps to Data Standardization." https://www.ringlead.com/4-steps-data-standardization.
10. Davoudi, Sion, Julie A. Dooling, Barbara Glondys, Theresa D. Jones, Lesley Kadlec, Shauna M. Overgaard, Kerry Ruben, and Annemarie Wendicke. "Data Quality Management Model (2015 Update)." Journal of AHIMA 86, no.10 (October 2015): expanded web version. http://library.ahima.org/doc?oid=107773.
11. Ibid.
12. Rinehart-Thompson, Laurie A. *Introduction to Health Information Privacy and Security*. Chicago: AHIMA Press, 2013, p. 24.
13. Ibid.
14. AHIMA. "Analysis of Modifications to the HIPAA Privacy, Security, Enforcement, and Breach Notification Rules under the Health Information Technology for Economic and Clinical Health Act and the Genetic Information Nondiscrimination Act; Other Modifications to the HIPAA Rules." 2013. http://library.ahima.org/PdfView?oid=106127.
15. AHIMA Practice Brief. "Authorization Requirements for the Disclosure of Protected Health Information (2013 update)." 2013. http://bok.ahima.org/doc?oid=107284.
16. Business Dictionary. "Gap Analysis." http://www.businessdictionary.com/definition/gap-analysis.html.
17. Business Dictionary. "Workflow Diagram (WFD)." http://www.businessdictionary.com/definition/workflow-diagram-WFD.html.
18. Khan, Vaseem. "Responsibilities of a Six Sigma Project Champion." Bright Hub Project Management, October 12, 2012. http://www.brighthubpm.com/resource-management/49034-responsibities-of-a-project-champion.
19. US Bureau of Labor Statistics. "Database Administrators." December 17, 2015. https://www.bls.gov/ooh/computer-and-information-technology/database-administrators.htm.
20. International Organization for Standardization. "ISO/IEC 11179–1:2004(en) Information technology—Metadata registries (MDR)—Part 1: Framework." https://www.iso.org/obp/ui/#iso:std:iso-iec:11179:-1:ed-2:v1:en.

21. AHIMA. *Health Data Analysis Toolkit*. 2014.
22. Tilley, Scott and Harry Rosenblatt. *Systems Analysis and Design*. Boston, MA: Cengage Learning, 2016.
23. Mon, Donald T. "Relational Database Management: What You Need to Know." *Journal of AHIMA* 74, no.10 (November 2003): 40–45. http://library.ahima.org/doc?oid=59721.
24. White, Susan. *A Practical Approach to Analyzing Healthcare Data*. Chicago: AHIMA Press, 2016.
25. Lucidchart. "What is an Entity Relationship Diagram?" https://www.lucidchart.com/pages/er-diagrams/b?er=1.
26. Lucidchart. "What is a Data Flow Diagram?" https://www.lucidchart.com/pages/data-flow-diagram.
27. Imel, Margo, and Kathy Giannangelo. "Essentials for Mapping from a Clinical Terminology." 2004 IFHRO Congress & AHIMA Convention Proceedings, October 2004. http://library.ahima.org/doc?oid=58226.
28. Glandon, G. L., D.J. Slovensky, and D.H. Smaltz. "HIT Service Management." In *Information Systems for Healthcare Management* (7th ed.). Chicago, IL: Health Administration Press, 2014, p. 215.
29. Glandon, G. L., D.J. Slovensky, and D.H. Smaltz. "HIT Service Management." p. 215.
30. Coplan, S., and D. Masuda. "Change Management." In *Project Management for Healthcare Information Technology*, pp. 226–227. New York, NY: McGraw-Hill, 2011.
31. Oachs, P.K. "Employee Training and Development." In *Health Information Management: Concepts, Principles, and Practice* (5th ed., pp. 761–762). Chicago, IL: AHIMA Press, 2016.
32. Oachs, P.K. "Employee Training and Development." p. 762.
33. Carey, Susan. "Project Management Tools & Techniques." Presentation at the AHIMA Convention and Exhibit, Salt Lake City, UT, October 2011. https://newsletters.ahima.org/newsletters/ICDTen/2011/December/Carey_2011%20ahima.pdf.

REFERENCES

Brittain, J. Michael, and A. C. Norris. "Delivery of Health Informatics Education and Training." *Health Information & Libraries Journal* 17, no.3 (2000): 117–128.

Carnegie Mellon University Open Learning Initiative. "Statistical Reasoning." https://oli.cmu.edu/courses/free-open/statistical-reasoning-course-details.

Covvey, H.D., and A.B. Pidduck. "Health Informatics Education." (1999). https://cs.uwaterloo.ca/research/tr/1999/CS-1999-24.pdf.

Datamark. "Unstructured Data in EHR Systems: Challenges and Solutions." October 2013. http://insights.datamark.net/white-papers/unstructured-data-in-electronic-health-record-systems-challenges-and-solutions.

HHS.gov. "HIPAA for Professionals." https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification.

Murphy, Jeannette, Katja Stramer, Susan Clamp, Penny Grubb, Julian Gosland, and Sue Davis. "Health Informatics Education for Clinicians and Managers—What's Holding Up Progress?" *International Journal of Medical Informatics* 73, no.2 (2004): 205–213.

Observational Health Data Sciences and Informatics. "Data Standardization." https://www.ohdsi.org/data-standardization.

Oracle Warehouse Builder User's Guide. "Identifying Data Sources and Importing Metadata." https://docs.oracle.com/cd/B31080_01/doc/owb.102/b28223/concept_integrate_import.htm.

Palace, Bill. "Data Mining." Technology note prepared for Management 274A, Anderson Graduate School of Management at UCLA, Spring 1996.

Progress Telerik Reporting. "Designing Reports." 2017. http://docs.telerik.com/reporting/designing-reports.

Techopedia. "Data Extraction." https://www.techopedia.com/definition/25328/data-extraction.

APPENDIX A: BIBLIOGRAPHY OF ADDITIONAL RECOMMENDED READING

Abrams, Kelly, Kathy Giannangelo, Sue Schneider, and Rita A. Scichilone. "New Name of the HIM Game." *Journal of AHIMA* 85, no.2 (February 2014): 34–39. http://bok.ahima.org/doc?oid=300241.

Acker, Beth A., and Kimberly A. Baldwin-Stried. "Get Ready for New Opportunities in Healthcare Standards." *Global Standards Update* (May 2013). http://bok.ahima.org/doc?oid=301596.

