A patient data can be defined as: ................................................................. 4
Types of Medical data: .................................................................................. 4
Uses of Medical Data ...................................................................................... 4
Weaknesses of the traditional medical record system: .................................. 5
Strategies of Medical data selection and use ................................................. 7
The Hypothetico deductive approach ............................................................ 7
Standard Development Process .................................................................... 8
Standards Development Organizations ........................................................... 9
Data Interchange Standards ........................................................................... 11
Data Interchange Standards ........................................................................... 12
Directions For Standards ............................................................................. 13
Patrick W. O’Carroll ...................................................................................... 14
T. Benson, Principles of Health Interoperability HL7 and SNOMED, Chapter 2: Why Interoperability is Hard .............................................................. 17
Principles of Health Interoperability HL7 and SNOMED UML and XML .......... 20
Law #1: Interoperability: it’s all about the people ........................................... 34
Law #2: Complexity: you can move it around, or externalize it, but you can’t make it go away ................................................................. 35
Law #3: Healthcare Software: Cheap, Flexible, Interoperable – you can have any two ........ 35
Driving Quality—A Health IT Assessment Framework for Measurement ............ 36
Session 2: JSON vs XML ............................................................................. 49
Simplicity ...................................................................................................... 49
Extensibility .................................................................................................. 49
Interoperability ........................................................................................... 49
Openness ......................................................................................................................................................... 49
Session 3: Desiderata for Controlled Medical Vocabularies in the Twenty-First Century.................................................. 50
Session 3 Development of the Logical Observation Identifier Names and Codes (LOINC) Vocabulary .... 53
Session 3: Cross-mapping clinical notes between hospitals: An application of the LOINC Document Ontology........................................................................................................................................... 57
Session 3 LOINC® - A Universal Catalog of Individual Clinical Observations and Uniform Representation of Enumerated Collections........................................................................................................................................ 59
Session 3: SNOMED CT® User Guide........................................................................................................................................... 62
Session 3: Normalized names for clinical drugs: RxNorm at 6 years .................................................................................. 67
Session 4: The Unified Medical Language System: An Informatics Research Collaboration.................................................... 70
Session 4: LexGrid: A Framework for Representing, Storing, and Querying Biomedical Terminologies from Simple to Sublime ........................................................................................................................................ 72
Session 4: The National Center for Biomedical Ontology ........................................................................................................ 76
Session 4: PHIN VADS - CDC Vocabulary Server .................................................................................................................. 78
PHIN VADS Series Lectures Given by CDC................................................................................................................ 78
Session 4: PHIN VOCABULARY ACCESS AND DISTRIBUTION SYSTEM (VADS) ......................................................... 79
Session 5: Chapter 6 HL7 Version 2 ............................................................................................................................................ 79
Session 5: Health Level Seven Version 2.x Message Profiling Specification Version 2.2 ......................................................... 86
  What is an HL7 V2.x Message Profile? ......................................................................................................................... 87
  HL7 V2.x Message Profile Components ....................................................................................................................... 87
    Use Case Model ..................................................................................................................................................... 87
    Static Definition of an HL7 V2.x Message Profile ..................................................................................................... 89
    Dynamic Definition of an HL7 V2.x Message Profile ............................................................................................ 93
Identification of HL7 Message Profiles Using ASN.1 ........................................................................................................ 95
  Static Profile Identifier ........................................................................................................................................... 95
  Dynamic Profile Identifier .................................................................................................................................... 96
Session 5 : Critical Factors Influencing Hospitals’ Adoption of HL7 Version 2 Standards: An Empirical Investigation .............................................................................................................................................. 97
Session 5: Data Interchange Standards in Healthcare IT—Computable Semantic Interoperability: Now Possible but Still Difficult, Do We Really Need a Better Mousetrap? ............................................................................ 98
  Context of Interoperability ....................................................................................................................................... 102
    Multimedia: Health Level 7 Version 2 .................................................................................................................. 103
    Multimedia: Health Level 7 Version 3 .................................................................................................................. 105
Session 6: Enabling Joint Commission Medication Reconciliation Objectives with the HL7 / ASTM Continuity of Care Document Standard

Session 6: The Clinical Document Architecture and the Continuity of Care Record: A Critical Analysis


Session 6: HITSP Summary Documents Using HL7 Continuity of Care Document (CCD) Component

Session 6: HHIITTSPP CCDDAA CCoonntteenntt MMMoodduulleess CCooommpoonneennt

Session 7: The “Meaningful Use” Regulation for Electronic Health Records

  Meaningful Use Stage 1

  Meaningful Use Stage 2

Session 7: Semantic integration in healthcare networks

Session 7: The Person Centered Coordination Plan

Session 7: Health Information Technology Automation of Quality Measurement: Quality Data Set and Data Flow

Session 7: Semantic Interoperability for Better Health and Safer Healthcare

Session 8: Health information exchange and patient safety

Session 8: Health information exchange policy and evaluation

Session 8: Health-information exchange: why are we doing it, and what are we doing?

Session 8: Protection Detail Protecting Against Breach of Electronic Protected Health Information

Session 8: Qualitative evaluation of health information exchange efforts

Session 8: The financial impact of health information exchange on emergency department care

Session 8: The Next Step in Health Data Exchanges Trust and Privacy in Exchange Networks

Session 8: The United Hospital Fund Meeting on Evaluating Health Information Exchange

Session 8: Using Health Information Exchange to Improve Public Health

Session 8: Case Study: Patient Matching Sharp HealthCare’s Journey
**Session 1: Chapter 2: Biomedical Data: Their Acquisition, Storage, and Use**

- Page 2.46
- All Medical care activities involve gathering, analyzing, or using data
- Medical datum: Any single observation of a patient
- Medical data are multiple observations

**A patient data can be defined as:**
1) The patient in question
2) The parameter being observed
3) The value of the parameter in question
4) The time of the observation

- Page 2.48
- The circumstances under which the data is observed
- Aka modifiers
- Uncertainty in the values of data
- Collect additional data to confirm or reject the concern

- Page 2.49

**Types of Medical data:**
1) Narrative data: Some narrative data are loosely coded with short hand conventions
2) Complete phrases have become loose standards of communication among medical personnel
3) Numerical measurements: precision becomes important
4) Visual images

- Page 2.53
- Who collects data:
  1) Doctors
  2) Nurses
  3) Office Staff
  4) Specialists
  5) Technological Devices

- Page 2.54
- Use of medical data:
  1) To support the proper care of patient
  2) Analysis of population

- Page 2.55

**Uses of Medical Data**

- **A) Create the basis for the Historical record:**
  1) Patient’s history
  2) Symptoms reported
  3) Physical signs during PE
4) How have signs and symptoms changed?
5) What tests performed?
6) What medications taken?
7) Reasons behind the management decisions

- To answer 3 questions:
  1) What was the nature of the disease
  2) What was the treatment decision
  3) What was the outcome of the treatment

- Central function to provide coordinated care to a patient over time

- Page 2.56
  - B) Support communication among providers
    - Importance of the central role of the medical records
    - Serves as communication mechanism among physicians and other medical personnel

- Page 2.57
  - C) Anticipate Future Health problems
  - D) Record Standard preventive measures
    - For interventions to prevent common or serious disorders
  - E) Identify deviations from expected trends
  - F) Provide a legal record
  - G) Support clinical research
    - Randomized clinical trial (RCT)

- Page 2.60

**Weaknesses of the traditional medical record system:**

1) **Pragmatic and logistical issues:**
   a. Can I find the data I need when I need them?
   b. Can I find the medical record in which they are recorded?
   c. Can I find the data in the record?
   d. Can I read and interpret the data
   e. Can I find what I need quickly
   f. Can I update the data reliably

- Page 2.62

2) **Redundancy and Inefficiency**
   a. MD developed a variety of techniques that provide redundant recording to match alternate modes of access
   b. Inefficiency from tension between opposing goals in the design of reporting forms

- Page 2.63

3) **Influence Clinical research**
   - Hard to flip through records for structured statistical analysis
   - Page 2.64
   - Retrospective Chart review vs prospective studies to collect future data relevant to the question
Double blind studies: Patients and researchers do not know which treatment is being administered

4) The passive nature of Paper records

- No warnings

- EHR system:
  - Monitor their contents and generate warnings or advice for providers
  - Provide automated quality control
  - Provide feedback on deviation from standards

Structure of Medical Data

- Need standardized Nomenclature:
- Use of coding systems:
  1) ICD: International Classification of Disease: Used for discharge coding and bills to insurance
  2) Systemized Nomenclature of Pathology (SNOP)
  3) Systemized Nomenclature of Medicine (SNOMED)
  4) Current Procedural Terminology (CPT): By American Medical Association to produce bills for services to patients

Health care professionals need standardized terms that can support pooling of data for analysis and can provide criteria for determining changes for individual patients

- None will be completely satisfactory
- Some want more specific
- Some want more aggregation

Develop a unified medical language system (UMLS): ties together various vocabulary

- Central focus: Information base that constitutes the substance of medicine

- 3 terms:
  1) Data: A single observation point that characterizes a relationship
  2) Knowledge: Derived through the formal or informal analysis of data
  3) Information: Encompass both organized data and knowledge

- Data are not information until they have been organized in some way for analysis or display
- Heuristics: A personal piece of knowledge that guides physicians in their decision making

- A database: A collection of individual observations without any summarizing analysis
- A knowledge base is a collection of facts, heuristics, and models that can be used for problem solving

Many decision support systems have been called knowledge base systems
Strategies of Medical data selection and use

1) Learn how to ask only the questions that are necessary to perform only the examination components that are required
2) And to record only those data that will be pertinent in justifying the ongoing diagnostic approach and in guiding the future management of the patient
3) Selectivity in data collection and recording

The Hypothetico deductive approach

1) Sequential staged data collection
2) Data collection
3) Hypothesis directed selection of the next appropriate data to be collected
4) As data are collected, they are added to the growing database of observations
5) Use to reformulate / refine the active hypotheses
6) Until one hypothesis reached a threshold of certainty
7) A management decision can be made

Differential Diagnosis: Set of active hypotheses
- The selection process is heuristic
- Safety measures to avoid missing important issues:
  1) Past medical history
  2) Family history
  3) Social history
  4) Review of systems

Focus hypothesis directed examination is augmented with general tests
- After asking questions, serves as the basis for a focused PE
- Treatment management plan can be developed
- Treatment itself is a datum point

Chronic disease management: a cycle of treatment and observation for a long time

Need to balance cost and risks

Sensitivity: The likelihood that a given symptom can be observed in a patient with a given disease
Pathognomonic tests: they evoke a specific diagnosis but they also immediately prove it to be true
Specificity: An observation is not seen in patients who do not have the disease
Prevalence of a disease: A measure of the frequency with which the disease can be found in a population
Predictive Value (PV): The post test probability that a disease is present based in the results of a test
- \[ PV^+ = \frac{(\text{sensitivity}) \times (\text{prevalence})}{((\text{sensitivity}) \times (\text{prevalence}) + (1 - \text{specificity}) \times (1 - \text{prevalence}))} \]
- Post test probability is low if the prevalence of that disease is low
- Test is pathognomonic: When specificity is 100%

- PV+: One of many forms of Bayes Theorem: A rule for combining probabilistic data attributed to Thomas Bayes

- Page 2.75
- Method for selecting questions and comparing test
- Page 2.76
- Computer and collection of data:
  - By MD
  - By Paid transcriptionist
  - By Medical staff
  - By device itself
  - By patients

---

**Session 1: Chapter 7: Standards in Biomedical informatics**

- Page 7.265
- Standards: A set of rules that specify how to carry out a process or produce a product

- Page 7.266
- The need for Health Informatics standards
- Little coordination and sharing of patient data
- Standardize Identifier for individuals, health care providers, health plan, and employers
- Mechanism for issuing identifiers:
  - CMS: National Provider Identifier (NPI)
    - Payer ID for Health Care plans
    - IRS Employer ID
- Person Identifier: Invasion of privacy
- Need Encoding of clinical knowledge using accepted standards
- Methods needed to transfer information from one to another

**Standard Development Process**

1) Ad Hoc Method
2) De Facto Method: Large vendor product make standard
3) Government Mandate method
4) Consensus Method: Group works in an open process to create a standard. E.g. HL7

- Page 7.269
- Process:
  1) Identification stage: Need
  2) Conceptualization Stage: Group formed to discuss
     - What must the standard do?
What is the scope?
What will be its format?

3) Discussion stage
4) Writing for draft standard
   - Page 7.271
   - Open policy
   - Open balloting policy
     - Negative ballots address
5) Early Implementation
   - Acceptance and rate of implementation are important
   - Need
     a. Conformance
     b. Certification

Standards Development Organizations
1) ANSI: American National Standards Institute
2) European Committee for Standardization Technical Committee 251 (CEN TC251)
   - 4 working groups
     a. Information Models
     b. Terminology
     c. Security, Safety and quality
     d. Technology for interoperability
3) International Standards Organization Technical Committee 215
   - Page 7.273
   - Health Informatics
   - 6 working groups
     a. Health records and modeling coordination
     b. Messaging and Communication
     c. Health concept Representation
     d. Security
     e. Smart Cards
     f. ePharmacy and Medicine

- Page 7.272
4) American Society for Testing and Materials (ASTM)

- Page 7.275
5) Health Care Informatics Standard Board
- Health Care Informatics Standard Planning Panel (HISPP)
- Identify Standards for:
  a. Health care models and EHR
  b. Interchange of health care data between organizations
  c. Health care codes and terminology
  d. Communication with diagnostic instruments
  e. Communication of HC protocols
  f. Privacy
g. Other Areas

- Page 7.276
  6) Health Care Information and Management Systems Society
  7) Computer Based Patient Record Institute (CPRI)
  8) Integrating the Healthcare Enterprise (IHE)
     o Stimulate the integration of HC Information resources
  9) National Quality Form (NQF)
 10) National Institute of Standards and Technology (NIST)
 11) Workgroup for electronic data Interchange (WEDI)
     a. Implementation of EDI

- Page 7.278
  - Health Insurance Portability and Accountability Act of 1996 (HIPAA)

- Page 7.279
  - Coded Terminologies, Nomenclature
  - Purpose
     1) Save system developers from reinventing the wheel
     2) Facilitate exchange of data among systems

- To discuss coding systems:
  1) Need to clarify the differences among a terminology, a vocabulary, and a nomenclature
  2) Determine the basic use of the terminology: 2 Levels:
     a. Abstraction: Examinations of the recorded data and then selection of items from a terminology to label the data
     b. Representation: process by which as much detail as possible is coded

- Page 7.280
  - Must consider
     1) Domain of Discourse: Must be good match with any standard selected for the purpose
     2) Content of the Standard
     3) The methods by which the terminology maintained method

- Specific Terminologies
  1) International Classifications of Disease (ICD)
     o Revised every 10 year
     o ICD9 inadequate
     o US published a set of clinical modifications (CM)
     o ICD9-CM
  2) DRG: Use in prospective payment in Medicare Program
      - Provide a small number of codes for patients hospitalization
      - DRG = Diagnosis related Groups
  3) International Classification of Primary Care (ICPC)
  - Post coordination of atomic terms
     o Coding through the use of multiple codes as needed to describe the data
  - Pre-coordination: Every type of pneumonia is assigned its own code
5) Diagnostic and statistical Manual of mental Disorder
6) Read Clinical Codes
7) SNOMED
   - Clinical Terms and its predecessors
     o SNOMED Clinical Terms (SNOMED CT)
8) Galen: Reference model for medical concepts using structures Meta Knowledge (SMK)
   o Terms are defined through relationships to other terms
   o Grammars to combine terms into sensible phrases
   o Page 7.288
   o To allow representation of information independent of the language being recorded and of the data model used by EHR
9) Logical Observations, Identifiers, Names, and Codes (LOINC): For Tests and Observations
10) Nursing Terminologies
    - North American Nursing Diagnosis Association (NANDA) codes
    - Nursing Intervention Classification (NIC)
11) Drug Codes
    - Drugs are classified according to the Anatomical Therapeutic Chemical (ATC)
    - Page 7.289
    - The National Drug Code by FDA
      o Collaboration to produce a representational model for drug terms called RxNorm
    - Page 7.291
12) Medical Subject headings (MeSH): Terminology by which the world medical literature is indexed
13) Bioinformatics Terminologies
    o Gene Ontology (GO)
    o Develop standard ways of using these terminologies to encode data
    o Distributed Annotation System
    o Minimal Information about a Microarray experiment (MIME)
    - Page 7.292
14) Unified Medical Language System (UMLS)
    - Metasaurus: Contains over 100 different sources and relate synonyms across different sources
15) Interchange Registration of coding system
    o Health Care Coding Scheme Designator (HCD)
    - Page 7.292

Data Interchange Standards
- By American Association for Medical Systems and Informatics (AAMSI)
- Developed standards for HL7 and IEEE Medical Data Interchange standard
- Page 7.296
- Purpose to permit one system, the sender, to transmit to another system, the receiver, all the data required to accomplish a specific communication or transaction set in a precise, unambiguous fashion
- Communication model: Open System Interconnection (OSI) reference model (ISO 7498-1) by ISO
- 7 Levels of requirements
1) Physical
2) Data Link
3) Network
4) Transport
5) Session
6) Presentation
   o Deals with how data are formatted
   o 2 philosophies
     a. Position dependent format
        ▪ Data content is specified and defined by position
     b. Tagged field format
        ▪ E.g SEX-M
7) Application
   o Deals with semantics of data content specification of a transaction set

- Page 7.297
- A transaction set is defined for a particular event called a trigger event

**Data Interchange Standards**

1) Digital Imaging and Communications in Medicine (DICOM)
   o For exchange of radiographic images
   o Includes
     ▪ Hardware Interface
     ▪ Data Dictionary
     ▪ A set of commands
   o Data elements are organized within the data dictionary into related groups
   o Groups and elements are numbered
   o Each data element consists of its group element tag, its length, and its value

- Page 7.298
2) ASTM International
   - E1238: Standard Specification for transferring Clinical observations between independent systems
   - Use position defined syntax
   - E1467

- Page 7.299
3) HL7: Reflect the application (7th) Level of OSI Reference model
   - Page 7.301
   - Most widely implemented HC data messaging standard
   - Message Based
   - Use Event trigger model version 3.0
   - Object oriented and Based on a reference Information Model (RIM)
   - A collection of
     a. Subject Area
     b. Scenarios
c. Classes
d. Attributes
e. Use Cases
f. Actions
g. Trigger Events
h. Interactions