Alakrawi, Zahraa M. "Clinical Terminology and Clinical Classification Systems: A Critique Using AHIMA's Data Quality Management Model." *Perspectives in Health Information Management* (Summer 2016): 1–19. http://perspectives.ahima.org/clinical-terminology-and-clinical-classification-systems-a-critique/.

Alakrawi, Zahraa M., Valerie J.M.Watzlaf, Jon Aborde, Carey Johnson, and Timothy Johnson. "Utilizing Open-Source Government Data Sets in Health Information Management Teaching: An Application in Statistics and Data Analytics." *Educational Perspectives in Health Informatics and Information Management* (Summer July 2016). http://eduperspectives.ahima.org/utilizing-open-source-government.

Alexander, Gregory L., Lori Popejoy, Vanessa Lyons, Sue Shumate, Jessica Mueller, Colleen Galambos, Colleen, Amy Vogelsmeier, Marilyn Rantz, and Marcia Flesner. "Exploring Health Information Exchange Implementation Using Qualitative Assessments of Nursing Home Leaders." *Perspectives in Health Information Management* (Fall 2016): 1–12. http://perspectives.ahima.org/exploringhealthinformationexchangeimplementation.

Ayatollahi, Haleh, Masoud Roozbehi, and Hamid Haghani. "Physicians' and Nurses' Opinions about the Impact of a Computerized Provider Order Entry System on Their Workflow." *Perspectives in Health Information Management* (Fall 2015): 1–16. http://perspectives.ahima.org/physicians-and-nurses-opinions-about-the-impact-of-a-computerized-provider-order-entry-system-on-their-workflow.

Bailey-Woods, Linda, Megan Munns, Anna Orlova, Harry B. Rhodes, and Diana Warner. "Guiding the Development of Health Information Technology Standards for HIM Practices." *Journal of AHIMA* 86, no.8 (August 2015): 40–43. http://bok.ahima.org/doc?oid=107722.

Baus, Adam, Keith Zullig, Dustin Long, Charles Mullett, Cecil Pollard, Henry Taylor, and Jeffrey Coben. "Developing Methods of Repurposing Electronic Health Record Data for Identification of Older Adults at Risk of Unintentional Falls." *Perspectives in Health Information Management* (Spring 2016): 1–32. http://perspectives.ahima.org/developing-methods-of-repurposing-electronic-health-record-data-for-identification-of-older-adults-at-risk-of-unintentional-falls.

Bhattacharyya, S.B., and Diana Warner. "Semantic Content in EHR Systems." *Journal of AHIMA* 87, no.6 (June 2016): 38–39. http://bok.ahima.org/doc?oid=301556.

Bradley, Paul. "Tying Together CDI, ICD–10, and Healthcare Data Analytics: How CDI Programs and ICD–10 Are Impacting Data Analytics—and The Benefits Healthcare Organizations Are Seeing." *Journal of AHIMA* 87, no.10 (October 2016): 28–30. http://bok.ahima.org/doc?oid=301901.

Butler, Mary. "Teaching the 'New' HIM: Educators Integrating Informatics, Data Analytics, and Information Governance Into HIM Programs." *Journal of AHIMA* 88, no.5 (May 2017): 16–19. http://bok.ahima.org/doc?oid=302102.

Butler, Mary. "How to Break Into Analytics and Informatics." *Journal of AHIMA* 87, no.10 (October 2016): 22–26. http://bok.ahima.org/doc?oid=301895.

Butler, Mary. "HIM's Professional Shake-Up Wake-Up: Moving Beyond Reality 2016 and 'Reimagining' HIM for a Quickly Changing Healthcare Industry." *Journal of AHIMA* 87, no.5 (May 2016): 16–21. http://bok.ahima.org/doc?oid=301433.

Cange, John R. "Preventing Hospital-Acquired Infections Starts with Data Collection." *Journal of AHIMA* 87, no.1 (January 2016): 44–46. http://bok.ahima.org/doc?oid=301324.

Coffey, Patricia S., Susan Postal, Susan M. Houston, and Jon W. McKeeby. "Lessons Learned from an Electronic Health Record Downtime." *Perspectives in Health Information Management* (Summer 2016): 1–7. http://perspectives.ahima.org/lessons-learned-from-an-electronic-health-record-downtime.

DuBravec, Daniel A. "Can EHRs Handle Genomics Data?" *Journal of AHIMA* 86, no.11 (November 2015): 28–31. http://bok.ahima.org/doc?oid=107793

Dolezel, Diane, Mary Clisbee, and Alexander McLeod. "Master of Health Information Management Students' First-Year Experiences: A Case Study." *Educational Perspectives in Health Informatics and Information Management* (Fall 2016). http://eduperspectives.ahima.org/masterofhimstudents/.

Dooling, Julie A., Shannon H. Houser, Raymound Mikaelian, and Clarice Pittillo Smith. "Transitioning to a Data-Driven, Informatics-Oriented Department." *Journal of AHIMA* 87, no.10 (October 2016): 58–62. http://bok.ahima.org/doc?oid=301903.

Eramo, Lisa A. "Becoming a Data Master." *Journal of AHIMA* 88, no.3 (March 2017): 18–21. http://bok.ahima.org/doc?oid=302050.

Eramo, Lisa A. "Standardizing Interoperability is a Team Effort." *Journal of AHIMA* 87, no.11 (November 2016): 18–23. http://bok.ahima.org/doc?oid=301914.

Feldman, Sue, and Shannon H. Houser. "Automating HIM through Health Information Exchange: An Informatics Approach." *Journal of AHIMA* 88, no.1 (January 2017): 32–35. http://bok.ahima.org/doc?oid=302021.

Gellert, George A, George S. Conklin, and Lynn A. Gibson. "Secure Clinical Texting: Patient Risk in High-Acuity Care." *Perspectives in Health Information Management* (Winter 2017): 1–5. http://perspectives.ahima.org/secureclinicaltexting.

Gibbs, David, Barbara Hewitt, and Alexander McLeod. "Gamification of Electronic Health Records: A Systematic Literature Review." *Educational Perspectives in Health Informatics and Information Management* (Spring 2016). http://eduperspectives.ahima.org/the-gamification-of-electronic-health-records-a-systematic-literature-review.

Glondys, Barbara, and Lesley Kadlec. "EHRs Serving as the Business and Legal Records of Healthcare Organizations (2016 Update)." *Journal of AHIMA* 87, no.5 (May 2016): extended online version. http://bok.ahima.org/doc?oid=301443.

Gordon, Lynne Thomas. "What Matters: Not Where We Are, But Where We Will Be." *Journal of AHIMA* 88, no.1 (January 2017): 13. http://bok.ahima.org/doc?oid=302027.

Gordon, Lynne Thomas. "Let's Talk About Informatics and HIM." J*ournal of AHIMA* 87, no.10 (October 2016): 19. http://bok.ahima.org/doc?oid=301909.

Grzybowski, Darice, and Anna Orlova. "Standards, Information Governance, Informatics: Essential Components of HIM Workforce Education." *Journal of AHIMA* 88, no.5 (May 2017): 48–50. http://bok.ahima.org/doc?oid=302133.

Harris, Susie T., Paul D. Bell, and James Anthony Woodward. "First Steps in Starting a Master of Science Program in Health Informatics and Information Management." *Educational Perspectives in Health Informatics and Information Management* (Fall 2016). http://eduperspectives.ahima.org/firststeps/.

Haugen, Mary Beth, Paula Dascher, and Barbara Manor. "Uniting HIM and IT: Lessons Learned Offer Tangible Takeaways to Bridge the Disconnect and Bring Departments Together." *Journal of AHIMA* 88, no.1 (January 2017): 24–27. http://bok.ahima.org/doc?oid=302020.

Hemmat, Morteza, Haleh Ayatollahi, Mohammad Reza Maleki, and Fatemeh Saghafi. "Future Research in Health Information Technology: A Review." *Perspectives in Health Information Management* (Winter 2017). http://perspectives.ahima.org/futureresearchinhit.

Houser, Shannon H., and Jean M. Meadows. "Transitioning to Electronic Clinical Quality Measures in the Informatics Era." *Journal of AHIMA* 88, no.2 (February 2017): 28–31. http://bok.ahima.org/doc?oid=302039.

Houser, Shannon H., Raymound Mikaelian, and Evita Payton. "HIM Connections in the Kaleidoscope of Informatics." *Journal of AHIMA* 87, no.11 (November 2016): 40–43. http://bok.ahima.org/doc?oid=301920.

Hoyt, Robert, Steven Linnville, Stephen Thaler, and Jeffrey Moore. "Digital Family History Data Mining with Neural Networks: A Pilot Study." *Perspectives in Health Information Management* (Winter, January 2016). http://perspectives.ahima.org/digital-family-history-data-mining-with-neural-networks-a-pilot-study/.

Hylock, Ray, and Susie T. Harris. "Healthcare Database Management for Health Informatics and Information Management Students: Challenges and Instruction Strategies—Part 1." *Educational Perspectives in Health Informatics and Information Management* (Winter 2016). http://eduperspectives.ahima.org/healthcare-database-management-for-health-informatics-and-information-management-students-challenges-and-instruction-strategies-part-1.

Just, Beth Haenke, David Marc, Megan Munns, and Ryan Sandefer. "Why Patient Matching Is a Challenge: Research on Master Patient Index (MPI) Data Discrepancies in Key Identifying Fields." *Perspectives in Health Information Management* (Spring 2016). http://perspectives.ahima.org/why-patient-matching-is-a-challenge-research-on-master-patient-index-mpi-data-discrepancies-in-key-identifying-fields.

Kadlec, Lesley. "Climbing Higher: Bridging the Gap to Advanced Degrees in HIM (2016 update)." (AHIMA Practice Brief, January 2016). http://bok.ahima.org/doc?oid=301355.

Kadlec, Lesley, and Annessa Kirby. "HIM as the Middleman to Informatics and Data Analytics." *Journal of AHIMA* web site, October 06, 2016. http://journal.ahima.org/2016/10/06/him-as-the-middleman-to-informatics-and-data-analytics/.

Khurshid, Anjum, Mark L. Diana, and Rahul Jain. "Health Information Exchange Readiness for Demonstrating Return on Investment and Quality of Care." *Perspectives in Health Information Management* (Fall 2015). http://perspectives.ahima.org/health-information-exchange-readiness-for-demonstrating-return-on-investment-and-quality-of-care.

Lee, Wei-Chen, Sreenivas P. Veeranki, Hani Serag, Karl Eschbach, and Kenneth D Smith. "Improving the Collection of Race, Ethnicity, and Language Data to Reduce Healthcare Disparities: A Case Study from an Academic Medical Center." *Perspectives in Health Information Management* (Fall 2016). http://perspectives.ahima.org/improvingthecollectionofrace/.

Marc, David. Janet Robertson, Leslie Gordon, Zakevia D. Green-Lawson, David Gibbs, Kayce Dover, and Michelle Dougherty. "What the Data Says About HIM Professional Trends." *Journal of AHIMA* 88, no.5 (May 2017): 24–31. http://bok.ahima.org/doc?oid=302104.

Martin, Melissa. "Bringing HIM to Every Corner of Healthcare." *Journal of AHIMA* 87, no.8 (August 2016): 6. http://bok.ahima.org/doc?oid=301833

McCann, Linda F., William A. Henciak, and Elizabeth Durand. "Reinventing HIM as Enterprise Content Management." *Journal of AHIMA* 87, no.7 (July 2016): 34–38. http://bok.ahima.org/doc?oid=301748.

Meigs, Stephen L., and Michael Solomon. "Electronic Health Record Use a Bitter Pill for Many Physicians." *Perspectives in Health Information Management* (Winter, January 2016). http://perspectives.ahima.org/electronic-health-record-use-a-bitter-pill-for-many-physicians.

Monga, Kapila. "Applying Predictive Analytics: The Role of Artificial Neural Networks in Predicting Alzheimer's Disease." (Journal of AHIMA website), December 22, 2016. http://bok.ahima.org/doc?oid=302091.

Orlova, Anna, and Kenneth Salyards. "Understanding Information in EHR Systems: Paving the Road for Semantic Interoperability through Standards." *Journal of AHIMA* 87, no.9 (September 2016): 44–47. http://bok.ahima.org/doc?oid=301861.

Orlova, Anna, Harry B. Rhodes, and Diana Warner. "HIT Standards for HIM Practices in an Interoperable World." *Journal of AHIMA* 87, no.2 (February 2016): 38–41. http://bok.ahima.org/doc?oid=301347.

Orlova, Anna. "Informatics and HIM: Enabling Semantic Interoperability and the Learning Health System." *Journal of AHIMA* 86, no.9 (September 2015): 46–49. http://bok.ahima.org/doc?oid=107747.