- To Provide a model for the creation of message specifications and messages for HL7

4) IEEE – MEDIX
- Standard of Medical Device Communications
- For medical Information Bus (MIB)
- For Bedside device in ICU, ER

5) National Council for Prescription Drug Program (NCPDP)
- For Pharmacy services
- 3 Standards:
  a. Telecommunication Standard
  b. SCRIPT Standard
  c. Manufacturer Rebate Standard

6) ANSI X12
- For PO data
- Invoice data
- Claims, Benefits, Claims Payment
- Define Business Transactions in a formal Transaction sets

- Page 7.305

7) American Dental Association (ADA)
8) Uniform Code Council: Bar Codes
9) Health Industry Business Communication Council (HIBCC)
- Health Industry Bar Code (HIBC) standards
10) Electronic Data Interchange for Administration, Commerce, and Transport (EDIFACT)
- For Trade in goods and services

- Page 7.306

Directions For Standards
- Need standards for Queries
- Must support scripting and data entry mechanisms to ensure that a data system can properly and accurately provide a response
- Require high speed query and response
- High computational speed
- Need To:
  1) Negotiation among vendors interface systems
  2) How closely the vendor’s implementation stick to the standard
  3) Need certification
  4) Lower standards change

- Page 7.309
Future Directions:
1) Crossing the Quality Chasm
2) Consolidated Health Informatics (CHI) to establish a portfolio of existing clinical terminologies and messaging standards to build interoperable federal health data systems

- Develop National health care Informatics Infrastructure to
  1) Improve Patient safety and quality
  2) Detect Bioterrorism
  3) Enhance the efficiency of HC System

- Problem: Conflict between a standard and the opportunity for a vendor to use creativity in a product to enhance sales
- Standards should encourage Creativity
  o Use of Mouse
  o Use of visual objects
  o Use of Icons

- Standards need to include all types of data representation
- Specifying the location of data
- Rules for the creation of data
- Tighter coupling of data

Session 1: Public Health Informatics and Information Systems: Chapter 5: Information Architecture: Patrick W. O’Carroll

Overview

- Implementation of the concept of information architecture provides a solution to the problems created by piecemeal systems development
- Information architecture is a metaphor for a systematic, planned approach to building enterprise wide information systems
- Benefits:
  o Enhanced system interoperability
  o Ease of support
  o efficiency
  o Reduced redundancy of data entry
- Information architectures can be developed through a process called information resource management (IRM) planning

Information Architecture

- Information architecture includes
  o databases
  o Applications
  o Standards
  o Procedures
  o Information use
  o Confidentiality policies
Benefits of an information Architecture

- Benefits of an information architecture
  - Provides a guiding plan across development projects
  - Promotes a component orientation to the development process (larger pieces of the system are built out of smaller units)
  - Simplifies systems by
    - decreasing redundancy of data entry and storage
    - providing a coherent approach to cross cutting systems like security and data backup
  - Promotes efficiency and interoperability through the incorporation of standards and through solving challenges once
  - Promotes planning and clarifies business processes
  - Returns the locus of control and decision making to the executive levels and take it away from the information technology community

- Public health executives can direct the development of information systems by means of the clear business specifications
- These specifications are developed through a variety of formal and informal processes, including a process called joint application design (JAD)
- The business views of the information architecture represent what processes need to be automated
- The information technology views of the architecture represent how these processes should be automated

Developing an information Architecture

Information Resource Management (IRM) and IRM Planning

- Information resource management (IRM) is a set of principles and practices by which an organization manages its information resources
- Central underlying ethic: Information is one of a public health organization's key strategic assets, along with financial resources and human resources
- IRM planning is the process by which an information architecture is developed
- Information Resource Management Planning:
  - Is business planning
  - Requires a high level review of the business – its goals, procedures, customers, organizational structure, etc
  - Requires ongoing executive level sponsorship. Failure of IRM planning is guaranteed without this element

Stakeholders in a Public Health Organization's IRM Planning

- Stakeholder is anyone or any group affected in some way by the actions of the organization

Steps in IRM Planning

- Steps
- Understand the business: To elucidate in concrete terms exactly what the public health agency does
- Simplify the business, through reorganization and the use of information technology
- Integrate the business: Data, software code, and technology can be shared across the agency

Information resource models: Views of the architecture

- The Zachman framework provides a model of how enterprise managers and their information technology departments can work together to design and change enterprises and the computer systems that support them
- Uses a 2 dimensional structure to describe the information architecture of an enterprise
  - the first dimension describes the roles involved in information systems design: Planner/owner, designer, builder, subcontractor
  - The second dimension specifies various attributes of the system: what (data), how (function), where (network)

<table>
<thead>
<tr>
<th>Objective/Scope (contextual)</th>
<th>Function</th>
<th>Network</th>
<th>People</th>
<th>Time</th>
<th>Motivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role: Planner</td>
<td>What</td>
<td>How</td>
<td>Who</td>
<td>When</td>
<td>Why</td>
</tr>
</tbody>
</table>
| List of things important in the business | List of Business Processes | List of Business Locations | List of Important Organizations | List of Events | List of Business Goal 
| Role: Designer               | Logical Data Model | System Architecture Model | Distributed Systems Architecture | Human Interface Architecture | Processing Structure |
| Role: Builder                | Technology Design Model | Technology Architecture | Presentation Architecture | Control Structure | Rule Design |
| Role: Programmer             | Data Definition | Program | Network Architecture | Security Architecture | Timing Definition |
| Role: User                   | Usable Data | Working Function | Usable Network | Functioning Organization | Implemented Schedule |

- Levels in this model:
The **enterprise (functional) model** describes what the enterprise does to meet its missions and objectives and describes what the enterprise needs to know to do it.

The **information model** identifies and defines the subjects about which the enterprise keeps information and also identifies the significant relationships between those entities.

The **application models** identify and define a set of applications that support the enterprise and information models.

The **distribution model** specifies the physical distribution of entities and applications of the models to physical locations.

The **technology model** specifies the blueprint for the development and integration of the information technology resources of the enterprise.

- Implementation and migration plan: lay out a stepwise process for moving from legacy systems to the new systems called for in the IRM plan.

---

**Session 1: T. Benson, Principles of Health Interoperability HL7 and SNOMED, Chapter 2: Why Interoperability is Hard.**


Principles of Health Interoperability HL7 and SNOMED Why Interoperability is Hard (Chapter 2)

**What is Interoperability?**

Interoperability is ability of two or more systems or components to exchange information and to use the information that has been exchanged.

The **HL7 EHR Interoperability Work Group** has developed a framework, which covers three different points of view (Gibbons et al. 2007):

- Technical interoperability
- Semantic interoperability
- Process interoperability

These concepts are interdependent, and all three are needed to deliver significant business benefits.

**Technical interoperability** moves data from system A to system B, neutralizing the effects of distance. It is domain-independent. It does not know or care about the meaning of what is exchanged.

**Semantic interoperability** ensures that system A and system B understand the data in the same way. It allows computers to understand, interpret, and use data without ambiguity. This is specific to domain and context and usually involves the use of codes and identifiers. Semantic interoperability is at the core of what we usually mean by healthcare interoperability.

**Process interoperability** coordinates work processes, enabling the business processes at the organizations that house system A and system B to work together.

**2.2 Benefits**
The benefits of joined-up health care, to provide the right information at the right time and place, are predicated on deploying and using standards that enable computer systems to exchange information in a way that is safe, secure, and reliable.

The benefits increase exponentially and more parties are involved, because the number of interfaces needed to connect N systems increases using the formula \((N^2 - N)/2\).

### 2.3 Need for a Lingua Franca

Problems begin because each computer system stores data internally in a different way. This means that to communicate, data has to be translated from one format or internal language into another. The solution is often achieved by translating to an intermediate *lingua franca* (such as a version of HL7) that is understood by each party.

There are two translations in any interchange; first from the native language of System A to the lingua franca; second, from the lingua franca to the native language of System B.

### 2.4 Electronic Health Records

![EHR Structure (CEN EN13606)](Fig. 2.3 CEN EN13606 EHR Structure)
Each clinical specialty has its own way of working. The grand vision of joined-up health care is predicated on the notion that patient records can be shared electronically between clinicians from different specialties.

First, we must recognize that it is difficult to share information between different computer applications even within the same specialty. This is because each computer application stores data in a different way and may use different internal codes.

Even within the same specialty, the information about an outpatient visit differs greatly from that about an elective surgical operation or inpatient discharge summary following an emergency admission.

Similarly, people sometimes talk about clinical laboratory reports as if they were homogeneous, but the content of each type of clinical laboratory report is quite different, as is the work done in each type of laboratory.

2.5 Analysis is Paramount

Much of the hard work involved in interoperability lies in teasing out the hundreds or thousands of different use cases.

One way to simplify the problem is to distinguish clearly between information that needs to be processed by computer and that which needs to be read and understood by human users. Computer processing is essential when data elements need to be identified, matched, retrieved, or counted. This type of information must be structured, complete, unambiguous, and validated. These are relatively few but important.

2.6 Complex Specifications Create Errors

The endeavor to be rigorous leads to errors caused by the sheer length and complexity of the specifications.

A different sort of problem arises when the domain experts (such as doctors, nurses, and managers) are unable to fully understand these specifications due to the complexity of language or simply the time it takes to read them.

Errors multiply according to the:

- Probability of misunderstanding any part of the specification,
- The length of specification: In a long specification, exactly the same idea may be presented in different ways in two places, but each may be understood differently
- Number of options permitted.
- Number of times different implementations to be made

Successful specifications avoid errors by limiting scope, being easy to understand, relatively short, and simple, with few if any options.

2.7 Users and Suppliers are both Guilty

Often, both users and suppliers genuinely believe that they are in full agreement until the moment when users try to use the final product. Problems lie on both sides.
2.8 Shared Meaning

Shared meaning between computers requires shared understanding between all of the human participants.

The challenge in interoperability is harder; it is to ensure understanding horizontally across business processes, which may be in different organizations (between domain expert and domain expert) and vertically within computer systems suppliers, between users and developers who speak different dialects (Fig. 2.4).

Session 1: Principles of Health Interoperability HL7 and SNOMED UML and XML

4.1 Unified Modeling Language

UML stands for Unified Modeling Language.

UML is now the standard modeling notation used throughout the IT industry.
UML is a specialized modeling language, not simply a notation for drawing diagrams. It includes a notation, which is used on diagrams, and a meta-model, which is of interest primarily to the developers of UML software tools.

UML makes a critical distinction between models and diagrams. A model is the sum of all the information held about a project in UML. Each diagram is a partial view of this model. When learning UML, it is a convenient simplification to regard a model as the sum of the diagrams. Each diagram shows a small part of the total design. The model is the sum total of all the specifications, comprising hundreds of diagrams and supporting text.

Fowler (Fowler 2004) identifies three main ways in which people use UML: sketch, blueprint, and programming language.

The distinction between a sketch and a blueprint is that sketches are incomplete and exploratory, while blueprints are complete and definitive. Serious modeling (blueprints rather than sketches) requires a specialized UML tool.

The step beyond blueprint is when programs are produced directly from the model. Here, UML becomes the source code for executable code.

UML allows beginners to do simple things simply, yet also supports highly complex applications, underpinned by a rigorous formal language.

UML is completely independent of the software used to implement computer applications and is not tied to any development methodology.

UML has a number of weaknesses

- Models and diagrams created using different tools cannot be imported and exported into and out of different tools reliably
- It does not have a neat way of specifying multiple choices, decision tables or other constraints
- A premise of UML is that no single diagram (or type of diagram) can provide, on its own, a full representation of what goes on, and so we need to use sets of related diagrams. Each type of diagram only shows certain aspects of a situation – everything else is ignored. This simplification provides both the power (it makes the situation understandable) and the weakness of diagrams (each diagram has a limited scope).

4.1.1 UML Diagrams

UML diagrams relate either to information structure or to behavior

One danger point is UML’s principle of suppressing information. This allows information to be omitted from any diagram in order to make it easier to understand. The corollary is that you should never infer anything from the absence of information in a diagram and that UML diagrams should not be read on their own without access to the rest of the model.
Some guidelines apply to all types of diagrams, such as:
• Diagrams should be laid out so that they can be read left-to-right and top-to-bottom.
• Avoid crossed, diagonal, and curved lines.
• Document diagrams using notes.
• Use the parts of UML that are widely understood, not the esoteric parts.
• Use color coding with discretion.
• Use common naming conventions such as UpperCamelCase (e.g., ClassName) and lowerCamelCase (e.g., attributeName).
• Do not put too much on a single diagram. Restrict diagram size to a single sheet of A4.
• Use consistent legible fonts.
• Show only what you need to show. It is good practice to suppress unnecessary detail.

4.2 Structure Diagrams

4.2.1 Class Diagrams

Class Diagrams are the most widely used UML diagrams. Class diagrams show the static structure of classes, their definitions, and relationships between classes. A class is a description of a group of objects with properties (attributes), behavior (operations), relationships to other objects (associations and aggregations), and semantics. Classes are shown as rectangles with one, two, or three compartments.
The top compartment shows the class name, the second shows attributes. Attributes describe the characteristics of the objects, while operations are used to manipulate the attributes and to perform other actions. Attributes and operations need not be shown on a particular diagram.

Class names should be a singular noun and the class name is conventionally written using UpperCamelCase notation (e.g., ClassName). This class has two attributes, documented and date, and one operation – create().

Figure 4.2 shows a simple class diagram representing a prescription. Each prescription has a prescriber (author) and relates to a single patient. It has one or more prescription lines. Each prescription line includes details of a drug and may have any number (zero to many) of dosage instructions. The arrowheads on the lines (associations) show navigation. The arrow from Prescription to Patient shows that, in this model, one goes from Prescription to Patient but not the other way round.

The notation for multiplicity, used in associations and attributes, is:
• [0..1] optional, no more than one is allowed
• [*] or [0..*] optional, any number of instances is allowed
• [1] mandatory always one required (this is the default)
• [1..*] mandatory to have at least one instance

![Simple class diagram](image)

If no multiplicity is shown, the default assumption is multiplicity of 1, meaning that exactly one is required, although caution must always be observed when inferring anything from the absence of data on diagrams.

Figure 4.3 illustrates two notations used for showing containment. The black diamond indicates composition between Document and Line. Each Document contains one or more Lines, but Line cannot exist independently of Document. The hollow diamond indicates aggregation between Document and Author. Each Document has one Author, but the Author can exist independently of Document. The multiplicities used in any diagram depend on its purpose. Figure 4.3 (above) is a Document-centric showing Document has just one Author (a one-to-one relationship), but an Author-centric diagram, Fig. 4.4 (below) shows that Author has one or more Documents (a one-to-many relationship). Both are right, it just depends on the purpose of the model.
The concepts of inheritance and attributes are illustrated in Fig. 4.5. Patient and Prescriber are both specializations of Person: Person is a generalization of Patient and Prescriber. The triangle arrowheads indicate that both Prescriber and Patient classes inherit the properties of Person. Patient has attributes: nhsNo, dateOfBirth, and gender, but also inherits the attributes name and address from Person. Similarly Prescriber, has attributes professionalID and organizationID, as well as the properties of Person. Attributes have several properties. For example, the notation + dateOfBirth: Date[0..1] indicates:
• Visibility (+) is public, meaning that it is fully accessible.
• Attribute name is “dateOfBirth.”
• The attribute type is “Date.”
• Multiplicity is [0..1] meaning that this attribute is optional with a maximum number of occurrences of one.
• Initial values and defaults may also be specified.

Attribute names are usually written in lowerCamelCase (e.g., attributeName). Operations implement the functionality of a software object. They are the actions that an object knows how to carry out. The syntax for operations includes:
• Visibility
• Name
• Parameter list (in parenthesis)
• Return type
• Property string

An object is a unique instance of a class. An object diagram, such as Fig. 4.6, shows the relationships between objects. Each object may have:
• Identity (name)
• State (attributes)
• Behavior (methods)

The object name is underlined (to distinguish it from a class) and comprises the object’s name, which is optional, followed by a colon and the class name (e.g., TimBenson:Author).

### 4.2.2 Package

Packages are used to divide up a model in a hierarchical way. Each package may be thought of as a separate name space. Each UML element may be allocated to a single package. Packages provide a useful means of organizing the model. Classes that are closely related by inheritance or composition should usually be placed in the same package. In Fig. 4.7, the Party package might include all classes related to people and organizations, including patients, doctors, and nurses. The Interaction package might include messages and entries in clinical records. The dashed arrow from Interaction to Party indicates that Interaction has a dependency on Party (Interactions involve Parties).

![Package diagram](image)

**Fig. 4.7** Package diagram

### 4.2.3 Deployment

The physical organization of computer systems is shown in deployment diagrams (Fig. 4.8). Each piece of the system is referred to as a node. The location of software can be shown as components.

![Deployment diagram](image)

**Fig. 4.8** Deployment diagram

### 4.3 Modeling Behavior

#### 4.3.1 Use Case

Use cases capture the behavioral requirements of business processes and provide a common linkage across all aspects of a project from initial analysis of requirements right through development, testing, and final customer acceptance. They show how people will ultimately use the system being designed. Each use case describes a specific way of using the system. Any real system has many use cases. Each use case constitutes a complete course of events, initiated by an actor (or trigger). A use case is essentially a special sequence of related transactions performed by an actor and the system in a dialogue. An actor is an external party, such as a person, a computer, or a device, which interacts with the system. Each actor performs one or more use cases in the system. By going through all of the actors and defining everything they are able to do with the system, the
complete functionality of the system is defined. Each use case is a description of how a system can be used (from an external actor’s point of view); it shows the functionality of the system, yielding an observable result of value to a particular actor. A use case does something for an actor and represents a significant piece of functionality that is complete from beginning to end. The collected use cases specify all the ways the system can be used. Nontechnical personnel can understand use cases intuitively. Thus they can form a basis for communication and definition of the functional requirements of the system in collaboration with potential users.

A simple use case diagram is shown in Fig. 4.9. Stickmen represent actors; ellipses represent use cases. In this diagram, the principal Actors are shown on the left. The arrowheads indicate the actor that initiates the use case. Use cases are fundamentally a text form and should be documented using simple templates, such as:

- Metadata, such as use case name, unique ID, author, date, version, and status
- Scope and context
- Primary and other actors
- Preconditions and trigger event
- Main success scenario describing the normal flow of events using numbered steps from trigger through to post-conditions
- Post-conditions
- Alternative flows, e.g., when errors occur
- Importance and priority
- Open issues

A scenario is an instance of a use case. It is one path through the flow of events for the use case and can be documented using an activity diagram or a storyboard free text description (Fig. 4.10).
4.3.2 Activity Diagram

Use activity diagrams to show the important business processes undertaken by each role, such as validation and database update. Each role may be shown in a separate “Swim Lane.” Transactions are communications that cross swim lanes.

Activity diagrams display a sequence of actions (including alternative execution paths) and the objects involved in performing the work. They are particularly useful for describing workflow and behavior that have branches and forks. Figure 4.11 shows a simplified activity diagram for the exchange of a referral and clinic letter between GP and hospital. It is organized in swim lanes to show who or what is responsible for each activity.
Activity diagrams can be used to show logical data flows. A branch has a single entry point, but a choice of exits depending on some condition. Only one route can be taken. Branches end at a merge. A fork has one entry and multiple exits, which can be undertaken in parallel, and the order of activities is not important. A fork ends at a join.

**4.3.3 Sequence Diagram**

Sequence diagrams (e.g., Fig. 4.12) show how objects interact with each other. Sequence diagrams show when messages are sent and received. A Sequence Diagram is a diagram that depicts object interactions arranged in time sequence, where the direction of time is down the page. The objects, which exchange information, are shown at the top of a vertical line or bar, known as the object's lifeline. An arrow between the lifelines of two objects represents each message.
4.3.4 Statechart Diagram
An object state is determined by its attribute values and links to other objects. A state is the result of previous activities of the object. A state is shown as rectangle with rounded corners. It may optionally have three compartments (like classes) for name, state variables, and activities (Fig. 4.13).

[Diagram showing state transitions]

A statechart diagram shows an object life cycle, and can be used to illustrate how events (messages, time, errors, and state changes) affect object states over time. State transitions are shown as arrows between states.

4.5 Business Process Modeling Notation (BPMN)
BPMN (Business Process Modeling Notation) is a notation for documenting complex business processes (White and Miers 2008). The BPMN notation is understandable by end users and it is also capable of including the technical detail needed to specify messages involved in web services delivery and the generation of XML-based Business Process Execution Language (BPEL).

BPMN is a standard for business process modeling with a notation that is similar to that used in UML Activity Diagrams. Some commentators regard the future of BPMN as a specialized “front end” to UML.

4.5.1 Activities
Activity is the generic term for Business Process, Process, Sub-Process, and Task. These have a hierarchical relationship. To use an analogy: a business process is a group of one or more trees; a process is a single tree; a subprocess is a branch (and may have further subbranches, sub-subbranches, and so on); task is a leaf, which is not subdivided further. Business Process is the top of the Activity hierarchy in BPMN. It is defined as a set of activities that are performed within an organization or across organizations, shown on a Business Process Diagram (BPD). Process is limited to the activities undertaken by one Participant (organization or role). Each Business Process may contain one or more Processes. A Process is an activity performed within an organization, and is depicted as a set of activities (Sub-Processes and Tasks) contained within a single Pool (see below Fig. 4.14).
Each Sub-Process may be expanded as a separate, linked diagram, showing its component Sub-Processes or Tasks. The facility to expand or consolidate Sub-Processes is a useful feature of BPMN. A Task is an atomic activity, showing that the work is not broken down to a finer level of detail. Sub-Processes and Tasks are shown as rounded rectangles. Sub-Processes, which can be expanded, are shown with a "plus sign" at the bottom center of the icon. Participants are each represented by a Pool, which may contain Lanes. Each Pool contains a single Process. A Pool may be subdivided into Lanes (like swim lanes in UML activity diagrams). Lanes may represent different roles within an organization. If a diagram contains a single Pool, the Pool boundaries need not be shown. A Pool is a container separating each Process from others and showing the Sequence Flow between activities. Shown as a small circular icon, an Event is something that happens during the course of a business process that affects the flow. Events may represent triggers for activities to begin or their outcomes. Start, Intermediate, or End events are indicated by the thickness of the circle perimeter. An additional icon inside the circle shows the type of Trigger or Result (Message, Timer, Error, Cancel, Compensation, Rule, Link, Multiple, or Terminate).

A Gateway, shown as a square diamond, is used to control branching, forking, merging, and joining of paths. An icon inside the diamond shows the type of control (exclusive XOR, inclusive OR, parallel AND, or complex). Connectors link the flow objects (Activity, Event, and Gateway). There are three types of Connectors:

- Sequence Flow (a solid line with arrow head) shows the order that activities are performed within a Process.
- Message Flow (a dotted line with arrow head) shows connections between Processes (crossing the boundary of a Pool).
- Association (dotted line, no arrow head) is used to associate information (such as Data Objects) and Annotations with Flow Objects.

4.5.2 Business Process Example

A complete business process from start to finish is shown in Fig. 4.15, which illustrates the traditional OP referral pattern for a patient suffering from a bowel problem. The Pools and Lanes show clearly who does what in what order. The dotted lines represent movement of information (messages) or of information sources (e.g., the patient). Each of the tasks shown could be represented as subprocesses and analyzed.
further in subsequent diagrams. Clinical care is essentially fractal and can usually be decomposed into smaller and more detailed subprocesses and tasks. Trigger events are shown as circles, with an icon indicating the type of trigger – an envelope indicates a message and a clock indicates a time trigger, such as an appointment slot.

4.6 XML

XML (eXtensible Markup Language) is a universal format for encoding documents and structured data, which is used in interoperability between different applications. XML documents are independent of the applications that create or use them (Bos 1999).
XML is strictly speaking a meta-language for formally describing a markup language.

4.6.1 Markup
Markup is a term that covers any means of making explicit an interpretation of a text. In electronic documents, including all word processors, the system does this by inserting special coded instructions into the text, which are not normally seen on a printed copy. A markup language specifies:
• What markup is allowed
• What is required
• How markup is to be distinguished from text
• What the markup means
XML provides a method of doing the first three; additional documentation is required for the last. Such specifications may be extensive. The real work is in the definition of message structures.

4.6.2 Descriptive and Procedural Markup
Most word processors include markup instructions embedded within their text.

XML is a descriptive markup scheme.

A descriptive scheme simply says what something is (for instance, a heading).

XML specifies how to render information using style sheets, coded using related style sheet languages called CSS (Cascading Style Sheets) and XSL (Extensible Stylesheet Language).

This separation of description (providing names for parts of a document) from procedure is the secret to platform-independence and universality, which are XML’s greatest strengths. A text marked up using a descriptive scheme can be processed in different ways by different pieces of software.

4.6.3 XML Files
XML and HTML are closely related. Indeed, XHTML is a version of HTML which is fully XML-compliant. Like HTML, XML makes use of tags (words bracketed by “<” and “>”) and attributes (of the form name = “value”). The key difference is HTML specifies what each tag and attribute means, and how the text between them will look on a browser, but XML uses the tags only to delimit pieces of data, and leaves the interpretation of the data to the application that reads them.

An XML document is said to be well-formed if it complies with a concise set of well-defined rules. One of the most important is that all nonempty elements are delimited by both a start-tag (e.g., <tag>) and a matching end-tag </tag>. Element names are case-sensitive.

4.6.4 XML Schema
The structure of an XML document is specified in a schema, which is also written in XML (van der Vlist 2001). The schema defines the structure of a type of document that is common to all documents of that type. It identifies the tags (elements) and the relationship among the data elements. This means that any document of a known type can be processed in a uniform way, including checks that all of the elements required are present and in the correct order.

Schema processing tools are used to validate XML messages or other XML documents using one or more schemas. Schema validation is applied to elements within a well-formed XML document. Three schema languages in widespread use are W3C’s Schema Definition Language (XSD), RELAX NG, and Schematron.

4.6.5 XML Element
In XML, documents are made up of elements. Each element is tagged using a start-tag and an end-tag. For example, a diagnosis element in a text might be tagged as follows: 
```
<diagnosis> Diabetes mellitus </diagnosis>.
```

Every XML document has a hierarchical tree-structure. Elements may be nested (embedded) within elements of a different type.

An XML element definition is a detailed specification of the form and content of an XML element, which includes the name (a generic identifier in XML terminology) for the element. A simple type contains no sub-element definitions. A complex type may include sub-elements.

### 4.6.6 XML Attribute

Attributes are used in XML to add information to the start-tag of an element to describe some aspect of a specific element occurrence. Attribute values are written in quotes and are separated from the attribute name by an equals sign. For example, a hypertext link in HTML is shown as `<A HREF = "URL" >`, where URL is the address of the uniform resource location (URL). Any number of attribute–value pairs may be defined for any element.

### 4.6.7 Entity

XML entities are named bodies of data, which can be referenced by an entity reference. Entity references always begin with “&” and end with “;”. A small number of entities are used to represent single characters that have special meanings in XML, such as `< (&lt;), > (&gt;) and & ( &amp; ). Numeric character references can be used to represent Unicode characters. For example, © is used to represent the © symbol. Other entities can be defined.

### 4.6.8 Namespace

A Namespace mechanism is provided to eliminate confusion when combining formats. This is used in XML schema to combine two schemas, to produce a third which covers a merged document structure.

### 4.6.9 The XML Family

XML is a family of technologies.
- CSS and XSL are style sheet languages.
- XPath provides a way to refer to individual parts of an XML document.
- XSLT is a transformation language for rearranging, adding, and deleting tags and attributes.
- DOM is a standard set of function calls for manipulating XML (and HTML) files from a programming language.
- SAX is a simple API for XML.
Session 2: Law #1: Interoperability: it's all about the people

That's because the real problems with interoperability are not technical. They've never been technical. It's the people that are the problem, because IT systems are only extensions of the people who wrote them, who maintain them, who use them.

Very often, interoperability projects are created to cause change in the business practices in the institution sponsoring them. This further increases the difficulty of getting people to align, since there is often contention due to the proposed change, and usually considerable lack of clarity over its exact impact on the business practices of the sponsor.

Session 2: Law #2: Complexity: you can move it around, or externalize it, but you can’t make it go away

Generally, the complexity arises from three different sources:

- The innate complexity of the problem, the real underlying scientific, clinical and business processes involved
- Diversity in approach and thinking between different instances of the same processes and data items due to different legacy practices, cultures and personalities
- Human factors (poor communication, poor motivation, failure to understand, fraudulent behaviour)

There’s two ways for a solution to be too simple: it just completely fails to account for the complexity of the problem, or it finds some way to externalize the complexity so it becomes Someone Else’s Problem.

The bottom line is that the complexity that does exist must be handled somewhere, by someone. It’s not going to go away

this implies that requirements gathering should actually capture the extent of the complexity of a project, and this is most often a total failure – in particular, people seem hardwired to express their requirements as simple principles with some special cases

Session 2: Law #3: Healthcare Software: Cheap, Flexible, Interoperable – you can have any two

there’s a trade-off between the upfront price of software and flexibility

The simplest pieces of software are not at all flexible: they solve exactly one problem in only one way

being flexible has a price – as software becomes more flexible, more skills are required to work with it,
Obviously there is an automatic cost for interoperability: adding the exchange feature cannot come for free.

However the general trade-off between interoperability and cost is only very loose.

The real tradeoff is between flexibility and interoperability. In order to use any information that is exchanged, someone has to understand it.

---

**Session 2: Driving Quality—A Health IT Assessment Framework for Measurement**

A CONSENSUS REPORT

National Quality Forum


The Health IT Utilization Expert Panel expands on HITEP’s Quality Data Set (QDS), a framework developed to clearly define concepts used in quality measures and clinical care to drive the use of quality measurement based on information available from an electronic health record (EHR).

The Expert Panel developed the *Health IT Utilization Assessment Framework* (framework). The framework is designed to define a method for expressing data that can be captured by health IT systems to understand and measure their effectiveness.

The framework is expected to support effective and meaningful EHR use assessment by:

- enabling development of measures of effective health IT use;
- understanding capabilities of the EHR and other health IT tools to meet meaningful use requirements;
- assessing unintended consequences of health IT usage;
- enabling information capture as a byproduct of clinical workflows;
- enhancing collaboration between health IT vendors, purchasers, implementers, and certifying bodies by encouraging the use of common health IT assessment strategies;
- enabling determination of high-priority health IT usage that supports certification of real-world implementations; and
- encouraging clinical effectiveness research regarding unintended consequences of health IT usage as well as research to determine effective health IT utilization.

Background
For health IT systems to fulfill their promise, however, they must support patient care directly—through clinical decision support and quality improvement—and support multiple uses of health information—through public reporting, public health surveillance, and clinical effectiveness research.

Identifying and developing measures of health IT activities and tracking performance on these metrics will be critical to assessing effective health IT use and driving more appropriate use where necessary. Measuring the quality of health IT use also requires an understanding of the system’s functions and the capabilities that track and monitor when and how it is used.

HITEP further defined how that information is captured within a clinical workflow with data flow attributes:

- **source** (e.g., the originator of the information, e.g., a clinician, patient, or device);
- **recorder** (e.g., a clinician, patient, or device and possibly different from the source);
- **setting** (e.g., hospital, home, ambulatory setting); and
- **health record field** (e.g., location in the EHR where the information should reside).

The QDS has been incorporated into the Healthcare Information Technology Standards Panel (HITSP) updates to the Quality Interoperability Specification, a standard that encodes electronic quality measure data, and the HITSP components to which it refers.

This report describes the Health IT Utilization Expert Panel’s approach to developing a framework to describe the information required to measure effective health IT utilization and presents the Expert Panel’s final output, the Health IT Utilization Assessment Framework (framework) itself. The framework is designed to help define a method for expressing data that can be captured by health IT systems to understand and measure their usage.

Health IT Utilization Assessment Framework Components

A health IT use measure must be able to identify:

- **Actor**: a person or electronic system that performs actions required in a measure of health IT utilization,
- **Content**: the concept on which an action is taken, and
- **Action**: something a measure recommends to a person or a computer programmed by a person.
The following proposed attributes were identified as common to all framework elements:

- **Source**: the originator of a data element, which may be an individual or a device
- **Recorder**: the individual or device that enters the data element in a health record field (may also be the source of the data)
- **Setting**: the physical location where a data element is captured, defining the encounter location where the data are expected to originate
- **Date/time**: the precise time and date recorded in an electronic health system
- **Method**: The manner in which the data element is captured (e.g., data entry by a user, a system query from one application to another, etc.)
- **Justification (for action or lack of action)**: information about why an action is taken or not taken, potentially taken from a pick list or a free-text entry
- **Content specific attributes**: information (or metadata) that provides additional detail about an individual element (e.g., for medications, the content-specific attributes are, for example, frequency, duration, dose, and route).

**Actor** An actor can be a member of the healthcare delivery team, a patient, a caregiver, or an electronic system. This component of a utilization data element captures who performs the specific health IT action being measured.

**Content** Content is the substance or subject matter on which actions are taken. This component of a utilization data element captures the information about which a health IT action is expected.

**Action** An action is an interaction with the health IT system that can be the product of human action or a programmed activity of the health IT system itself. The action is the health IT functionality that is measured or the health IT intervention that is called for in a quality measure.
Measuring CDS Effectiveness

The framework provides direction to standardize this process by defining data context for action, actor, and content.
Using these two new utilization data elements (depicted in Figures 8a and 8b) in a measure can result in the incorporation of existing QDS data elements into the measure’s denominator to define the population and the incorporation of the utilization elements to determine the numerator.

As shown in Figure 9, a measure of CDS effectiveness can identify the population of patients in the denominator using existing QDS concepts:

Each of these denominator elements uses the QDS Model as previously described. The numerator contains new data elements to indicate that either of the following has occurred: 1) an order action by a provider actor for aminoglycoside (using RxNorm) content or 2) a documentation action by a provider actor for reason and justification content.

Recommendations and Future Work
The Health IT Utilization Assessment Framework is expected to evaluate effective and meaningful EHR use and help avoid unintended consequences by:

- enabling development of measures of effective health IT use;
- understanding capabilities of the EHR and other health IT tools to meet meaningful use requirements;
- assessing unintended consequences of health IT use;
- enabling information capture as a byproduct of clinical workflows;
- enhancing collaboration among health IT vendors, purchasers, implementers, and certifying bodies by encouraging the use of common health IT assessment strategies;
- enabling determination of high-priority health IT use that supports certification of real-world implementations; and
- encouraging clinical effectiveness research regarding unintended consequences of health IT use as well as research to determine effective health IT utilization.

The framework put forth in this report provides a unique approach to identifying and measuring:

1) the use of health IT applications,
2) whether the workflow (driven by the system’s user interface) occurs as designed, and
3) that such use improves care processes, quality, and safety.

However, there is currently no standard mechanism to determine the appropriate utilization of EHRs. To accomplish such measurement requires incorporating metrics of usage directly within the EHR infrastructure.

A standard framework of user interactions with eHRs and transactions among them is required to

- generate measures of use,
- enable clinical effectiveness research,
- determine unintended consequences of EHR use, and
- evaluate the real-world usability of EHR vendor products.

User-centered design is a common approach to developing software products.

Four components of user-centered design:

1) specify the context of use,
2) specify requirements,
3) create design solutions, and
4) evaluate designs.

Framework Refinement and Evolution Recommendations

1. NQF should incorporate the Health IT Utilization Assessment Framework into the QDS Model and continue to manage its evolution with public consensus as existing standards and quality measures evolve.
2. The Department of Health and Human Services’ Federal Advisory Committees should consider the adoption of health IT utilization metrics as an EHR certification requirement; specifically, adopt the capture, logging, and evaluation by EHRs of each of the 20 action triplets identified in this report as EHR certification requirements.

3. The health IT use data requirements should be incorporated into standard frameworks such as the HL7 EHR Functional Model to encourage further evaluation and efforts by the appropriate Standard Development Organizations (SDOs).

4. Standards development organizations, professional societies, workflow planners, and other key stakeholders and entities should collaborate to standardize, harmonize, and identify definitions of and gaps in roles for all users of clinical applications and health IT systems.

5. NQF should pursue a call for measures of health IT utilization

---


*Integration* is the process by which different information systems are able to exchange data in a fashion that is seamless to the end user. The physical aspects of joining networks together are not nearly as complicated as getting unlike systems to exchange information in a seamless manner. Traditionally, communication between and among most disparate systems has been the result of costly, time-consuming efforts to build interfaces. In other words, interface programs are the tools used to achieve integration. An *interface* is a computer program that tells two different systems how to exchange data.