Overgaard, Shauna M. "Examining the Computation of the Chi-square Statistical Test." *Journal of AHIMA* web site, October 22, 2015. http://journal.ahima.org/2015/10/22/examining-the-computation-of-the-chi-square-statistical-test.

Overgaard, Shauna M., and Julie A. Dooling. "Measurement Brings Meaning to Population Health." *Journal of AHIMA* website, August 27, 2015. http://journal.ahima.org/2015/08/27/measurement-brings-meaning-to-population-health.

Potter, Andrew J., Marcia M. Ward, Nabil Natafgi, Fred Ullrich, A. Clinton MacKinney, Amanda L. Bell, and Keith J. Mueller. "Perceptions of the Benefits of Telemedicine in Rural Communities." *Perspectives in Health Information Management* (Summer 2016). http://perspectives.ahima.org/perceptions-of-the-benefits-of-telemedicine-in-rural-communities.

Sawatsky, Elaine, and Diana Warner. "Managing PHI Content Standards." *Journal of AHIMA* 87, no.4 (April 2016): 38–41. http://bok.ahima.org/doc?oid=301415.

Taylor, Nathan Patrick. "Getting Your Feet Wet with Open-Source Statistical Languages." *Journal of AHIMA* web site, January 25, 2016. http://journal.ahima.org/2016/01/25/getting-your-feet-wet-with-open-source-statistical-languages.

Taylor, Nathan Patrick. "The 'Ins' and 'Outs' of Data Conversions." *Journal of AHIMA* 86, no.11 (November 2015): 46–47. http://bok.ahima.org/doc?oid=107794

Viangteeravat, Teeradache, and Naga Satya V. Rao Nagisetty. "Giving Raw Data a Chance to Talk: A Demonstration of Exploratory Visual Analytics with a Pediatric Research Database Using Microsoft Live Labs Pivot to Promote Cohort Discovery, Research, and Quality Assessment." *Perspectives in Health Information Management* (January 2014). http://perspectives.ahima.org/giving-raw-data-a-chance-to-talk-a-demonstration-of-exploratory-visual-analytics-with-a-pediatric-research-database-using-microsoft-live-labs-pivot-to-promote-cohort-discovery-research-and-quality.

VanOver, Jennifer, and Katherine Lusk. "From the HIE Point of View: When is Data Too Old to Use? It Depends." *Journal of AHIMA* web site, November 16, 2016. http://journal.ahima.org/2016/11/16/from-the-hie-point-of-view-when-is-data-too-old-to-use-it-depends.

Wang, Tiankai, and Diane Dolezel. "Usability of Web-based Personal Health Records: An Analysis of Consumers' Perspectives." *Perspectives in Health Information Management* (Spring 2016). http://perspectives.ahima.org/usability-of-web-based-personal-health-records-an-analysis-of-consumers-perspectives.

Warzel, Denise B. and Dianne M. Reeves. "Cross-walking Health Content Standards Using the ISO/IEC 11179 Metadata Registries Standard." *Journal of AHIMA* 87, no.7 (July 2016): 46–49. http://bok.ahima.org/doc?oid=301752.

Wendicke, Annemarie. "What Makes Data Meaningful? The Important Role of Data Structures." *Journal of AHIMA* 87, no.3 (March 2016): 34–36. http://bok.ahima.org/doc?oid=301394.

Wiedemann, Lou Ann. "In Data Analytics and Informatics, One is the Loneliest Number." *Journal of AHIMA* 88, no.5 (May 2017): 44–47, 68. http://bok.ahima.org/doc?oid=302132.

Williams, Patricia A.H. "Standards for Safety, Security, and Interoperability of Medical Devices in an Integrated Health Information Environment." *Journal of AHIMA* 88, no.4 (April 2017): 32–34. http://bok.ahima.org/doc?oid=302078.

APPENDIX B: SAMPLE ROLE DESCRIPTIONS—INFORMATICS ANALYSTS

Informatics analysts can be found in a variety of roles and their job titles vary according to the type of healthcare organization in which they are employed and the variety of responsibilities they are required to perform. An informatics professional working in a healthcare research department may be called a research informatics analyst, while the one working in an EHR deployment may be titled an analyst or a clinical informatics analyst. These professionals are involved in the organization and management of health information and in ensuring that it can be accessed wherever and whenever needed to fill its requirements.

There are a variety of informatics analyst roles in healthcare. Below are some basic descriptions.

Health Insurance Informatics Analyst: The health insurance informatics professional may conduct analysis of risk data using descriptive and inferential statistics to extend cost-effective solutions that increase the quality and utilization of claims data and assist in applying a predictive model on trends for rate development.

Clinical Informatics Analyst or Implementation Analyst: The clinical informatics analyst is concerned with carrying out the role of developing, implementing, organizing, and maintaining the health information and associated reporting requirements in such a way as to document and explain the operational and financial performance of the organization along with health outcomes. The analyst will work in collaboration with quality staff, executives, clinicians, health information professionals, and others to design both ad hoc and scheduled reports. They will assist in monitoring the quality programs and make efforts to improve outcomes. They will compile health data from multiple sources including the EHR and are involved in importing health data into appropriate databases. They are responsible for selecting the appropriate tools and using appropriate methodology to maintain all existing informatics programs and ensuring appropriate data transfer so all necessary reports are well designed and documented to increase organizational efficiency.

Health Informatics or Revenue Cycle Analyst: May be responsible for integrating the analysis of healthcare data through use of a variety of resources such as medical claims data, pharmacy claims, health information systems, electronic health records, etc. The health informatics analyst may develop tools and applications and will study data through the use of various statistical approaches and by ensuring that the presented information is clear and concise in order to meet the needs of the organization. They may make recommendations for improvement of existing policies and the development of new policies that ensure cost effectiveness and efficiency. They communicate results of their findings, retrieve, analyze data, and provide the necessary recommendations for quality information, such that data issues are significantly reduced, resulting in enhanced data processes and methodologies.

ESSENTIAL SKILLS

Informatics analysts should have a depth of understanding of analytical methodologies, health technology systems, as well as a strong knowledge of all basic statistical concepts. Excellent problem-solving skills along with good organizational and research skills are required. Exceptional communication skills are needed to handle a diverse population of staff and clinicians. Analysts are required to interact with staff and management at all levels within the organization. Must be well versed with computers and technology and have the capability to identify trends, opportunities, and risks or other issues.

EDUCATIONAL REQUIREMENTS

The educational requirements vary depending on the requirements of the organization. A minimum of an associate degree is typically required. A bachelor's degree in statistics, computer science, business, math, health information or informatics management, or a related field is helpful, and may be required in some instances. Additional training, certification, or experience in a similar position is helpful and may be required.