**BOX 11-1 The Benefits of Integration**

- Allows instant access to applications and data
- Improves data integrity with single entry of data
- Decreases labor costs with single entry of data
- Facilitates the formulation of a more accurate, complete client record
- Facilitates information tracking for accurate cost determinations

There are two general types of interfaces: point-to-point and those using integration engine software. A *point-to-point interface* is an interface that directly connects two information systems. Communication and transfer of data take place only between these systems. Historically these were the first types of interfaces used
More recent technology uses **interface engine** software to create and manage interfaces. This provides the ability to transfer information from the sending system to one or many receiving systems and allows users of different information systems to access and exchange information both in real-time and batch processing. **Real-time processing** occurs immediately, or with only a slight delay, whereas **batch processing** typically occurs once daily. In this situation, data are often not processed until the end of the day and therefore are not available to users until that time. Although batch processing was very common in the past,

The interface engine provides seamless integration and presentation of information results. Interface engines work in the background and are not seen by the user. This technology allows applications to interact with hardware and other applications. Interface engines allow different systems that use unlike terminology to exchange information without the need to build expensive point-to-point interfaces (Freedman 2007). This is done through the use of translation tables to move data from each system to the **clinical data repository**, a database where collective data from all information systems are stored and managed. The clinical data repository provides data definition consistency.

dors (Akhtar et al. 2005). Mapping is the process in which terms defined in one system are associated with comparable terms in another system. The major impact of using interface engines is that mappings for multiple receiving systems can be built for each sending system. For example, a client registration system can send

---

**BOX 11-2 Interface Engine Benefits**

- Improves timeliness and availability of critical administrative and clinical data
- Decreases integration costs by providing an alternative to customized point-to-point interface application programming
- Improves data quality because of data mapping and consistent use of terms
- Allows clients to select the best system for their needs
- Preserves institutional investment in existing systems
- Simplifies the administration of healthcare data processing
- Simplifies systems integration efforts
- Shortens the time required for integration
- Improves management of care, the financial tracking of care rendered, and efficacy of treatment

---
longitudinal health records speak of interoperability. **Interoperability** is the ability of two entities, whether those are human or machine, to exchange and predictably use data or information while retaining the original meaning of that data (Freedman 2007; Mead 2006; Schwend 2007). While the terms *interface* and *interoperability* are sometimes used interchangeably the interface engine routes information from one system to another but stops short of enabling the second system to understand and use that information. There are two types of interoperability: syntactic and semantic. Syntactic interoperability is the ability to exchange the structure of the data, but not necessarily the meaning of the data. It is also referred to as functional interoperability. Web pages built with HTML illustrate this type of interoperability. Semantic interoperability guarantees that the meaning of the exchanged data remains the same on both ends of the transaction. This is critical for clinical data. There have been several standardization efforts to achieve interoperability for EHRs. These include: the Health Level Seven (HL7) Clinical Document Architecture (CDA); the European Committee for Standardization (CEN) EN 13606-1, also known as EHRcom; and the openEHR. All provide specifications for how information should be exchanged (Kilic 2007). HL7 relies upon XML markup language for the storage and move-

The openEHR initiative is an international effort to provide semantic interoperability through the creation of specifications, open source software, and tools (Kalra 2007; Kilic & Dogac 2007; openEHR Foundation 2007). It builds upon more than 15 years of international research on the EHR. In the clinical arena, it strives to create high-quality, reusable clinical models of content and process that are defined by clinicians and known as archetypes, as well as formal interfaces to terminology. Each archetype contains a header, definition, and ontology section. The header has a unique code identifying the clinical concept defined as well as descriptive information such as author, version, and status. The definition includes restrictions obtained from the information model. The ontology section contains codes representing the meanings of nodes and constraints on text or terms, as well as linkages to terminologies such as SNOMED or LOINC. The Archetype Definition Language (ADL) is a formal language for the expression of archetypes.

2007). Integrating Healthcare Enterprise (IHE®) is a global initiative dedicated to improving patient care by promoting the way that computer systems in healthcare share information by promoting the adoption of standards such as DICOM and HL7 and interoperability (HIMSS 2007). IHE®, which was con-
Recent literature focuses upon yet another potential solution to interoperability, namely service-oriented architecture (SOA) (Bridges 2007; Cohen, Amatayakul, & Zeng 2007). SOA calls for placing key functions into modules that, along with new capabilities, may be re-used, similar to object-oriented programming. SOA defines a service as a self-contained unit of work that has well defined and understood capabilities. A service may be an entire process, a function supporting a process, or a step of a business process. Services may be built into a library that can be used to address enterprise needs. SOA does not require re-engineering of existing systems. SOA goes beyond other technologies in that it is vendor and technology neutral and does not require specific equipment or standards for operation. SOA can support information exchange among systems that use different programming languages. Its other characteristics hold promise for streamlining health information exchange. These include the ability to maintain a registry of services at the enterprise level that can be invoked after lookup, and the ability to provide quality service that includes security requirements such as authentication, authorization, and reliable messaging and policies.

**BOX 11-3  Factors That Slow Integration**

- **Unrealistic vendor promises.** Vendors often promise that their information systems are interoperable with other systems. Many customers find, however, that they face difficult, lengthy, and costly integration efforts after they have already purchased the system.
- **Unrealistic institutional timetable.** This is often based on a lack of understanding of the complexity of the integration process.
- **Changing user specifications.** As the integration process proceeds, users frequently request additional capabilities or change their minds regarding initial specifications.
- **Lack of vendor support.** Vendors may not provide enough support and assistance to facilitate the integration efforts.
- **Insufficient documentation.** Information regarding existing systems and related programming is imperative for achieving successful integration.
- **Lack of agreement among merged institutions.** Individual facilities within a merged enterprise may wish to continue use of their existing systems. This means there are more systems to integrate.
- **All components of a vendor's products may not work together.** Although difficulties are expected in attempts to integrate competing vendors' products, there may also be problems in integrating products developed by the same vendor.
THE NEED FOR INTEGRATION STANDARDS

The need to exchange client data is rapidly increasing in response to the demands placed by managed care as well as consumer demands for improved levels of healthcare. To derive the utmost benefit from data, it must have a consistent or standard meaning across institution, enterprise, and alliance boundaries, facilitating the exchange of client data. This is the basis for developing a data dictionary within an enterprise and a uniform language for use on a national and global scale. It is becoming increasingly important for hospitals and information system vendors to adopt and use uniform standards for the electronic exchange of clinical information (Ball & Farish-Hunt 2003). Use of uniform standards will provide safer and more efficient healthcare delivery systems and also play a critical role in compliance with government healthcare regulations.

Data Dictionary

The data dictionary defines terminology to ensure consistent understanding and use across the enterprise. Terms defined in the data dictionary should include synonyms found in the various systems used within the enterprise. This may be achieved, in some cases, through the use of the interface engine. For example, a

Master Patient Index

The integration process may require enhancements to the data dictionary, the clinical data repository, and the master patient index. The master patient index (MPI) is a database that lists all identifiers assigned to one client in all the information systems used within an enterprise. It assigns a global identification number for each client and allows clients to be identified by demographic information provided at the point of care. The MPI may use first and last names, birthdates, Social Security numbers, and driver’s license numbers. It cannot

Some of the key features of an effective MPI include the following:

• It locates records in real time for timely retrieval of information.
• It is flexible enough to allow inclusion of additional identification.
• It is easily reconfigured to accommodate network changes.
• It can grow to fit an organization of any size.
Uniform Language

One step in the integration process is the development of a uniform definition of terms, or language. This is essential for the easy location and manipulation of data. Many efforts to develop uniform languages are under way in the healthcare arena. For example, the National Library of Medicine developed the Unified Medical Language System (UMLS), which includes the Uniform Nursing Language. In addition, the American Nurses Association sponsors the

In addition to nursing, coding systems are used in other areas of healthcare to communicate information about medical diagnoses and procedures performed. This information is most commonly captured using the ICD-9/ICD-10 and CPT-4 systems. ICD-9 (International Classification of Disease—Ninth Revision) provides a classification for surgical, diagnostic, and therapeutic procedures (Thede 1999). This information is used for hospital billing and third-party payment throughout the United States. ICD-10 is an enlarged version of ICD-9 that is generally used in Europe at this time. The ICD-9/ICD-10 systems are published by the National Center for Health Statistics. Another commonly used classification is CPT-4 (Current Procedural Terminology—Fourth Revision), which is published annually by the American Medical Association. This system lists medical services and procedures performed by physicians and is used for physician billing and payer reimbursement.

SNOMED (Systemized Nomenclature of Human and Veterinary Medicine) is a classification system created by the College of American Pathologists (Thede 1999). This system includes signs and symptoms of disease, diagnoses, and procedures and is meant to represent the full integration of all medical information in an electronic medical record.
**HL7** A major standard for the exchange of clinical data for integration is Health Level 7 (HL7). HL7 refers to both an organization and its standards for the exchange of clinical data. The mission of the organization is to provide standards for the management and integration of healthcare data (Health Level Seven 2007). In particular, these standards address definitions of data to be exchanged, the timing of the exchanges, and communication of certain errors between applications. HL7 provides a structure that defines data and elements and specifies how the data are coded. The structure of the data element must follow HL7 rules, such as those specifying the length of the fields and the code nomenclature. Use of HL7 standards in individual applications improves the integration of these applications with other applications or systems using an interface engine. Benefits include easier and less costly integration within an organization and more accurate and useful data integration nationally and globally. Integration efforts and the development and use of integration

**DICOM** DICOM (Digital Imaging and Communications in Medicine) is a global Information technology standard first developed for the transmission of medical images and their associated information. It is now used in nearly all hospitals worldwide for the production, display, storage, retrieval, and printing of medical images and derived structured documents, as well as to manage related workflow

**BENEFITS OF INTEGRATION AND INTEROPERABILITY**

One major benefit of integration and the ability to exchange client data is the development of the electronic health record. In this case, integration allows data from many disparate information systems to be accessed from one point by the user, providing a complete record for each client. The clinical data repository is a key element of the EHR. It provides a storage facility for clinical data over time. The data in the clinical data repository may be gener-
Hospitals and healthcare enterprises also benefit from integration. Integration strategies permit data exchange within each hospital and across healthcare networks or enterprises, allowing them to find trends in financial and clinical data. Integration opens up a realm of possibilities to trend data, such as by provider, by diagnosis, or by cost. In this way, healthcare providers obtain improved information, making them better able to react to market changes and maintain a competitive edge.

Session 2: JSON vs XML
http://www.json.org/xml.html

Let's compare XML and JSON on the attributes that the XML community considers important.

From http://www.simonstl.com/articles/whyxml.htm

Simplicity

XML is simpler than SGML, but JSON is much simpler than XML. JSON has a much smaller grammar and maps more directly onto the data structures used in modern programming languages.

Extensibility

JSON is not extensible because it does not need to be. JSON is not a document markup language, so it is not necessary to define new tags or attributes to represent data in it.

Interoperability

JSON has the same interoperability potential as XML.

Openness

JSON is at least as open as XML, perhaps more so because it is not in the center of corporate/political standardization struggles.
many standards have been proposed, but their adoption has been slow. Why? System developers generally indicate that, while they would like to make use of standards, they can't find one that meets their needs.

However, systems developers, as users of controlled vocabularies, are like users everywhere: they may not always articulate their true needs.

**Desiderata**

The task of enumeration of general desiderata for controlled vocabularies is hampered in two ways. First, the desired characteristics of a vocabulary will vary with the intended purpose of that vocabulary and there are many possible intended purposes.

the desired vocabulary must be *multipurpose*. Some of the obvious purposes include: capturing clinical findings, natural language processing, indexing medical records, indexing medical literature, and representing medical knowledge.

A second obstacle to summarizing general desiderata is the difficulty teasing out individual opinions from the literature and unifying them

some of the characteristics that seem to be emerging from recent vocabulary research.

**Content, Content, and Content**

One approach to increasing content is to add terms as they are encountered, responding as quickly as possible to needs as they arise.[16] In this approach one adds complex expressions as needed rather than attempting a systematic, anticipatory solution.

An alternative approach is to enumerate all the atoms of a terminology and allow users to combine them into necessary coded terms,[17] allowing compositional extensibility

The real issue to address in considering the "content desideratum" is this: a formal methodology is needed for expanding content

we need formal, explicit, reproducible methods for recognizing and filling gaps in content

**Concept Orientation**

_Concept orientation_ means that terms must correspond to at least one meaning ("nonvagueness") and no more than one meaning ("nonambiguity"), and that meanings correspond to no more than one term ("nonredundancy").

"Myocardial Infarction" has a meaning which can be expressed in terms of a particular pathophysiologic process which affects a particular anatomic site. Now, if we use this concept to encode patient data, the meaning of the data will vary with the context ("Myocardial Infarction", "Rule Out Myocardial Infarction", "History of Myocardial Infarction", "Family History of Myocardial Infarction", "No Myocardial Infarction", etc.).
Concept orientation, therefore, dictates that each concept in the vocabulary has a single, coherent meaning, although its meaning might vary, depending on its appearance in a context.

**Concept Permanence**

*concept permanence*: the meaning of a concept, once created, is inviolate.

Its preferred name may evolve, and it may be flagged inactive or archaic, but its meaning must remain. This is important, for example, when data coded under an older version of the vocabulary need to be interpreted in view of a current conceptual framework.

**Nonsemantic Concept Identifier**

If each term in the vocabulary is to be associated with a concept, the concept must have a unique identifier. If a concept may have several different names, one could be chosen as the preferred name and the remainder included as synonyms.

Because many vocabularies are organized into strict hierarchies, there has been an irresistible temptation to make the unique identifier a hierarchical code which reflects the concept's position in the hierarchy.

There are several problems with using the concept identifier to convey hierarchical information. First, it is possible for the coding system to run out of room. A decimal code, such as the one described above, will only allow ten concepts at any level in the hierarchy and only allow a depth of four. Once assigned a code, a concept can never be reclassified without breaking the hierarchical coding scheme.

If a concept belongs in more than one location in the hierarchy (see “Polyhierarchy”, below), a convenient single hierarchical identifier is no longer possible.

It is desirable, therefore, to have the unique identifiers for the concepts which are free of hierarchical or other implicit meaning (i.e., *nonsemantic concept identifiers*) semantic connotations; such information should be included as attributes of the concept.[

**Polyhierarchy**

Different vocabulary users, however, may demand different, equally valid, arrangements of concepts. It seems unlikely that there can ever be agreement on a single arrangement that will satisfy all; hence the popular demand for multiple hierarchies.

Concept classification should be based on the *essence* of the concepts, rather than arbitrary descriptive knowledge.

However, unless there can be agreement on what the essence of concepts should be, there can never be agreement on what the appropriate hierarchy should be. Furthermore, if the essence of a concept is defined by its being the union of the essence of two other concepts, its classification becomes problematic.

**Formal Definitions**

Many researchers and developers have indicated a desire for controlled vocabularies to have *formal definitions* in one form or another.[23,25,26,27,36,38,39,40,41,42,43,44,45,46] Usually, these definitions are expressed as some collection of relationships to other concepts in the vocabulary.
This information can be expressed in a number of ways, including frame-based semantic networks[39] and conceptual graphs.

Many researchers have included in their requests that the definitional knowledge be made explicitly separated from and assertional knowledge which may also appear in the vocabulary.[25,40,42,45] For example, linking "Pneumococcal Pneumonia", via the "caused by" relationship, to "Streptococcus pneumoniae" is definitional, while linking it, via a "treated with" relationship, to "Penicillin" would be assertional.

Reject "Not Elsewhere Classified"

The controlled vocabularies should reject the use of "not elsewhere classified" terms.

Multiple Granularities

Associated with that purpose, usually implicitly, is some preconception of a level of granularity at which the concepts must be expressed

Multiple granularities are needed for multipurpose vocabularies. Vocabularies which attempt to operate at one level of granularity will be deemed inadequate for application where finer grain is needed and will be deemed cumbersome where coarse grain is needed. Insistence on a single level of detail within vocabularies may explain why they often are not reusable.[54] It also conflicts with a very basic attribute of medical information: the more macroscopic the level of discourse, the coarser the granularity of the concepts.

Multiple Consistent Views

If a vocabulary is intended to serve multiple functions, each requiring a different levels of granularity, there will be a need for providing multiple views of the vocabulary, suitable for different purposes.

We must be careful to confine the ability to provide multiple consistent views, such that inconsistent views do not result. For example, if we create a view in which concepts with multiple parents appear in several places in a single hierarchy, care must be taken that each concept has an identical appearance within the view.

Beyond Medical Concepts: Representing Context

Many researchers have expressed a need for their controlled vocabulary to contain context representation through formal, explicit information about how concepts are used.

One possible solution is to view the recording of patient information from an "event" standpoint, where each event constitutes some action, including the recording of data, occurring during an episode of care which, in turn occurs as part of a patient encounter.

Evolve Gracefully

An important desideratum is that those charged with maintaining the vocabulary must accommodate graceful evolution of their content and structure. This can be accomplished through clear, detailed descriptions of what changes occur and why,[64] so that good reasons for change (such as simple addition, refinement, precoordination, disambiguation, obsolescence, discovered redundancy, and minor name changes) can be understood and bad reasons (such as redundancy, major name changes, code reuse, and changed codes) can be avoided.
Recognize Redundancy

Synonymy is a type of redundancy which is desirable: it helps people recognize the terms they associate with a particular concept.

On the other hand, the ability to code information in multiple ways is generally to be avoided.

A good application will allow the user to add more detail to the coded problem, either through the addition of a coded modifier, through the use of unconstrained text, or perhaps a combination of both.

As vocabularies evolve, gracefully or not, they will begin to include this kind of redundancy. Rather than pretend it does not happen, we should embrace the diversity it represents while, at the same time, provide a mechanism by which can recognize redundancy and perhaps render it transparent.

Discussion

Shareability of vocabulary has become important as system builders realize they must rely on vocabulary builders to help them meet the needs of representing large sets of clinical terms.

The multipurpose nature of vocabularies refers to their ability to be used to record data for one purpose (such as direct patient care) and then be used for reasoning about the data (such as automated decision support), usually through a variety of views or abstractions of the specific codes used in data capture.

Session 3 Development of the Logical Observation Identifier Names and Codes (LOINC) Vocabulary

STANLEY M. HUFF, MD, ROBERTO A. ROCHA, MD, PHD, CLEMENT J. MCDONALD, MD, GEORGES J. E. DE MOOR, MD, PHD, TOM FIERS, MD, W. DEAN BIDGOOD, JR., MD, MS, ARDEN W. FORREY, PHD, WILLIAM G. FRANCIS, WAYNE R. TRACY, DENNIS LEAVELE, MD, FRANK STALLING, BRIAN GRIFFIN, PAT MALONEY, DIANE LELAND, LINDA CHARLES, KATHY HUTCHINS, JOHN BAENZIGER, MD


Background

There are at least three activities that would be made possible by a consistent representation of medical concepts: 1) the real-time exchange of medical data; 2) the exchange of decision-support programs including alerts, protocols, care pathways, and care plans; and 3) the pooling of data for outcomes research, quality assurance programs, and clinical research.

One proposed model of medical concept representation, which we adopt here, breaks the representation model into three components: a medical vocabulary or lexicon, a semantic data model (an information model), and a knowledge base.

The first component, a structured medical vocabulary, is an organized set of terms or words with associated codes.
The second component of medical concept representation, a semantic data model (SDM), is a description (template or data structure) of how vocabulary items can be combined to make a valid representation of medical information. Semantic data models have also generated a large number of publications.3,14 – 21 Often unrecognized as SDMs are the health care data exchange standards, 

In Europe, an important distinction is made between an interchange format and message syntax. An interchange format is a syntax-independent description of the structure and content of information within a message. As used in this article, the SDM for a message is the same thing as an interchange format.

Message syntax, however, is the specific way a message is encoded for transmission. Examples of message syntaxes include: EDIFACT, BER (Binary Encoding Rules) of ASN.1 (Abstract Syntax Notation One), and the delimited record format of HL7

HL7 message syntax defines a simple SDM

In the HL7 standard, vocabulary elements are linked to the message structure by the use of identifier (ID) and coded element (CE) data-type fields

Coded element data fields allow the set of values for a field to be defined by reference to a coded vocabulary that is external to the HL7 standard. This allows reuse of the many existing medical vocabularies and spares the HL7 group the time and effort of recreating terms that have been painstakingly created by other organizations.

It is the combination of the SDM and a structured medical vocabulary that makes the representation of a medical concept complete

The interdependence of the vocabulary and data model are especially evident in the context of data exchange standards

The LOINC Development Process

Problem Selection

An obvious but important first step for the LOINC committee was to select a domain of interest

The initial need we chose to address was the set of codes that could be used as observation identifiers in the HL7, ASTM 1238, and CEN TC251/PT3-008 and PT3-022 result messages. The observation identifier is the part of the result message that expresses what kind of observation is being made.

There were at least four reasons why we chose the creation of observation identifiers as our initial task. First, creating a consistent set of observation identifiers appeared to have an immediate benefit

Second, when the LOINC committee began, no structured vocabulary existed that was appropriate for naming clinical, laboratory, or physiologic measurements in result messages and that also had the correct degree of granularity to represent the names of procedures as commonly defined in laboratory information systems and in clinical information systems

Third, good vocabularies, such as SNOMED, did exist for expressing many of the coded values of the observations, so the creation of sets of coded values was not a high priority.
Fourth, while no standard set of observation identifiers existed, good starting lists were available.

**Additional Vocabulary Principles**

additional principles related to aspects of a good vocabulary
n There should be version control associated with the coding system. The version identifier should be a prominent component in each distribution.

n The code associated with a term should have no intrinsic or embedded meaning.

n The term component should stand on its own without additional components being required to convey the meaning of the concept.

n There should be an organization capable of extending, correcting, and maintaining the coding system. This should include evidence of organizational stability and adequate funding for the coding activity.

n There should be easy access to the vocabulary. It should be available by both paper and electronic distribution.

n There should be no limitation on who can acquire copies of the coding system.

n There should be no cost or minimal cost associated with access to and use of the coding system.

Likewise, maintenance fees are undesirable.

n The coding system should have training materials, such as tutorials, training syllabi, and printed user manuals. People with in-depth understanding of the coding system should be available to provide help and consultation.

n The coding system should be acceptable for use internationally.

n The coding system should be extensible. There should be no inherent limits on the number or types of codes that can be created.

n The coding system should be compositional. The creation of new codes should be guided by a data model.

**Organizational Structure**

There were at least three possible ways in which a vocabulary committee could have been organized: 1) as part of an existing standards development organization 2) as a commercial venture to produce a for-profit product; or 3) as a group of interested individuals with support from governmental or private sources.

**Selection of Participants**

This mixture of individuals proved to be essential to ensuring that the codes and names produced by the group would be appropriate for use in private hospitals, Veterans Health Administration hospitals, and commercial reference laboratories.

**Focusing the Scope**

Once the participants had been selected, the initial meetings commenced. The first meetings were used to focus the scope of the activity and to mutually educate one another.

Another key decision was to limit the initial scope to the representation of names and codes for clinical laboratory result observations.

**Developing the Model and the Vocabulary Content**

General Approach
The process of developing LOINC content was, and continues to be, an iterative combination of both bottom-up (empirical) and top-down (conceptual modeling) approaches.

The Initial Model
Given the conceptual background as defined in preceding paragraphs, we began with a simple four-axis model. The four axes were the component being measured, the specimen type (system), the precision of the measurement, and the method by which the measurement or observation was made.

Rapid Evolution of the Model

In the preliminary stages of LOINC development, the model evolved rapidly as the committee gained understanding and experience.

Chemical Subspecies and Kind-of-Property

Challenge Tests and Observation Methods

This iterative process of examining data, matching to the model, and modifying the model when necessary is the core process in creating the LOINC vocabulary.

Policies for Creating LOINC Names Using the Model

Once a reasonably robust model was in place, policies and procedures were instituted that made creation of names using the model consistent.

The LOINC Model in a Formal Notation

One of the strengths of the LOINC terminology is its underlying SDM.

The main objective of having a formal model is to make explicit all the discrete domains that are used by the different axes of the LOINC fully specified names.

Distribution of and Additions to the LOINC Vocabulary

Implementing and Evaluating the LOINC Vocabulary

For purposes of evaluation, the potential utility of LOINC centers on two propositions: the use of LOINC as a universal coding scheme, and the use of LOINC to improve mapping between systems.

The question of whether LOINC is a suitable universal vocabulary relates directly to adequate coverage. Despite the initial consideration of various clinical laboratory vocabularies as sources of terms and the actual size of the LOINC vocabulary.

The second proposition is that the use of LOINC can improve the process of mapping. Improvements could be in the form of greater accuracy in mapping or decreased time and costs. These benefits could be realized whether the mapping were done manually or by an automated process. Evaluations of the second proposition then would focus on the speed, efficiency, and accuracy of mapping between systems using LOINC.

Discussion

One key issue was the recognition of alternative styles of representing data in messages.

We have intentionally included both variable and value style names in the LOINC vocabulary, since both styles are in common use in systems today.
This leads to either inconsistent representation of the data in patient databases or to increased complexity in computer-to-computer interfaces.

A second issue became apparent as we began to examine patient data outside the clinical laboratory. It became clear that different systems used different levels of aggregation (pre-coordination) in the names used to describe clinical measurements.

**Conclusion**

The accurate exchange of clinical data between computer systems requires the combination of an SDM and a vocabulary. The SDM provides structure and context that are essential for the correct interpretation and understanding of the terms that the structure contains. The creation of the LOINC vocabulary has lead to a greater understanding of the interdependency of the SDM and the vocabulary.

---

**Session 3: Cross-mapping clinical notes between hospitals: An application of the LOINC Document Ontology**

Li Li, MS, C. Paul Morrey, PhD, David Baorto, MD, PhD  
New York Presbyterian Hospital, New York, NY

LOINC (Logical Observation Identifiers Names and Codes), a nationally recognized source of terminology for laboratory and other observations, has developed the LOINC Document ontology (DO) for standardizing clinical documentation with terms in a hierarchical structure. It consists of five axes: Kind of Document, Subject Matter Domain, Type of Service, Role and Setting. Each of the axes has a set of possible entry values which can be combined with values from the other axes to create LOINC-compatible combinations. A legitimate name according to LOINC DO should consist a value from “Kind of Document” and at least one value from the other four axes.
Discussion

We found that the LOINC DO provides a thorough semantic structure for representation of clinical documents.

One of the objectives of this project was to test the use of ontology – modeled notes for document interoperability between campuses. For such a use case, the issue of granularity came into play.

Addressing the possibility of using LOINC codes as a bridge between local documents and outside institutions yielded the finding that a relatively small percentage of our local documents find exact matches to existing LOINC codes (18-22%).

This indicates that most of the matching documents only matched at one or two axes. So the granularity of specification intercedes in precise matching to LOINC codes. This implies that, at least at the current stage, using these codes for exact document exchange may not be feasible, however, the specification of classes of documents certainly is.

Conclusion
The LOINC document ontology model can provide a robust representation of clinical documents. This can be exploited to assist in common naming conventions, finding notes by traversing the axis categories of Subject, Setting, Role and Type of Service. However, the granularity of representation can be an issue in mapping across sites using the LOINC DO and mapping to specific LOINC codes.

Session 3 LOINC® - A Universal Catalog of Individual Clinical Observations and Uniform Representation of Enumerated Collections

Daniel J. Vreeman, PT, DPT, MSc.
Assistant Research Professor and Research Scientist, Indiana University School of Medicine and Regenstrief Institute, 410 W. 10th Street, Suite 2000, Indianapolis, IN 46202


Introduction

We can build bridges across these islands of data much more efficiently by using data exchange standards (C J McDonald, 1997). LOINC® (Logical Observation Identifiers Names and Codes) is a universal code system for identifying laboratory and clinical observations that facilitates exchange and pooling of results for the clinical care, research, and outcomes management (Clement J McDonald et al. 2003). When used in conjunction with widely adopted messaging standards such as Health Level 7 (HL7), vocabulary standards like LOINC can be an essential ingredient for efficient electronic processing and storage of clinical data that comes from many independent sources.

Within LOINC we make a distinction between a) panels such as the “complete blood count” or “Braden scale”, which are collections that have enumerated discrete contents, and b) documents such as a “physical therapy evaluation note” or “discharge summary”, which are general information collections whose contents are not definitively enumerated

Methods

Overview of LOINC

LOINC constructs “fully specified” names according to an established model that contains six main axes (Table 1). The fully specified LOINC names contain sufficient information to distinguish among similar clinical observations, but do not carry all possible information about the testing procedure and result

Thus, different LOINC codes are assigned to observations that measure the same attribute but have different clinical meanings.

Hierarchical panel structure—LOINC represents panels, whether laboratory batteries or assessment instruments, by creating LOINC panel terms that are linked to an enumerated set of child elements. A complete hierarchical structure can be represented because the child elements themselves can be panel terms, which enables full nesting

Attributes of individual variables—The main LOINC table contains the LOINC code, fully-specified name, and fields for many other additional attributes about the terms
Answer lists—The clinical meaning of many questions on assessment instruments are inextricably tied to the allowable answer options, and thus LOINC contains a data structure for linking LOINC observation codes to answer lists.

Panel-specific attributes of variables—The LOINC model for representing assessment content supports reuse of variables across panels, but also enables some attributes to be stored at the level of the instance of the variable within a particular panel.

Results
A Growing Universal Instrument and Item Bank
With the iterative refinements made to the LOINC model for representing panels, we have successfully represented a wide variety of content. Over time we have continued to add new content to LOINC, including many patient assessments.

The LOINC model accommodates panels with categorical variables that have enumerated answer lists as well as other clinical variables that report physical quantities, like height, weight, or systolic blood pressure using the typical LOINC terms. In addition to the content available in the structured export, LOINC also includes several other standardized collections of variables in the same data model.

Regenstrief is also creating LOINC content in collaboration with developers of two innovative clinical research variable sets: the Phenotypes and eXposures (PhenX) measures and the Patient-Reported Outcomes Measurement Information System (PROMIS). PhenX (PhenX n.d.) is funded by the National Human Genome Research Institute to develop and distribute a set of high priority measures that will enable cross-study comparisons and analyses in genome-wide association and other clinical studies. PROMIS (PROMIS n.d.) is funded by the National Institutes of Health Roadmap for Medical Research Initiative to develop publicly available computer-adaptive tests for measuring patient-reported symptoms such as pain, fatigue, physical functioning, and other aspects of health-related quality of life across a wide variety of chronic conditions.

By collecting the details about individual variables and the panels that contain them, LOINC makes it easy for system implementers to access the content in a common format. The Personal Health Record being developed by the National Library of Medicine is an early example of a system that has the capability to read the LOINC panel definition and dynamically create electronic data collection forms (Lister Hill National Center for Biomedical Communications - U.S. National Library of Medicine - National Institutes of Health, n.d.). Having a standard for patient observations of all kinds also makes it possible to construct interoperable electronic result messages that blend routine clinical data with results from formal research questionnaires.

Enabling Interoperability Together with Other Health Information Technology Standards
LOINC’s standardized representation of assessment content is an important enabling component of interoperable exchange between electronic systems and has been adopted by several large U.S. initiatives.

Additionally, the LOINC model was incorporated into the HL7 Draft Standard for Trial Use “CDA Framework for Questionnaire Assessments and CDA Representation of the Minimum Data Set Questionnaire Assessment” (Health Level Seven International, n.d.). HL7’s questionnaire assessment draft standard filled an important gap by providing an implementation guide for patient assessments. This standard includes both an internationally-applicable component that supports exchange of any assessment represented in LOINC and a detailed guide for implementing the exchange of the MDS version 3.
Variation Across Panels
When adding the content from these assessments to LOINC, we found substantially more variation across panels than we had initially expected, and some of it could have been avoided. Many of the variables in the OASIS, MDS, and CARE instruments are very similar, but not directly comparable. The lack of direct comparability is also present between different versions of the same instrument.

Discussion
LOINC contains a well-developed model for representing variables, answer lists, and the collections that contain them. With continued growth, LOINC is expanding as a large “master observation file” that provides a uniform representation of the essential attributes for items in data collection forms. The level of standardization achieved by modelling this content in LOINC provides an important component of enabling interoperable data exchange, storage, and processing. By creating a uniform representation and distributing it worldwide at no cost, LOINC aims to lower the barriers to interoperability among systems and make this valuable data available across settings when and where it is needed.

LOINC’s inclusion of assessments aims to achieve a convergence of codes and vocabulary for observations by providing a uniform and standardized representation. This approach complements the current efforts to build metadata repositories and other clinical information models by providing the lingua franca that can populate the models and be used for exchanging data between and among clinical and research systems.

Lessons and Recommendations

**Variation abounds and limits comparability**—As we modelled various assessment instruments in LOINC, we were struck by the degree of variation among observations measuring similar clinical characteristics. The lack of comparability between the assessment instruments required for payment by CMS in post acute care settings creates obstacles for caring for often-fragile patients; the information on one assessment cannot be used to directly populate another.

Before inventing yet another variant, the possible benefits should be weighed against the loss of data comparability. The larger the amount and generalizability of existing data, the more carefully we should consider any potential modifications. Having a large collection of panels and variables in LOINC’s uniform format should make it easier to review and reuse existing content.

**Intellectual property restrictions can be large barriers**—Prior to being able to include a copyrighted instrument in LOINC, Regenstrief must negotiate separate (often resource consuming) agreements with each copyright holder.

**A master catalogue and uniform representation is a step forward**—Building a master catalogue of panels and variables in LOINC is an enabling step towards interoperable data exchange, but much work remains.
2.1. What is SNOMED CT?
SNOMED Clinical Terms (SNOMED CT) is a comprehensive clinical terminology that provides clinical content and expressivity for clinical documentation and reporting. It can be used to code, retrieve, and analyze clinical data. SNOMED CT resulted from the merger of SNOMED Reference terminology (SNOMED RT) developed by the College of American Pathologists (CAP) and Clinical Terms Version 3 (CTV3) developed by the National Health Service (NHS) of the United Kingdom. The terminology is comprised of concepts, terms and relationships with the objective of precisely representing clinical information across the scope of health care. Content coverage is divided into hierarchies.

2.2. SNOMED CT uses
Health care software applications focus on collection of clinical data, linking to clinical knowledge bases, information retrieval, as well as data aggregation and exchange. Information may be recorded in different ways at different times and sites of care. Standardized information improves analysis. SNOMED CT provides a standard for clinical information. Software applications can use the concepts, hierarchies, and relationships as a common reference point for data analysis. SNOMED CT serves as a foundation upon which health care organizations can develop effective analysis applications to conduct outcomes research, evaluate the quality and cost of care, and design effective treatment guidelines. Standardized terminology can provide benefits to clinicians, patients, administrators, software developers and payers. A clinical terminology can aid in providing health care providers with more easily accessible and complete information pertaining to the health care process (medical history, illnesses, treatments, laboratory results, etc.) and thereby result in improved patient outcomes. A clinical terminology can allow a health care provider to identify patients based on certain coded information in their records, and thereby facilitate follow-up and treatment.
3.1. Concepts
In the context of this document, a “concept” is a clinical meaning identified by a unique numeric identifier (ConceptId) that never changes. Concepts are represented by a unique human-readable Fully Specified Name (FSN). The concepts are formally defined in terms of their relationships with other concepts. These logical definitions give explicit meaning which a computer can process and query on. Every concept also has a set of terms that name the concept in a human-readable way.

3.1.1. Concept granularity
Support for multiple levels of granularity allows SNOMED CT to be used to represent clinical data at a level of detail that is appropriate to a range of different uses. Concepts with different levels of granularity are linked to one another by | is a | relationships. This enables appropriate aggregation of specific information within less detailed categories.

3.1.2. Concept Identifiers
Each SNOMED CT Concept has a permanent unique numeric Identifier which is known as the ConceptId. The sequence of digits in a ConceptId does not convey any information about the meaning or nature of the Concept. The meaning of Concept is represented in human-readable forms by Descriptions and in a computer processable form by Relationships with other Concepts.

3.2. Descriptions
Concept descriptions are the terms or names assigned to a SNOMED CT concept. “Term” in this context means a phrase used to name a concept. A unique DescriptionId identifies a description. Multiple descriptions might be associated with a concept identified by a ConceptId.