SALARY DETAILS

Salary will vary and will be dependent upon the location, size, and type of healthcare organization.

CAREER OUTLOOK

Career outlook is strong, as employment opportunities are on rise in healthcare and the rate of growth rate is expected to remain high for healthcare informatics professionals of all types.

APPENDIX C: GLOSSARY OF TERMS

Source: *AHIMA Health Data Analysis Toolkit, 2014.*

CALCULATIONS AND STATISTICS

Alpha level: The probability of making a type I error. The alpha level should be based on the cost of incorrectly rejecting the null hypothesis. In clinical studies, the alpha level is typically set low (1 percent or 5 percent)

Analysis of variance (ANOVA): Test used to determine the differences among two or more means

Average length of stay (ALOS): The mean length of stay for hospital inpatients discharged during a given period

Bar chart: A graphic technique used to display frequency distribution data that fall into categories

Chi-square test: A statistical calculation used to determine whether proportions in a randomly drawn sample are significantly different from the underlying or theoretical population proportions

Confidence interval (CI): An interval that has a certain probability (confidence level) of including the true value of a population parameter. CIs may be calculated for population means, proportions, and standard deviations, etc.

Correlation: The existence and degree of linear relationship among factors

Dependent variable: A measurable variable in a research study that depends on an independent variable

Derived attribute: An attribute whose value is based on the value of other attributes (e.g., current date minus date of birth yields the derived attribute age)

Descriptive statistics: A set of statistical techniques used to describe data such as means, frequency distributions, and standard deviations; statistical information that describes the characteristics of a specific group or a population

F test: The ratio of the between-group variance to the within-group variance in the ANOVA procedure. If the *F* ratio is statistically significant (the *F* value equals or exceeds the critical value of *F*), the observed differences between the group means of the independent variables under study will be significantly different from each other. See also **analysis of variance (ANOVA)**

Frequency distribution: A table or graph that displays the number of times (frequency) a particular observation occurs

Geometric mean length of stay (GMLOS): Statistically adjusted value of all cases of a given DRG, allowing for the outliers, transfer cases, and negative outlier cases that normally would skew the data. The GMLOS is used to compute hospital reimbursement for transfer cases

Hypothesis: A statement that describes a research question in measurable terms

Independent variable: An antecedent factor that researchers manipulate directly

Inferential statistics: Statistics that are used to draw conclusions regarding a population parameter on the basis of a sample

Length of stay (LOS): The total number of patient days for an inpatient episode, calculated by subtracting the date of admission from the date of discharge. If the admission and discharge are on the same day, the LOS generally is set to one day

Line graph: A graphic technique used to illustrate the relationship between continuous measurements; it consists of a line drawn to connect a series of points on an arithmetic scale and often is used to display time trends.

HEALTH DATA ANALYSIS

Mean: A measure of central tendency that is determined by calculating the arithmetic average of the observations in a frequency distribution

Measures of central tendency: The typical or average numbers that are descriptive of the entire collection of data for a specific population

Median: A measure of central tendency that shows the midpoint of a frequency distribution when the observations have been arranged in order from lowest to highest

Mode: A measure of central tendency that consists of the most frequent observation in a frequency distribution

Normal distribution: A theoretical family of continuous frequency distributions characterized by a symmetric bell-shaped curve, with an equal mean, median, and mode; any standard deviation; and half of the observations above the mean and half below it

Null hypothesis: A hypothesis that states there is no association between the independent and dependent variables in a research study. The null hypothesis often represents the status quo or a state of no statistical difference in a study

Pie chart: A graphic technique in which the proportions of a category are displayed as portions of a circle (like pieces of a pie)

***P* value:** The probability that the observed difference could have been obtained by chance alone, given random variation and a single test of the null hypothesis

Qualitative analysis: In healthcare data, determining that the data accurately portray the care that was administered and that the content is correct

Quantitative analysis: In healthcare data, analyzing aggregate data for patterns. Typically, quantitative analysis requires numeric data for calculations

Range: A measure of variability that is the difference between the smallest and largest observations in a frequency distribution

Rank: Denotes a score's position in a group relative to other scores that have been organized in order of magnitude

Rate: A measure used to compare an event across time; it is a comparison of the number of times an event did happen (numerator) with the number of times an event could have happened (denominator). See also **ratio**

Ratio: A calculation performed by dividing one quantity by another. It is also a general term that can include a number of specific measures such as proportion, percentage, and rate

Regression analysis: Statistical technique that uses an independent variable to predict the value of a dependent variable. In the inpatient psychiatric facility prospective payment system (IPF PPS), patient demographics and length of stay (independent variables) were used to predict cost of care (dependent variable)

Sample: A set of elements drawn from and analyzed to estimate the characteristics of a population

Sample size: The number of subjects needed in a study to represent a population

Scatter diagram: A graph that visually displays relationships among factors

Standard deviation: A measure of variability that describes the deviation from the mean of a frequency distribution in the original units of measurement—the square root of the variance

Stratified random sampling: The process of selecting the same percentages of subjects for a study sample as they exist in the subgroups (strata) of the population

Systematic sampling: The process of selecting a sample of subjects for a study by drawing every *n*th unit on a list

***t*-distribution:** The *t*-distribution is a probability distribution that is bell shaped and centered at zero, much like the standard normal distribution. The distribution is defined by the degrees of freedom, which determines the spread (or width) of the distribution. The *t*-distribution typically is used to determine the significance of a *t*-test of hypotheses regarding a population mean

***t*-test:** Assesses whether the means of two groups are statistically different from each other; appropriate when comparing the means of two groups

Trend: A long-term movement in an ordered series, say a time series, which may be regarded, together with the oscillation and random component, as generating the observed values[1]

Type I error: An error in which the researcher erroneously rejects the null hypothesis when it is true

Type II error: An error in which the researcher erroneously fails to reject the null hypothesis when it is false

Variable: A factor or quantity capable of assuming any of a set of values

Variance: A measure of variability that gives the average of the squared deviations from the mean—the difference between the planned or expected value and the actual value

DATABASE TERMS

Clinical data repository (CDR): A central database that focuses on clinical information

Data dictionary: A descriptive list of the data elements to be collected in an information system or database, the purpose of which is to ensure consistency of terminology

Data element: An individual fact or measurement that is the smallest unique subset of a database

Data mart: A well-organized, user-centered, searchable database system that usually draws information from a data warehouse to meet the specific needs of users

Data mining: The process of extracting information from a database and then quantifying and filtering discrete, structured data

Data repository: An open-structured database that is not dedicated to the software of any particular vendor or data supplier, in which data from diverse sources are stored so that an integrated, multidisciplinary view of the data can be achieved; also called a "central data repository" or, when related specifically to healthcare data, a "clinical data repository"

Data warehouse: A database that makes it possible to access information from multiple databases and combine the results into a single query and reporting interface. See also clinical data repository

Database: An organized collection of data, text, references, or pictures in a standardized format, typically stored in a computer system for multiple applications

Decision support system (DSS): A computer-based system that gathers data from a variety of sources and assists in providing structure to the data by using various analytical models and visual tools to facilitate and improve the ultimate outcome in decision-making tasks associated with non-routine and non-repetitive problems

Edit: A condition that must be satisfied before a computer system can accept data

Interface engine: A computer program that isolates the task of transferring data from one database to another

Metadata: Data about data that describe a specific item's content

Query: The process of making a logical inquiry or request from a database

Structured Query Language (SQL): A fourth-generation computer language that includes both data definition language and data manipulation language components and is used to create and manipulate relational databases

1. Dodge, Yadolah, ed. *The Oxford Dictionary of Statistical Terms.* New York: Oxford University Press, 2003.

APPENDIX D: STATISTICAL TESTS AND CHART TYPES FOR GRAPHICAL REPRESENTATION

Statistical testing provides a methodology for making quantitative decisions regarding processes. The purpose is to help determine whether there is enough evidence to "reject" the hypothesis that has been speculated about the process. The speculative statement is called the null hypothesis. Determining not to reject the null hypothesis may be the correct result if we want to continue to act as if we suspect that the null hypothesis is a true statement. Or we may find that we have a different result than what we initially suspected to be true, thereby indicating we do not have sufficient data to "prove" something by rejecting the null hypothesis.

Statistical Tests: The decision of the type of statistical test to use is dependent upon the type of variable (quantitative or qualitative), relationship (difference or correlation), and quantity of variables being analyzed.

Test	Use	Null Hypothesis	Use Example
Chi square test	Testing for a relationship between two categorical variables (nominal/ordinal)	There is no association between the categories. (They are independent,)	Samples of studies involving results comparison between males vs. females
One sample z-test for proportions	Testing to see if a population proportion is different from a standard value	The rate of the population of the sample is the same as the standard rate.	Testing the proportion of men who are smokers and have lung cancer against the rate of all men with lung cancer
Two sample z-test for proportions	Testing for a difference between proportions from two populations	The proportion rates are the same between the two categories.	The proportion of men who smoke compared against the proportion of women who smoke
One sample t-test for mean	Testing to see if a population mean is different from a standard value	The sample population mean is the same as the mean of the standard.	Study of the average number of falls at Hospital A against the national benchmark value
Two sample t-test for mean	Testing to see if a population mean is different between two populations. Measurements from distinct populations	The means of two populations are the same.	The comparison of the average length of stay for a disease between men and women.
Paired t-test	Pre/post testing for a difference in means – same subjects in pre/post data	There is no difference between the pre- and post-test values.	The mean of a measurable parameter of a sample group as a control is compared to the mean of the parameter of the same sample group after a treatment or drug.
ANOVA	The analysis of variance between more than two populations	The means between the sample groups ``are equal	The comparison of the complication rate in adults from a surgery in all age groups
Pearson correlation	Linear correlation strength measure as "r" in rage from -1 to 1	There is no correlation between two samples.	Testing the correlation between age and the incidence of a disease

Source: White, S.E. *A Practical Approach to Analyzing Healthcare Data* (2nd ed.). Chicago, IL: AHIMA Press, 2013.

CHART TYPES FOR GRAPHICAL REPRESENTATION

Graphic visuals can summarize large amounts of data for greater understanding analysis of the distribution and trending. The type of graph chosen depends upon the variable type and the purpose of the analysis. The following table is a representation of some common types of graphs.