3.2.1. Types of descriptions
3.2.1.1. Fully Specified Name
Each concept has one Fully Specified Name (FSN) intended to provide an unambiguous way to name a concept. The purpose of the FSN is to uniquely describe a concept and clarify its meaning.

3.2.1.2. Preferred Term
Each concept has one Preferred Term in a given language dialect. The Preferred Term is a common word or phrase used by clinicians to name that concept.

3.2.1.3. Synonym
A synonym represents a term, other than the FSN or Preferred Term, that can be used to represent a concept in a particular language or dialect.

3.3. Relationships
Relationships link concepts in SNOMED CT. There are four types of relationships that can be assigned to concepts in SNOMED CT:
• Defining
• Qualifying
• Historical
• Additional

3.3.1. Relationships and concept definitions
Each concept in SNOMED CT is logically defined through its relationships to other concepts. Every active SNOMED CT concept (except the SNOMED CT Concept Root concept) has at least one | is a | relationship to a supertype concept.

3.3.2. IS A Relationships
| is a | relationships are also known as “Supertype - Subtype relationships” or “Parent - Child relationships”. | is a | relationships are the basis of SNOMED CT's hierarchies, as illustrated below.

A concept can have more than one | is a | relationship to other concepts. In that case, the concept will have parent concepts in more than one sub-hierarchy of a top-level hierarchy. Subtype relationships can be multi-hierarchical.

3.3.3. Attribute Relationships

An attribute Relationship is an association between two concepts that specifies a defining characteristic of one of the concepts (the source of the Relationship). Each Attribute Relationship has a name (the type of Relationship) and a value (the destination of the Relationship). For example, The combination of the attribute Relationships and | is a | relationships associated with a concept represent the logical definition of that concept. The logical concept definition includes one or more supertypes (represented by | is a | relationships), and a set of defining Attributes that differentiate it from the other concept definitions.

Example:
Since pneumonia is a disorder of the lung, the logical definition of the concept | Pneumonia (disorder) | in SNOMED CT includes the following Relationship. The Attribute | Finding site | is assigned the value | Lung structure (body structure) |

• | Finding site | = | Lung structure (body structure) |

4 Attributes Used in SNOMED CT

This part of the guide provides an overview of the defining attributes used by the SNOMED CT Concept Model. Further details are provided in the chapters dedicated to each hierarchy.

4.1. Introduction

SNOMED CT currently uses over 50 defining attributes to model concept definitions. Each SNOMED CT attribute can usually be applied to one hierarchy and for a few attributes to more than one hierarchy. The hierarchy or hierarchies to which an attribute can be applied are referred to as the “domain” of the
attribute. Each attribute can be given a limited set of values; this set of values is called the “range” of the attribute.

4.1.1. Domain
The Domain is the hierarchy to which a specific attribute can be applied. The Domain of the attribute | ASSOCIATED MORPHOLOGY | is the | Clinical finding | hierarchy. A | Procedure | cannot have an | ASSOCIATED MORPHOLOGY |. A | Procedure | has a | PROCEDURE MORPHOLOGY |.

4.1.3. Range
The Range is the set of values allowed for each attribute. For example, the Range for | ASSOCIATED MORPHOLOGY | is | Morphologically abnormal structure (morphologic abnormality) | and its descendants, and the Range for | FINDING SITE | is | Anatomical or acquired body structure (body structure) | and its descendants in the | Body structure | hierarchy.

4.2. Attribute Hierarchies in SNOMED CT
Selected SNOMED CT attributes have a hierarchical relationship to one another known as “attribute hierarchies”. In an attribute hierarchy, one general attribute is the parent of one or more specific subtypes of that attribute. Concepts defined using the more general attribute can inherit concepts modeled with the more specific subtypes of that attribute.

4.2.1. Attribute hierarchies used in modeling Procedures
Three groups of attributes are organized as a simple two-level hierarchy. The three top level attributes are | PROCEDURE SITE |, | PROCEDURE DEVICE |, and | PROCEDURE MORPHOLOGY |. Each has a sub-attribute to represent the direct object, and another to represent the indirect object. In addition, | PROCEDURE DEVICE | can be specialized by the attributes | USING DEVICE | and | USING ACCESS DEVICE |.

4.4.1. PROCEDURE SITE
The | PROCEDURE SITE | attribute describes the body site acted on or affected by a procedure.

4.4.3. METHOD
This attribute represents the action being performed to accomplish the procedure.

4.4.2. PROCEDURE MORPHOLOGY
| PROCEDURE MORPHOLOGY | is the attribute used to specify the morphology or abnormal structure involved in a procedure.

4.4.4. PROCEDURE DEVICE
| PROCEDURE DEVICE | is a general attribute used to model devices associated with a procedure.

4.12. Relationship groups in SNOMED CT
Multiple attributes and their values can be grouped together into "Relationship groups" to add clarity to concept definitions. A Relationship group combines an attribute-value pair with one or more other attribute-value pairs. Relationship groups originated to add clarity to | Clinical finding | concepts which require multiple | ASSOCIATED MORPHOLOGY | attributes and multiple | FINDING SITE | attributes and to | Procedure | which require multiple | METHOD | attributes and multiple | PROCEDURE SITE | attributes. However, Relationship groups are not limited to | Clinical finding | and | Procedure | concepts.
Chapter 5 Hierarchies

SNOMED CT concepts are organized into hierarchies. There are two special Codes referred to as the | Root Concept Code | and the | Root Metadata Code |. They are at the "root" of the two hierarchies that contain all Concept Codes in SNOMED CT. The root named "SNOMED CT Concept" subsumes (is the supertype of) the top-level concepts and all the concepts beneath them (their subtypes), and the root named “SNOMED CT Model component” subsumes all the metadata components. As the hierarchies are descended, the concepts within them become increasingly specific (or granular). A brief description of the content in each hierarchy is given below.

5.1. Summary of Top Level Hierarchies
5.1.1. Top Level Concepts

Table 71: Top Level Concepts

| Physical force |
| Event |
| Environment or geographical location |
| Social context |
| Situation with explicit context |
| Staging and scales |
| Physical object |
| Qualifier value |
| Record artifact |
| Clinical finding |
| Procedure |
| Observable entity |
| Body structure |
| Organism |
| Substance |
| Pharmaceutical / biologic product |
| Specimen |
| Special concept |
| Linkage concept |

Chapter 6

Structure and Technology Considerations

The structure and technology behind SNOMED CT enables organizations to implement it and integrate it into their own clinical and business processes and applications. SNOMED CT offers additional capabilities to facilitate customization of an implementation to meet the unique requirements of an organization.

SNOMED CT data structure: SNOMED CT data components and their relationships, including the core table structure, as well as:

- History
- Subsets
- Cross Mapping
- Extensions
- SNOMED CT applications and services
6.2. History
The content of SNOMED CT evolves with each release. The types of changes made include new Concepts, new Descriptions, new Relationships between Concepts, new Cross Maps, and new Subsets, as well as updates and retirement of any of these Components. Drivers of these changes include changes in understanding of health and disease processes; introduction of new drugs, investigations, therapies and procedures; and new threats to health, as well as proposals and work provided by SNOMED users.

6.4. Cross Mappings
Cross Mappings enable SNOMED CT to effectively reference other terminologies and classifications. Each cross map matches SNOMED CT concepts with another coding scheme that is called the “target scheme”. The Cross Mapping mechanism enables the distribution of Cross Maps from SNOMED Clinical Terms in a common structure.

Session 3: Normalized names for clinical drugs: RxNorm at 6 years
Stuart J Nelson, Kelly Zeng, John Kilbourne, Tammy Powell, Robin Moore
J Am Med Inform Assoc 2011;18:441e448. doi:10.1136/amiajnl-2011-000116 441

INTRODUCTION
RxNorm is a standard nomenclature developed by the United States National Library of Medicine (NLM) in the field of medications.

The RxNorm vocabulary is available at no cost from http://www.nlm.nih.gov/research/umls/ rxnorm/

The creation of RxNorm was motivated by the need for a single, standard, multipurpose terminology for representing medications.

Many clinical information tasks can benefit from the use of a standard terminology for representing drug information, including creation of electronic medical records (EMR), automated decision support, quality assurance, healthcare research, reimbursement, and mandatory reporting.

A key concept in the model was the notion of a clinical drug as it appears in a provider’s medication order, comprised of active ingredient, strength, and dose form.
RxNorm is built upon what is already availabled various drug vocabularies commonly used in pharmacy management and drug interaction software. It is built by creating normalized drug names based on the information and names in the contributing source vocabularies. It has a limited and controlled scoped the domain of medications expressible as clinical drugs.

Today, RxNorm continues to evolve as a standard for clinical information exchange. Any source vocabulary included in RxNorm can be used to achieve compliance with the ‘Meaningful Use’ requirements for electronic health records (EHR); such designation establishes an important bridge to full RxNorm adoption. Recommended by the Healthcare Information Technology Standards Panel (HITSP), RxNorm is the designated vocabulary to represent ‘Medication Brand Name,’ ‘Medication Clinical Drug Name,’ and ‘Allergy/Adverse Event Product’ (if the product causing the adverse event is a medication).

CONTENT WITHIN RXNORM
No single drug vocabulary provides complete and interoperable drug names, codes, and relevant information. RxNorm takes the multiple drug names, using them to complement each other and reconcile the conflicts among them. By aggregating and organizing content from various source drug vocabularies, RxNorm can derive a more complete and consistent representation of drug names, codes, and relevant information.

Unique identifiers and normalized names
Unique identifiers and normalized names are the primary mechanism to group together semantically equivalent terms and codes from various source vocabularies. Different names and codes can be mapped to each other if they share the same RxNorm Concept Unique Identifier (RxCUI).

Multiple levels of description and relationships
Multiple levels of description and relationships enable RxNorm to represent drugs from multiple purposes or various perspectives. To accommodate these various needs and contexts, RxNorm assigns term types (TTY) to organize various levels of description, representing drugs not only at the clinical drug level (SCD) {ingredient+strength+dose form}, but also at levels of ingredient, which includes single ingredient (IN), multiple ingredients (MIN), and precise ingredient (PIN), clinical drug component (SCDC) {ingredient+strength}, clinical drug dose form (SCDF) {ingredient+dose form}, and pack (multiple clinical drugs or clinical drugs designed to be administered in a specified sequence) as well.

Attributes
Attributes are relevant information about a drug at an appropriate level of abstraction. This relevant information indicates various aspects of a drug.

The National Drug File Reference Terminology (NDF-RT)
The National Drug File Reference Terminology (NDF-RT) was integrated into RxNorm as a source vocabulary beginning with the RxNorm June 2010 monthly release. NDF-RT is a resource developed by the Department of Veterans Affairs (VA) Veterans Health Administration, as an extension of the VA National Drug File. The inclusion of NDF-RT has provided RxNorm with an additional type of information. Besides drug names and codes, providers now can find within the RxNorm data the clinical properties associated with certain drugs.

Source content
Source content, if provided to the NLM and indicated as releasable, is present in RxNorm release files. Users wishing to use a source vocabulary integrated with RxNorm may find that, at times, content from a particular source is not available in released RxNorm for a given drug.

TYPICAL USES OF RXNORM
Typical uses of RxNorm include using RxNorm standard names and codes to capture drug product information in EHR, cross mapping among disparate drug vocabularies, and facilitating medication-related clinical decision support.

Recommended by HITSP, RxNorm is the designated vocabulary to represent 'Medication Brand Name,' 'Medication Clinical Drug Name,' and 'Allergy/Adverse Event Product' (if the product causing the adverse event is a medication). RxNorm names and codes are used to represent medication names in the Continuity of Care Document (CCD). CCD is a Clinical Document Architecture (CDA) Release 2 implementation that maps the Continuity of Care Record (CCR) elements into a CDA representation, harmonizing CCR and CDA into a common framework.12 CCD is a content standard for patient summary records.

Semantic interoperability
RxNorm has been used in the patient data exchange between the VA and the Department of Defense (DoD).13 By mapping from VUID (VA Unique IDentifier) or NCID (Numeric Concept ID used by DoD) to RxCUI, real-time bi-directional encoded data exchange between the two agencies is enabled.

Medication-related decision support
Standard drug names and codes, and cross-mapping among various drug terminologies can be considered as a necessary first step toward medication-related decision support.

NLM-PROVIDED RXNORM-RELATED SERVICES
To facilitate the use of RxNorm, the NLM has provided a few additional services, mainly developed as research projects, including RxNav, RxTerms, MyMedicationList, and MyRxPad.

RxNav, the RxNorm Navigator, was originally primarily a browser for RxNorm. It is a software application that displays RxNorm names and codes and the relationships among them based on a user’s search input. It serves as a supplemental tool to help users browse through RxNorm in a visually friendly and interactive manner.

RxTerms17 is a drug interface terminology derived from RxNorm to facilitate CPOE. It reorganizes RxNorm names and codes into a two dimensional representation tailored for prescription writing; it eliminates certain drug names that are less likely to be needed in a prescribing environment as well.

MyMedicationList (MML) is an application that helps patients create, update, and save their medication lists.18 MyRxPad is a prototype application that is intended to help prescribers lower some of the e-prescribing adoption barriers and encourage an early positive experience of e-prescribing.

MML and MyRxPad provide an alternative approach to medication reconciliation across points of care.

FUTURE WORK
Users often use RxNorm names and codes to record drug names, and RxNorm attributes for various purposes in their PHR or EMR applications. For these specific uses, users may find names and codes from sources other than RxNorm less relevant. Thus, a pre-processed RxNorm subset that contains only data that are pertinent to users’ needs seems desirable. The subset can also facilitate the adoption of RxNorm as providers prepare themselves to meet the requirements for ‘Meaningful Use’ Stage 2 when full adoption of RxNorm is expected.

RxNorm e-prescribing subset. The subset intends to provide a comprehensive nomenclature that includes all prescribable clinical drugs and packs on the market for use in e-prescribing systems. The subset would
Session 4: The Unified Medical Language System: An Informatics Research Collaboration

The project is focused on overcoming two important barriers to the development of information systems that can help health professionals make better decisions. These barriers are the disparity in the terminologies used in different information sources and by different users, and the sheer number and distribution of machine-readable information sources that might be relevant to any user inquiry.

The UMLS project has produced and widely disseminated four multipurpose knowledge sources designed for system developers: the Metathesaurus, the Semantic Network, the Information Sources Map, and the SPECIALIST Lexicon and associated lexical programs.

Genesis of the UMLS Project

The objective of this program . . . is to solve what is the most fundamental barrier to the application of computers in medicine; namely, the lack of a standard language in medicine. We will attempt to build that vocabulary, a language that will cross between the biomedical literature and the observations on the patient, as well as the educational applications in the school, a language which allows those areas to be interrelated.

—Donald A. B. Lindberg, M.D., March 19, 1985

The initiation, funding, and ultimately the maintenance of such a system were therefore proposed as a reasonable undertaking for a federal research agency.

The NLM team exemplified a key characteristic of UMLS research—its dependence on contributions from many disciplines, including medicine, biomedical science, medical informatics, computer science, library and information science, and linguistics.

The NLM UMLS team selected the “task-order” research contract as the most appropriate vehicle for funding extramural UMLS research collaborators. Task-order research contracts involve a series of research tasks that are defined and negotiated throughout the life of the contract.

That “the principal barrier to effective integrated access to biomedical information is the tremendous array of classification and representation schemes used in major information sources: the published biomedical literature, patient records, medically related data banks, and medical knowledge bases.

include Federal Medication Terminologies components (RxNorm, UNII, NDC, NDF-RTclasses) as well as important attributes in RxNorm.

Several challenges remain in constructing such a subset. One obstacle is to identify drugs that are on the market.
The solution to this fundamental medical information problem is the development of conceptual links among disparate classification schemes.

UMLS collaborator engaged in two types of work: tasks jointly developed with NLM and other UMLS collaborators to assist in defining, building, and testing central UMLS components and individually designed and motivated projects related to UMLS research goals and local research priorities.

Making Key Project Decisions

First, NLM would select a basic approach following an assessment of elements of the work and recommendations of all UMLS participants. This would then be discussed, refined, and inevitably improved by all collaborators, including those who may have strongly favored a different basic approach.

In the short term, NLM’s approach led to early editions of the UMLS Metathesaurus with content seen as insufficient for a number of important clinical applications.

the Library’s decision not to generate a new clinical vocabulary

The UMLS and the Advance of Information Technology

Early versions of UMLS components will be relatively simple structures, offering modest enhancements to current systems with respect to their representation of the interrelationships among biomedical terms and concepts. Complexity will be added in subsequent versions as actual use shows it to be necessary.

UMLS effort is to build “middle-ware” that enables advanced capabilities in many different health information systems.

The Web provides a readily accessible vehicle for distribution of current UMLS fact sheets and documentation, and many new UMLS users now “discover” the UMLS on the Web.

The Impact of the UMLS

Technological advances have improved the methods for investigating UMLS research questions without diminishing the importance of these questions. Over the years, the UMLS Knowledge Sources have matured into significant research and development tools. The UMLS project has also offered the incentive and opportunity for stimulating inter-institutional collaboration.

HL7 has recently selected the UMLS Metathesaurus as an appropriate vehicle for recording and distributing its planned decisions about the vocabularies that are valid for specific parts of the HL7 clinical messaging standard. These two developments reflect the fact that the Metathesaurus provides access to a large and increasing number of important vocabularies in a common and explicit database format.

It has expanded understanding of desirable vocabulary features, including concept organization, multiple hierarchical perspectives, and unique concept identifiers without embedded meaning. The UMLS project has also raised consciousness about the need to represent changes in vocabularies explicitly and about the problems associated with keeping local systems synchronized when changes to externally developed vocabularies occur.
However, in practice, health care providers and EMR system vendors alike confront the difficulties of incorporating elaborate ontologies and vocabularies into clinical workstations and record system clients.

- **Incompatible Representational Formats**: Terminological resources, in general, are developed by an individual or a group of individuals, and are stored using different representations (e.g., text files, XML files, relational databases) depending on various factors ranging from application requirements to ease of use, or just as a matter of preference.

- **Multiple Ontology Modeling Languages**: this leads to redundant, inefficient, often expedient, and frequently semantically distorting efforts to create interfaces to ontologies which proliferate across a multitude of vendors, applications, provider institutions, and even departments or user groups.

- **Lack of Common Tools and Application Programming Interfaces**: Along with multiple ontology modeling languages and formats, multiple and unique tools and application programming interfaces get developed virtually all efforts at interface development had been proprietary or integrally bound with application environments. The consequence was that redundant, insular, noncomparable, nonscalable, and nonincremental terminology-access solutions prevailed.