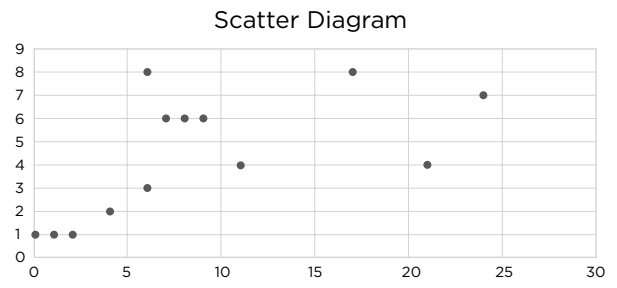
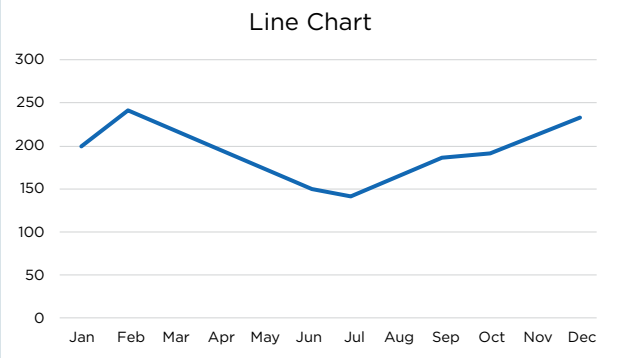
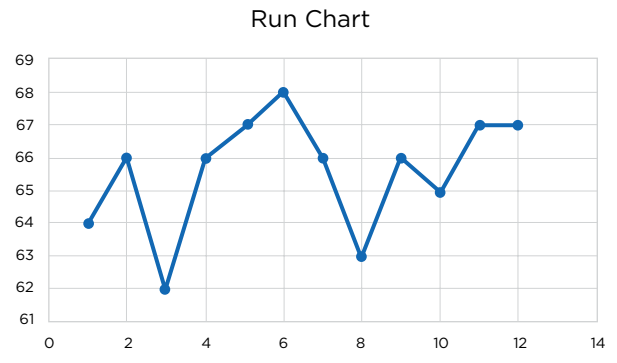
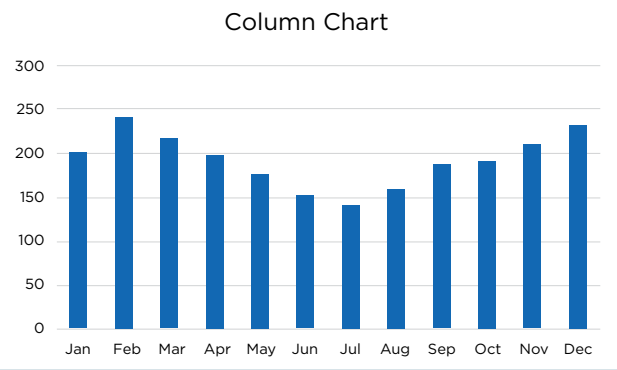
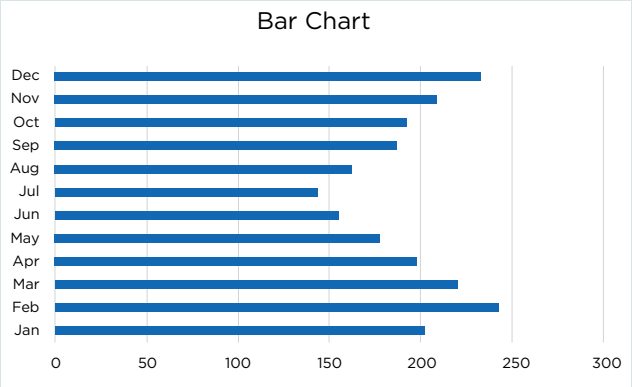
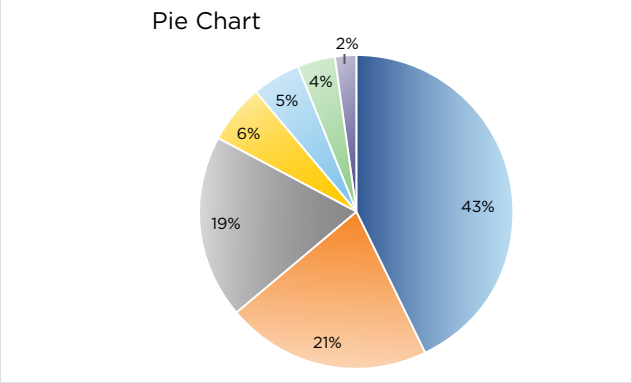
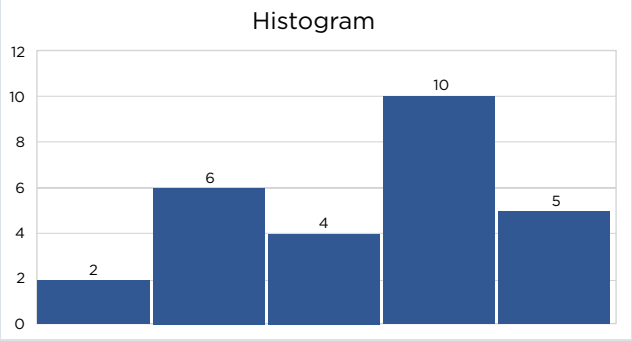
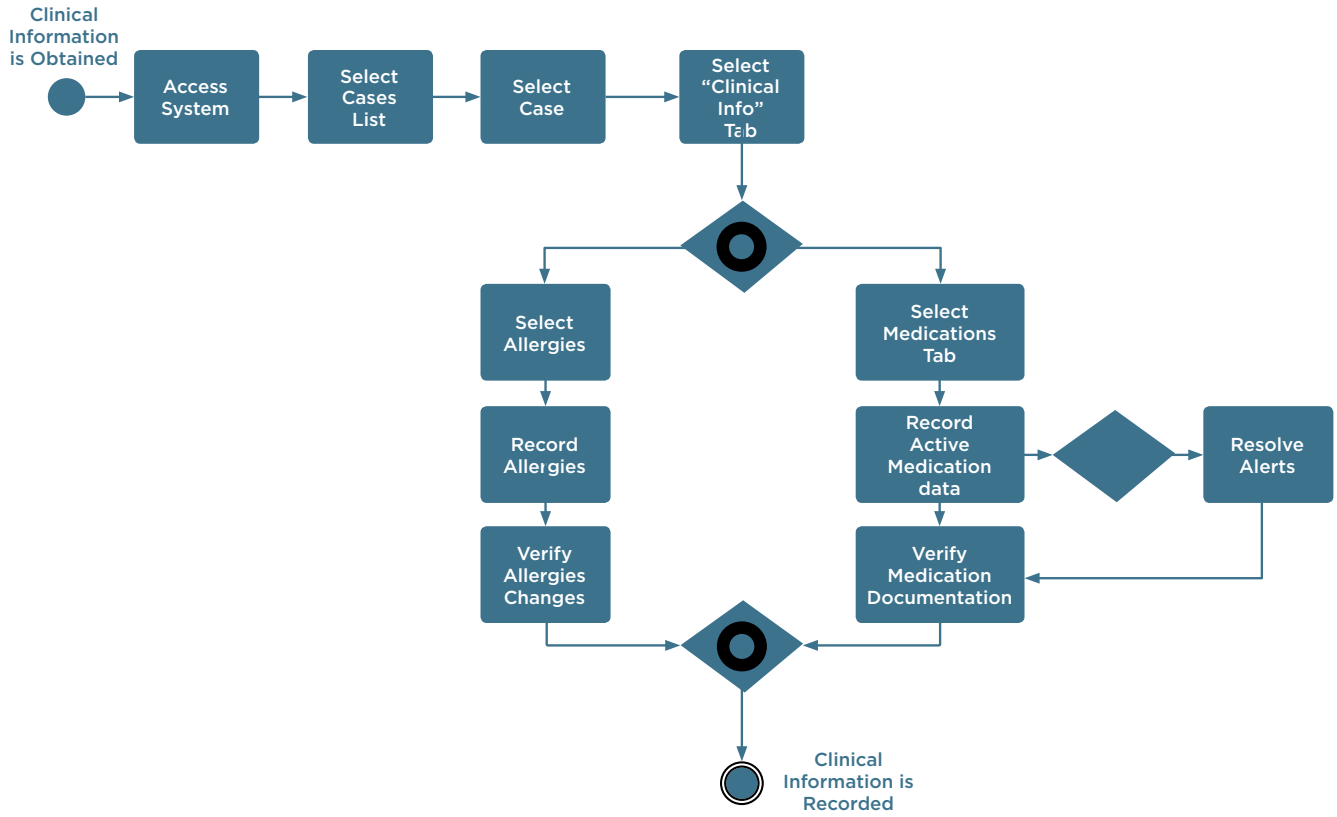
Chart Type	Best Use	Example of Use	Visual Example
Scatter Diagram (Scatter Plot)	To determine the correlation or distribution between two quantitative values.	Is the length of stay related to	
Line Graph	To show values over time.	Trend Analysis, Daily Census	
Run Chart	To determine the variation in a process using limits and a center line.	Statistical Process Control (SPS) Charts	
Column Chart	Vertical columns that show visual comparison; may be stacked to show part-to-total composition.	To show bed occupancy rate by the time component	

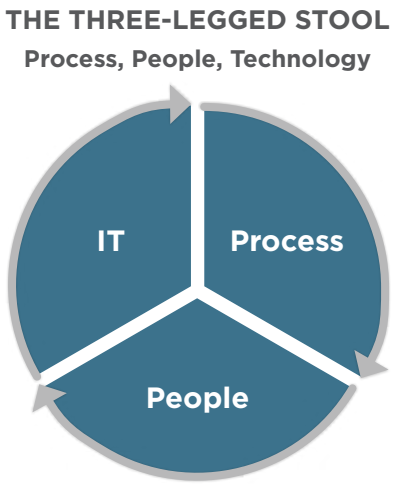
Chart Type	Best Use	Example of Use	Visual Example
Bar Chart	Horizontal rows that show visual comparison; may be stacked to show part-to-total composition.	To show bed occupancy rate	
Pie Chart	Comparison of the percentage of values in a category. Slices must add up to 100%.	To show the method of payment for all patients in percentage of each type.	
Histogram	To determine the distribution of variables. Frequency distribution.	Evaluation of Likert Scale responses by frequency of response.	

SAMPLE OF VISIO WORKFLOW DIAGRAM



THE PROJECT MANAGEMENT "THREE-LEGGED STOOL"

The three-legged stool refers to a basic project management philosophy. As shown in the graphic, if one of the legs of the stools gets longer (processes are delayed, more people are needed, the IT costs rise, etc.), then the other two legs usually get longer as a result. Keeping all three components on time and budget are essential to ensure that the project meets its goals.



The components of a project charter can vary from project to project. However, the most common ones are:[1]

- Date of project charter creation
- Date of revision
- Project manager's name
- Project name
- Type or commodity (product or service)
- Project statement (why we are doing this project)
- Current state
- Desired state (what success looks like)
- Contract expiration (if applicable)
- Diversity supplier potential opportunity
- Stakeholders (both negatively and positively affected by the project)
- Saving opportunity (quantify it)
- Describe current process or metrics (what's going wrong)
- Describe future process or metrics (what the goal is, mention numbers)
- Customers (list who the customers are, internal, external, all of them)
- Customer requirements (what do they expect, remember we do this for them)
- Risks (every project has a risk, quantifiable or not)
- Estimated project expense (travelling expenses, subcontracting, etc)
- Team members
- Executive sponsors (who will approve the charter and modifications to the charter)
- Steering committee members (some names may cross with the stakeholders list)
- Project timeline: milestone (phase 0, phase 1, etc.), deliverable (example: prepare contract), status (on time, past due, etc.)
- Approval signatures by phase

1. Schafer, Anna. "Components of a Project Charter." The Project Management Hut. March 20, 2013.

APPENDIX E:
SAMPLE PROJECT CHARTER CONTENT SAMPLE PLAN

SAMPLE PROJECT CHARTER CONTENT

Project Name

Prepared By

Document Owner(s)	Project/Organization Role
(Author)	Project Manager
	Project Sponsor

Project Charter Version Control

Version	Date	Author	Change Description

Background and Purpose

Project Executive Summary

Project Scope

- Goals and Objectives
- Project Deliverables
- Deliverables Out of Scope
- Project Estimated Costs and Duration

Project Conditions

- Project Assumptions
- Priority Criteria
- Project Risks
- Project Constraints

Project Structure Approach

Project Team Organizational Plans

Approvals

Approvals	Role	Date
Prepared by:	Project Manager	05/30/20XX
Approved by:_____	Project Sponsor/Executive sponsor	

APPENDIX F:
SAMPLE INTERNAL BUSINESS USE REQUIREMENTS

A sample form can be found here.

APPENDIX G: STATISTICAL TEST TABLE

A sample statistical test table can be found here.

APPENDIX H:
HIPAA PHI IDENTIFIERS

18 PHI IDENTIFIERS PER HIPAA:

1. Names
2. All geographic subdivisions smaller than a state, including street address, city, county, precinct, zip code, and their equivalent geocodes, except for the initial three digits of the zip code if, according to the current publicly available data from the Bureau of the Census:
   The geographic unit formed by combining all zip codes with the same three initial digits contains more than 20,000 people; and
   The initial three digits of a zip code for all such geographic units containing 20,000 or fewer people is changed to 000
3. All elements of dates (except year) for dates that are directly related to an individual, including birth date, admission date, discharge date, death date, and all ages over 89 and all elements of dates (including year) indicative of such age, except that such ages and elements may be aggregated into a single category of age 90 or older
4. Telephone numbers
5. Vehicle identifiers and serial numbers, including license plate numbers
6. Fax numbers
7. Device identifiers and serial numbers
8. E-mail addresses
9. Web Universal Resource Locators (URLs)
10. Social security numbers
11. Internet protocol (IP) addresses
12. Medical record numbers
13. Biometric identifiers, including finger and voice prints
14. Health plan beneficiary numbers
15. Full-face photographs and any comparable images
16. Account numbers
17. Any other unique identifying number, characteristic, or code, except as permitted by paragraph (c) of this section [Paragraph (c) is presented below in the section "Re-identification"]
18. Certificate/license numbers

APPENDIX I: CONCEPTS OF HIPAA'S SECURITY RULE

- Administrative safeguards (https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/administrative/securityrule/adminsafeguards.pdf)
  - » Security management process
  - » Assigned security responsibility
  - » Workforce security
  - » Information access management
  - » Security awareness and training
  - » Security incident procedures
  - » Contingency plan—evaluation
  - » Business associate contracts and other arrangements administrative safeguards

- Physical safeguards (https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/administrative/securityrule/physsafeguards.pdf)
  - » Facility access controls
  - » Workstation use
  - » Workstation security
  - » Device and media controls

- Technical safeguards (https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/administrative/securityrule/techsafeguards.pdf)
  - » Access control
  - » Audit controls
  - » Integrity
  - » Person or entity authentication
  - » Transmission security
- Organizational, policies and procedures and documentation requirements (https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/administrative/securityrule/pprequirements.pdf)
  - » Business associate contracts or other arrangements
  - » Requirements for group health plans