- **Functional Characteristics**:
  - **Direct concept retrieval**: the ability to retrieve various concept attributes such as definitions or context synonyms when the concept identifier is known.
  - **Associated text-based concept retrieval**: the ability to retrieve concept attributes when one knows only terms or phrases that are similar to the concept sought. This is a “smart” term look-up.
  - **Hierarchy traversal**: the ability to navigate relationships between concepts (e.g., finding sub- or super-concepts of a given concept), or to recursively return all the “children” of a given concept (e.g., all drugs in the penicillin family).
  - **Metadata access**: the ability to retrieve information about versions, allowed relationships, languages, and other capabilities of a particular ontology.
  - **Complex queries**: the ability to access information combinatorially.

- **Performance Characteristics**:
  - **Static performance**: how much time does it take to respond to a single request, absent competing user system load?
  - **Dynamic performance**: how a system responds to different loads? Does the system behave in a predictable fashion when subjected to different loads?
- **Architectural Characteristics:**
  - **Fault resilience:** can services deployed in a data-grid fashion sustain hardware failure and recover from software failure without crashing?
  - **Security:** can the services prevent unauthorized alteration or disruption of content?
  - **Distributed:** how rapidly can the changes introduced into a master server be synchronized to slave/local copies along a chain of distribution?
  - **Federated:** is it possible to maintain cross linkages between and among components of a single, large terminology or related terminologies with cross-referenced content?
  - **Platform and language independence:** can the services be operated independent of the programming language used for implementation, operating system and hardware platform?

The key principle of LexGrid is to accommodate multiple vocabulary and ontology distribution formats and support of multiple data stores for federated vocabulary and ontology access. This is facilitated via a common terminology model that defines how ontologies are formatted and represented programmatically, and is intended to be flexible enough to accurately represent a wide variety of ontologies developed in different languages and other lexically based knowledge resources.

The LexGrid model provides the semantic basis for the development of consistent and rich APIs to access multiple vocabularies that support lexical search queries, hierarchy navigation, and various other features. Existing implementations include the LexBIG API as well as a reference implementation of the HL7 Common Terminology Services (CTS) ([http://informatics.mayo.edu/LexGrid/downloads/CTS/specification/ctsspec/cts.htm](http://informatics.mayo.edu/LexGrid/downloads/CTS/specification/ctsspec/cts.htm)) specification that provides programmatic access via Java, Web, and Grid services.

**Related Work**

**OpenGALEN**
OpenGALEN ([http://opengalen.org](http://opengalen.org)) provides access to the GALEN terminological content which is a computer-based multilingual coding system for medicine.

**Lexicon Query Services**
The Lexicon Query Service (LQS), published by the Object Management Group (OMG), specifies a set of Interface Description Language (IDL) interfaces to be used in querying and accessing computerized medical terminology services.

**Health Level 7 (HL7)**
Health Level 7 (HL7) ([http://www.hl7.org](http://www.hl7.org)) is an ANSIaccredited standards organization that is focused on creating standards for the exchange, management and integration of data that support clinical patient care and the management, delivery and evaluation of health care services.

CTS 2 is an Application Programming Interface(API) specification that is intended to describe the basic functionality that will be needed by HL7 Version 3 software implementations to query and access terminological content.

**Apelon Distributed Terminology System**
The Apelon Distributed Terminology System (DTS) is an integrated set of open-source components that provides comprehensive terminology services in distributed application environments.
**The UMLS Knowledge Source Server**
The UMLS project develops "Knowledge Sources", namely the UMLS Metathesaurus, the Semantic Network, and the Specialist Lexicon, that can be used by a wide variety of applications to overcome retrieval problems caused by differences in terminology and the scattering of relevant information across many databases.

**Semantic Web Technologies**
The Semantic Web is an extension of the existing World Wide Web where semantics of the information and services on the Web are defined, making it possible for the Web to understand and satisfy the requests of people and machines to use Web content

- *Protégé Ontology Development Environment*: Protege\textsuperscript{12,31} is an extensible, platform-independent environment for creating and editing ontologies and knowledge bases

- *The OWL API*: the OWL API\textsuperscript{14,15}, formerly known as the WonderWeb API, is a Java interface and implementation for the W3C web ontology language (OWL), and provides features similar to the Protege OWL API

**The LexGrid Terminology Model**

The LexGrid Model is a community-driven proposal for the standard storage of controlled vocabularies and ontologies whose current development is being coordinated by the Mayo Clinic.

It provides the core representation for all data managed and retrieved through the API, and can represent vocabularies provided in numerous source formats

this common terminology model is a critical component of the LexGrid project

**Key Elements of the LexGrid Model**
LexGrid is based on a model-driven architecture. The master LexGrid model is maintained in XML Schema which is then transformed into UML and other representational forms for publication.

*Service*: A service represents an access point for a collection of coding schemes and/or value domains and their accompanying history and provenance information

*Coding Scheme*: A coding scheme represents particular version of a lexicon, code set, classification system, thesaurus, ontology, mapping or other terminological resource in a uniform and consistent fashion

*Properties*: An important aspect of the LexGrid model is its notion of a "property" (see Fig 6, available as an online data supplement at http://www.jamia.org\textsuperscript{}).
The LexGrid model was designed with two primary purposes—(1) to provide a common representation and semantics for terminological entities and attributes that are commonly used across many terminological resources, and (2) to faithfully represent the remainder of the terminological content in a way that remains faithful to the original resource

*Coded Entry*: A typical coding Scheme defines a collection of coded entries (see Fig 7, available as an online data supplement at http://www.jamia.org\textsuperscript{}). A coded Entry has the following characteristics:

1. Every coded Entry is associated with a code or an identifier that is unique within the namespace of the containing coding scheme.
2. Each coded Entry code is intended to represent a class, category, individual or association within the context of the containing coding scheme
**Concepts:** In the context of the LexGrid model, a concept defines a group of individuals that belong together because they share some properties or characteristics.

**Instances:** In addition to the concepts, each coding Scheme may define their members (see Fig 9, available as an online data supplement at http://www.jamia.org). We normally think of these as individuals in our universe of things, and represent them by defining one or more instances containers.

**Relations:** Each code system may define one or more containers to encapsulate relationships between concepts and instances. Each named relationship (e.g., “subClassOf”, “Part of”, “Same as”) is represented as an association within the LexGrid model (see Figs 10 and 11, available as an online data supplement at http://www.jamia.org). Each relations container must define one or more association.

**Available Representations**
The master representation of the LexGrid model is provided in XML Schema Definition (XSD) format and can be accesses from http://informatics.mayo.edu/schema/2008/01/LexGrid/.

**LexGrid Architecture and Implementation**

A *LexGrid Node* represents both software and a backing data store that provide a logical persistence layer for accessing, storing and managing vocabulary content according to a formalized information model described in the previous section.

The *Import Toolkit* provides an API and a set of administration tools to load, index, publish, and manage vocabulary content for the vocabulary server.

The *Export Toolkit* provides an API and set of administration tools to export content in a standard format from a LexGrid node.

**LexGrid Deployment Projects**

**The LexBIG Project**
LexBIG (https://cabig-kc.nci.nih.gov/Vocab/KC/index.php/ LexBIG) is a project that applies the LexGrid vision and technologies to the requirements of the Cancer Biomedical Informatics Grid (caBIG) Community.

The major components of LexBIG include:

- A set of service management programs to load, index, publish, and manage content for the vocabulary server.
- An Application Programming Interface providing Java interfaces to various functions including lexical queries, graph representation and traversal, and NCI change event history.
- A graphical user interface providing access to administrative and API functions.
- Documentation consisting of API JavaDocs, administrator and programmer guides.
- Numerous examples providing sample source code for common vocabulary queries.
- A test suite to validate the software package installation.

**The LexBIO Project**
LexBIO (http://informatics.mayo.edu/LexGrid/index.php? page_lexbio) provides development and support of Lex- Grid-based terminology software (services, persistence, tooling) to the National Center for Biomedical Ontology (NCBO; http://bioontology.org).

**The LexWiki Project**
LexWiki (https://cabig-kc.nci.nih.gov/Vocab/KC/index.php/LexWiki) is a project that is being introduced and applied within the biomedical informatics community for collaborative terminology authoring and curation tasks.

Discussion

An important feature that we plan to incorporate within the LexGrid infrastructure is the ability to reason with ontologies. Ontology reasoning, this implies using a software tool, called a reasoner (e.g., Pellet, Racer, FaCT), to retrieve answers to certain queries, identify logical inconsistencies, integrate and align multiple ontologies and vocabularies, among others. Thus, to provide reasoning services within LexBIG, our objective is to enhance the existing LexBIG API such that it can leverage existing off-the-shelf open-source reasoners such as Pellet.

Session 4: The National Center for Biomedical Ontology

Mark A. Musen,1 Natasha F. Noy,1 Nigam H. Shah,1 Christopher G. Chute,2 Margaret-Anne Storey,3 Barry Smith,4 and the NCBO team
1Center for Biomedical Informatics Research, Stanford University, Stanford, CA USA
2Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN USA
3Department of Computer Science, University of Victoria, Victoria, BC Canada
4Department of Philosophy, University at Buffalo, Buffalo, NY USA

MISSION

In biomedicine, one of the most fruitful approaches has been to use software tools and knowledge resources known as ontologies—machine processable descriptions of scientific domains—that can promote the integration of disparate data sources. We have shown that such resources can enable data aggregation, improve search, and allow the detection of new associations that were previously not detectable. It is now possible to demonstrate computationally correlations among genes, diseases, treatments, and outcomes, to use these correlations to efficiently direct research into potentially fruitful areas, and to translate the insights from this research to the practice of medicine.

Achieving these integrative analyses requires software systems that take advantage of the semantics of these areas and that can intelligently negotiate domains and knowledge sources, identifying commonality across systems that use different and conflicting vocabularies, while understanding apparent differences that may be concealed by the use of superficially similar terms.1 An appropriate ontology provides the cornerstone of software for bridging systems, domains, and resources.2 Ontologies are the foundation of all semantic technologies in e-science, and are a critical component of multi-disciplinary and translational research in biomedicine.

OUTPUTS OF THE CENTER

Repository of biomedical ontologies
The NCBO’s BioPortal provides access to more than 270 biomedical ontologies and controlled terminologies.
Users come to the BioPortal Web site to browse biomedical ontologies and to search for specific ontologies that have terms that are relevant for their work. BioPortal enables users to navigate ontologies using a standard tree browser. Users also can visualize resources in BioPortal using special tools that offer cognitive support for understanding the complexities of large ontologies.

When users need to understand the relationships between terms in two different ontologies, BioPortal provides mappings between the ontologies to enable direct comparisons. The mappings between ontologies in BioPortal not only allow users to compare the use of related terms in different ontologies, but also allow analysis of how whole ontologies compare with one another. They allow us to identify ontologies that cluster together and to identify the degrees of overlap among ontologies. The mappings in BioPortal form the basis for what we refer to as the NCBO mega-thesaurus.

We have created the system as the nexus of an online community of ontology developers and ontology users who use BioPortal to view, to comment on, and to discuss the content of biomedical ontologies. BioPortal thus adopts Web 2.0 conventions to allow its users to communicate with one another about the NCBO's hosted ontologies in a highly interactive manner. The outcome of these capabilities is that BioPortal offers the equivalent of online, open, community-based peer review for the BioPortal ontology content.

We are developing BioPortal so that computer-based ontology-development tools can access all its content programmatically—including the mappings between ontology terms and the notes about the ontology content contributed by members of the user community.

**Tools and Web services**

**NCBO Annotator.** Perhaps the most widely used tool created by the NCBO is one that maps arbitrary keywords and natural-language text to standardized ontological terms.

**NCBO Resource Index.** A common use of the NCBO Annotator is to ascribe ontological terms to the textual metadata that are associated with experimental data sets. The result is an enormous database of all the terms (and abstractions of those terms) that relate to the textual metadata (or text descriptions) found in a growing set of online data resources (such as the microarray data sets in the Gene Expression Omnibus or the individual protein descriptions in UniProt).

The Index offers the biomedical community a common interface for information retrieval, linking the dozens of ontologies in BioPortal to dozens of biomedical data resources.

**NCBO Ontology Recommender Service.** The NCBO Ontology Recommender Service takes as input representative textual data relevant to a domain of interest and returns as output an ordered list of ontologies available in BioPortal the terms of which would be most appropriate for annotating the corresponding text.

**NCBO Lexicon Builder.** Users frequently turn to the terms of biomedical ontologies to create the *value sets* that constitute the basis of "pick lists" that allow users to make selections from menus when filling in computer-based forms.
FUTURE GOALS
A major initiative of the NCBO in the coming years will concentrate on ensuring the scalability of our technology

Session 4: PHIN VADS - CDC Vocabulary Server


PHIN VADS Series Lectures Given by CDC

http://www.cdc.gov/phin/activities/vocabulary.html

PHIN Vocabulary

PHIN Vocabulary Standards is a key component in supporting the development and deployment of standards-based public health information systems. PHIN Vocabulary Services seeks to promote the use of standards-based vocabulary within PHIN systems and foster the use and exchange of consistent information among public health partners. The use of PHIN Vocabulary Standards ensures that vocabularies are aligned with PHIN standards and with appropriate industry and Consolidated Health Informatics Initiative (CHI) vocabulary standards. These standards are supported by the PHIN Vocabulary Access and Distribution System (VADS) for accessing, searching, and distributing standards-based vocabularies used within PHIN to local, state and national PHIN partners. It promotes the use of standards-based vocabulary within PHIN systems to support the exchange of consistent information among Public Health Partners.

PHIN Vocabulary Services strives to enable the consistent and accurate representation of information by encouraging and supporting the use of Vocabulary Standards to promote semantic interoperability among public health systems. Working with Standard Development Organization (SDOs) terminology experts and other subject matter experts, PHIN Services actively participates in the development and identification of vocabularies important to the public health arena. SDOs include Systematic Nomenclature for Medicine (SNOMED), Logical Observation Identifiers Names and Codes (LOINC), Health Level 7 (HL7), Consolidated Health Informatics Initiative (CHI), Health Information Technology Initiative (HITSP) and American Health Information Community (AHIC).

PHIN Vocabulary Access and Distribution System (PHIN VADS)

PHIN VADS is a web-based enterprise vocabulary system for accessing, searching, and distributing vocabularies used within the PHIN.
Session 4: PHIN VOCABULARY ACCESS AND DISTRIBUTION SYSTEM (VADS)

CODE SYSTEM REPRESENTATION IN PHIN VADS

1/30/2009

http://www.cdc.gov/phin/activities/vocabulary.html

https://phinvads.cdc.gov/vads/DownloadCodeSystemRepresentation.action

Session 5: Chapter 6 HL7 Version 2

T. Benson, *Principles of Health Interoperability HL7 and SNOMED*, HI,
DOI 10.1007/978-1-84882-803-2_6, © Springer-Verlag London Limited 2010

HL7 Version 2 is the most widely used healthcare interoperability standard in the world. It is used in over 90% of all hospitals in the USA and is widely supported by healthcare IT suppliers worldwide.

During its long development period the scope of HL7 Version 2 has increased enormously, but the basic principles have hardly changed. The Version 2.6 standard now has 1,965 pages and 717,000 words. It contains an enormous amount of knowledge and experience about health informatics. One of the basic principles of HL7 V2 has been the preservation of backward compatibility, while the standard has evolved by addition. The idea being that a system, which can understand a new message in a new version, should also be able to understand a previous version. Ideas, which have been superseded, are flagged as being deprecated, but not replaced.
Message syntax describes the overall structure of messages and how the different parts are recognized. Each message is composed of segments in specified sequence, each of which contains fields also in a specified sequence; these fields have specified data types. Data types are the building blocks of the fields and may be simple, with a single value, or complex, with multiple components. These components themselves have data types, which can be simple or complex, leading to subcomponents (Fig. 6.2).
6.1 Message Syntax

HL7 V2 messages are sent in response to trigger events. The message name is derived from the message type and a trigger event. The message type is the general category into which a message fits.

The full message name is ADT^A01 (the “^” is the HL7 field component separator). The message name is always entered in the ninth field of the message header segment (MSH-9).

Each HL7 V2 message comprises a set of segments. The overall structure and allowable content of each message is defined in an abstract message syntax table, which lists segments in the order in which they occur.

The abstract message syntax also shows which segments are optional and which can be repeated. A segment listed on its own is mandatory and may not repeat. Optional segments are surrounded by square brackets [...]. Segments that are allowed to repeat are indicated using curly braces {...}. If a segment is both optional and repeatable, it has both brackets and braces [{...}]. Note that the order is not important: [{...}] and {...} are equivalent.

6.1.1 Delimiters
Delimiters (such as field separators, component separators, and subcomponent separators) are used to indicate the boundaries between these elements. The term element is used to refer to a field, a component, or a subcomponent.

Most HL7 V2 implementations use default encoding with the delimiters to terminate segments and to separate components and subcomponents. The delimiters are defined in the first two fields of the MSH segment (MSH-1 and MSH-2). There is also an XML representation (not described here).

Symbol Usage

- Field separator
- Component separator
- Repetition separator
- Escape character
- Subcomponent separator
- Segment terminator

The field separator (|) is always the fourth character of each segment. Fields are named according to their sequential position within a segment. For example, MSH-9 is the ninth field in the MSH segment and is preceded by nine field delimiters. Two adjacent field separators (||) indicate an empty field. If an application wishes to state that a field contains null and expects the receiving system to act on this, then an explicit null is represented as |‖‖|.

The component separator (^) separates the components of a field. Components are referred to by the segment, field, and position in the field (e.g., MSH-9.1). For example, the MSH-9 field contains two components: MSH-9.1 (message type) and MSH-9.2 (trigger event) and might be represented as ADT^A01. The field separator truncates any components, not needed at the end of a field. For example, the following two data fields are equivalent: |ABC^DEF^^| and |ABC^DEF|.

The repetition separator (~) is used to separate the first occurrence or repetition of a field from the second occurrence and so on. The escape character (\) is used mainly in text elements to bracket text for special processing. The escape character can be used to send delimiters within a message. Symbol Escape sequence

- |F\^
- ^\S^
- ~\R^
- \E^
- & \T e.g., |Marks \T Spencer|

The escape character may also be used to indicate certain formatting commands, such as \.br\ to indicate line break, or \.sp 3\ to skip 3 spaces in the formatted text (FX) data type. The subcomponent separator (&) is used to separate subcomponents within components, providing an additional level of granularity. Each segment is ended with an ASCII carriage return < CR > character.

6.2 Segment Definition

Each segment is defined in a table such as that shown below for the MSH Message Header segment. All HL7 V2 messages begin with a single MSH segment and this provides an example of how segments are defined.
<table>
<thead>
<tr>
<th>SEQ</th>
<th>LEN</th>
<th>DT</th>
<th>OPT</th>
<th>RP/#</th>
<th>TBL#</th>
<th>ITEM #</th>
<th>ELEMENT NAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>ST</td>
<td>R</td>
<td></td>
<td>00001</td>
<td></td>
<td>Field Separator</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>ST</td>
<td>R</td>
<td></td>
<td>00002</td>
<td></td>
<td>Encoding Characters</td>
</tr>
<tr>
<td>3</td>
<td>180</td>
<td>HD</td>
<td>O</td>
<td></td>
<td>00003</td>
<td></td>
<td>Sending Application</td>
</tr>
<tr>
<td>4</td>
<td>180</td>
<td>HD</td>
<td>O</td>
<td></td>
<td>00004</td>
<td></td>
<td>Sending Facility</td>
</tr>
<tr>
<td>5</td>
<td>180</td>
<td>HD</td>
<td>O</td>
<td></td>
<td>00005</td>
<td></td>
<td>Receiving Application</td>
</tr>
<tr>
<td>6</td>
<td>180</td>
<td>HD</td>
<td>O</td>
<td></td>
<td>00006</td>
<td></td>
<td>Receiving Facility</td>
</tr>
<tr>
<td>7</td>
<td>26</td>
<td>TS</td>
<td>O</td>
<td></td>
<td>00007</td>
<td></td>
<td>Date/Time Of Message</td>
</tr>
<tr>
<td>8</td>
<td>40</td>
<td>ST</td>
<td>O</td>
<td></td>
<td>00008</td>
<td></td>
<td>Security</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>CM</td>
<td>R</td>
<td></td>
<td>00009</td>
<td></td>
<td>Message Type</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>ST</td>
<td>R</td>
<td></td>
<td>00010</td>
<td></td>
<td>Message Control ID</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>PT</td>
<td>R</td>
<td></td>
<td>00011</td>
<td></td>
<td>Processing ID</td>
</tr>
<tr>
<td>12</td>
<td>8</td>
<td>ID</td>
<td>R</td>
<td>0104</td>
<td>00012</td>
<td></td>
<td>Version ID</td>
</tr>
<tr>
<td>13</td>
<td>15</td>
<td>NM</td>
<td>O</td>
<td></td>
<td>00013</td>
<td></td>
<td>Sequence Number</td>
</tr>
<tr>
<td>14</td>
<td>180</td>
<td>ST</td>
<td>O</td>
<td></td>
<td>00014</td>
<td></td>
<td>Continuation Pointer</td>
</tr>
<tr>
<td>15</td>
<td>2</td>
<td>ID</td>
<td>O</td>
<td>0155</td>
<td>00015</td>
<td></td>
<td>Accept Acknowledgment Type</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td>ID</td>
<td>O</td>
<td>0155</td>
<td>00016</td>
<td></td>
<td>Application Ack Type</td>
</tr>
<tr>
<td>17</td>
<td>2</td>
<td>ID</td>
<td>O</td>
<td></td>
<td>00017</td>
<td></td>
<td>Country Code</td>
</tr>
<tr>
<td>18</td>
<td>6</td>
<td>ID</td>
<td>O</td>
<td>Y3</td>
<td>0211</td>
<td>00692</td>
<td>Character Set</td>
</tr>
<tr>
<td>19</td>
<td>60</td>
<td>CE</td>
<td>O</td>
<td></td>
<td>00693</td>
<td></td>
<td>Principal Language Of Message</td>
</tr>
</tbody>
</table>

The columns of this table show:
- SEQ: Field sequence number
- LEN: Maximum field length
- DT: Data type
- OPT: Optionality: R (required), O (optional), C (conditional), B (deprecated but retained for backward compatibility)
- RP/#: Repeatable field. If "y" can repeat any number of times; a number indicates the maximum number of repeats
- TBL#: The reference number of the HL7 table which contains a controlled vocabulary from which values can be taken
- ITEM#: HL7's internal database item number
- ELEMENT NAME: Human readable name of the field
6.2.1 Message Header MSH
The report header (MSH) contains common metadata found in most messages, irrespective of subject. The first two fields of the MSH segment specify the delimiters used (see above).

6.2.2 Patient Identification Details (PID)
PatientID refers to the patient identifiers (one or more), which are used by the healthcare facility to uniquely identify a patient (e.g., hospital number, NHS number).

6.2.3 Patient Visit (PV1)
The PV1 (patient visit) segment is used in this example for both the patient’s GP and the patient location at which the sample was taken.

6.2.4 Request and Specimen Details (OBR)
The laboratory allocates each specimen an accession number, which is used to identify that specimen and any derivatives. In HL7 this is referred to as the Filler Order Number and is provided in field OBR-3.1.

6.2.5 Result Details (OBX)
Each separate result is entered as a separate OBX segment, which relates to a single observation or observation fragment. It represents the smallest indivisible unit of a report. Each result is represented as an attribute–value pair.
6.2.6 Z-Segments
HL7 V2 provides a facility for any users to develop their own segments, message types, and trigger events using names beginning with Z. Z-segments are widely used and are one of the main reasons why there are so many different variants of HL7 V2 messages.

6.3 Data Types
Data types are the basic building blocks used to construct or constrain the contents of each element. Every field, component, and subcomponent in HL7 V2 has a defined data type, which governs the information format in the element, what sub-elements it can contain and any vocabulary constraints. Some data types are Simple others are Complex. HL7 V2 has 89 data types in all, but most applications use only a small number of common data types.

The abstract syntax of the HL7 V2 message is:
MSH Message header
PID Patient Identification Details
PV1 Patient Visit
106 6 HL7 Version 2
OBR Results header
{OBX} Results detail (repeats)
All segments are required.
The structure of an HL7 V2 message which meets these requirements is:
MSH|delimiters||sender|||dateTime||messageType|messageID
|processingStatus|syntaxVersion
PID|||patientID^^^source^IDtype||familyName^givenName||dateOfBirth|sex|||streetAddress^addressLine2^^^postCode
PV1|||patientLocation|||||patientsGP
OBR|||accessionNumber|testCode^testName^codeType|||specimenDate||||||specimenSource^^^bodySite^siteModifier
|requester
OBX||valueType|observableCode^observableName|observationSubID|valueCode^valueText^valueCodeType|||abnormalFlag|||result status
OBX ...
A populated example is:
MSH|^~\|^123457^Labs|||200808141530||ORU^R01|123456789|P|2.4
PID||123456^^SMH^PI||MOUSE^MICKEY||19620114|M||14 Disney Rd^Disneyland^MM1 9DL
PV1||5N|G123456^DR SMITH
OBR||54321|666777^CULTURE^LN||20080802||SW^^^FO
OT^RT|C987654
OBX||CE|0^ORG|01|STAU|||||F
OBX||CE|500152^AMP|01|||R|||F
OBX||CE|500155^SXT|01|||S|||F
OBX||CE|500162^CIP|01|||S|||F
This could be rendered as:
Report from Lab123457, 15:30 14-Aug-2008, Ref 123456789
Patient: MICKEY MOUSE, DoB: 14-Jan-1962, M
Address: 14 Disney Rd, Disneyland, MM1 9DL
Specimen: Swab, FOOT, Right, Requested By: C987654,
Location: 5N
Patients GP: Dr Smith (G123456)
Organism: STAU
Susceptibility: AMP R
SXT S
CIP S

Session 5: Health Level Seven Version 2.x Message Profiling Specification Version 2.2

November 30, 2000
Copyright © 2000, by Health Level Seven, Inc.
HL7 V2.x Message Profiling provides a guideline for documenting particular uses of HL7 messages. A defined V2.x message profile will be registered with HL7 and may be reused by other HL7 users, moving the HL7 V2.x standard closer to “plug and play” interfaces.

With consistent and complete V2.x Message Profile documentation, HL7 V2.x interface partners explicitly understand:

- What data will be passed
- The format in which the data will be passed
- The acknowledgement responsibilities of the sender and receiver.

**What is an HL7 V2.x Message Profile?**

An HL7 V2.x Message Profile is a precise and unambiguous specification of a standard HL7 message that has been analyzed for use within a particular set of requirements. It is a particular style or usage of a standard HL7 message, driven by use case analysis and interaction modeling.

An HL7 V2.x Message Profile defines both the **static** structure and content of the message and the **dynamic** interaction, which involves the communication of the message from the sending application to one or more receiving applications.

HL7 V2.x Message Profiles must consist of the following components:

- **Use Case Model** - this may be a use case diagram supported with text or just a textual description
- **Static Definition** – consisting of Message Level Profile, Segment Level Profile, and Field Level Profile
- **Dynamic Definition** – consisting of an Interaction Model and Dynamic Profile

**HL7 V2.x Message Profile Components**

**Use Case Model**
A Use Case Model (Figure 2.1) documents the scope and requirements for an HL7 V2.x Message Profile or set of Message Profiles. The model includes a diagram and detailed text.

The Use Case Model must:

- Provide a **name** that clearly and concisely defines the exchange
- Define the **actors**, including the sending and receiving applications
- Define the **responsibilities** of these actors
- Document the **situations** in which the exchange of a particular HL7 Message Profile is required
- Document the **purpose** for each message exchange
Static Definition of an HL7 V2.x Message Profile

The static definition of an HL7 V2.x Message Profile is directly associated with its corresponding message in HL7 V2.x Standard. A complete HL7 V2.x Message Profile shall be defined at the message, segment, and field levels.

A static profile identifies only those specific elements of a standard HL7 message that are used in the exchange.

A static profile removes all instances of optionality, defining explicitly:

- Segments, segment groups, fields and components
- Cardinalities
- Value sets and coding systems.

As Figure 2.2 depicts, think of the HL7 Message Profile as an overlay of the HL7 Message Structure that is further constrained. For example, where the HL7 Message Structure shows unlimited number of NK1 Segments, the HL7 Message Profile allows for only three repetitions. Additionally, fields that are optional in the HL7 Message Structure may be required within the HL7 Message Profile.

Message Level Profile

The set of segments included within the message of an HL7 V2.x Message Profile shall be defined.

Any segments that are required by HL7 shall be included
Some segments within HL7 Standard Messages are allowed to repeat. The cardinality of all the segments within the message shall be defined.

- The **minimum** cardinality shall be specified.
- Where known, the **maximum** cardinality shall also be specified, or specified as unlimited by using the asterisk symbol (*).
- Allowable Values:
  - [0..0] - Segment not Used
  - [0..1] - Segment is Optional, but can only have one Occurrence
  - [1..1] - Segment is Required, only one Occurrence
  - [0..n] - Segment is Optional, or may repeat n times.
  - [1..n] - Segment is Required, and may repeat up to n times
  - [0..*] - Segment is Optional, or may repeat unlimited number of times
  - [1..*] - Segment is Required, and may repeat unlimited number of times

**Segment Level Profile**

The set of fields of each instance of an HL7-defined segment within the HL7 V2.x Message Profile shall be specified.

The result of this definition is a **segment profile** (Figure 2.4). If a segment occurs multiple times within a message profile, it may be represented by different segment profiles. This shall be explicitly defined within the Message Profile specification.

The segment level profile shall be documented using the **tabular format** employed for the HL7 segment definitions.

- The **length column** shall be updated to accurately reflect the maximum allowed length for the field within this profile.
- The **R/O column** shall be updated to reflect the usage of the field within the particular segment of the message profile (see the following paragraph, **Field Usage**).
• The **RP/# column** shall accurately reflect the maximum number of repetitions of the field allowed for this segment within this message profile.

The **usage** of the field shall be defined using one of the following allowed values:

<table>
<thead>
<tr>
<th>Usage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>R</strong></td>
<td><strong>Required.</strong>&lt;br&gt;A conforming sending application shall provide a valid value for all &quot;R&quot; fields. The value shall be of the specified type and within the range specified for the field. For complete compatibility with HL7, any field designated as “Required” in a standard HL7 message definition shall also be required in all HL7 Message Profiles of that standard message.</td>
</tr>
<tr>
<td><strong>RE</strong></td>
<td><strong>Required but may be empty.</strong>&lt;br&gt;A conforming sending application shall be capable of providing a valid value for all &quot;RE&quot; fields. If the conforming sending application knows the value for this field, then a field value shall be provided of the specified type and within the range specified for the field. If the conforming sending application does not know the value for this field, then the field value shall be specified as empty. For this usage, empty is a distinguished value.</td>
</tr>
<tr>
<td><strong>C</strong></td>
<td><strong>Conditional.</strong>&lt;br&gt;There is a predicate associated with this field that identifies the conditions under which the value of the field shall be specified. The predicate must be based on other field values within this message. This predicate may be expressed as a mathematical expression or in text and may utilize operators such as equivalence, logical AND, and logical OR. The conforming sending application shall evaluate the predicate. If the predicate is satisfied, then the conforming sending application shall provide a value of the specified type and within the range specified for the field. If the predicate is not satisfied, then the field value shall be specified as empty.</td>
</tr>
</tbody>
</table>
Conditional but may be empty.

There is a predicate associated with this field which identifies the conditions under which the value of the field shall be specified. The predicate must be based on other field values within this message. This predicate may be expressed as a mathematical expression or in text and may utilize operators such as equivalence, logical AND, and logical OR. The conforming sending application shall evaluate the predicate.

If the predicate is satisfied and the conforming sending application knows the value for the field, then the conforming sending application shall provide a value of the specified type and within the range specified for the field. If the predicate is satisfied but the conforming sending application does not know the value for this field, the field value shall be specified as empty. If the predicate is not satisfied, then the field value shall be specified as empty.

Not supported.

These fields will not be supported. A conforming sending application will not create a message with a value for these fields. A conforming receiving application will not obtain the value of this field contained within the message. In the case of HL7 V2.x Encoding Rules, these fields are expected to be empty.

Field Level Profile

Each individual field within a segment shall be completely defined to eliminate any possible ambiguity.
The allowed value set for many fields within the HL7 V2.x Standard is specified as user-defined or containing only HL7 suggested values.

In these cases, the **exact allowed value set** shall be specified. These values shall be defined by agreement between the sending and receiving application vendors.

**Coded Entry (CE)** type fields are specified as being populated based on coding systems. For each of these fields, the specific coding system used shall be identified. (See Figure 2.6 for an example of a CE type field.)

Many fields in HL7 V2.x are defined to be **Composite Data (CM)** types. Each component within these composite fields shall be profiled. This requires defining the usage, length, data type, and cardinality of each of the components. Where there are sub-components of a component, each of the sub-components shall also be profiled using the same criteria.

**Dynamic Definition of an HL7 V2.x Message Profile**
The dynamic definition of an HL7 V2.x Message Profile identifies the acknowledgment mode supported for the interaction between the sending application and the receiving application(s).

**Interaction Model**
An Interaction Model (Figure 2.7) shall be included with the HL7 V2.x Message Profile dynamic specification. This model defines specific interactions between the applications that support message profile communication requirements.

The Interaction Model includes **interaction diagrams** that illustrate the sequence of trigger event and resulting message flows between the sending and receiving applications.
Figure 2.7
Interaction Model Example – ADT^A01
**Dynamic Profiles**
The specific HL7 acknowledgments required and/or allowed for use with the specified static definition of the HL7 V2.x Message Profile shall be defined. Specifically, the dynamic profile shall identify whether an accept and/or application level acknowledgment is allowed or required.

The dynamic profile shall define the **conditions** under which an accept and/or application level acknowledgments is expected.

Allowed conditions include:

- Always
- Never
- Only on success
- Only on error.

The specific success or error conditions must be specified.

**Identification of HL7 Message Profiles Using ASN.1**
HL7 Message Profiles shall be uniquely identified with static and dynamic profile identifiers.

The **sending application** uses the profile identifiers to determine the specific HL7 Message Profile to send.

The **receiving application** uses the profile identifiers to determine:

- Which HL7 Message Profile it has received
- What data to expect from the sending application
- Its responsibility as a receiver.

**Static Profile Identifier**
The static profile identifier is a means to uniquely identify a message profile. The static profile identifier is expressed as an ASN.1 Object Identifier (OID).

```
{ joint-iso-ccit(2) country(16) US(840) organization(1) hl7(113883) v2.3(5) static-profile(1) send(2) at(1) null(0) null(0) vl(1) }
```

For efficient communication, the identifier may be specified using the integer form:

```
[ 16 840 1 113883 5 1 1 0 0 1 ]
```

**Figure 2.8**
Static Profile Identifier Example

Note: The tree structure below is not a representation of the Figure 2.8 example.
Dynamic Profile Identifier

The dynamic profile identifier is a means to uniquely identify the dynamic aspects of a message profile.

The dynamic profile identifier is expressed as an ASN.1 Object Identifier (OID).

For efficient communication, the identifier may be specified using the integer form:

\[2 16 840 1 13883 5 2 6\]
Session 5 : Critical Factors Influencing Hospitals' Adoption of HL7 Version 2 Standards: An Empirical Investigation

Chi-Hung Lin & I-Chun Lin & Jin-Sheng Roan & Jehn-Shan Yeh

DOI 10.1007/s10916-010-9580-2

Adopting the HL7 standard could reduce the complexity of communication interfaces employed by different systems. However, the adoption of a standard inevitable also has some impacts on the existing practices of an organization. This might explain the slow progress of HL7 adoption in Taiwan.

Prior studies on HL7 or healthcare information exchange focused largely on technical challenges, such as designing the algorithm and the interface engine for parsing messages, or establishing the middleware for exchanging the specific healthcare information (2,9–11).

Therefore, an increasing number of hospitals had implemented innovative technologies and medical information-related standards, mainly to improve the delivery of their healthcare services and operational performance, especially in terms of heterogeneous information exchange. HL7 is one of the well-known standards for this type of text and numeral-related data exchange. Future e-hospitals will integrate all the stakeholders into a seamless network, allowing data to be shared (1). HL7 provides an important foundation for future electronic medical records (EMR) implementation and system integration. The adoption of HL7 is also closely related to the development of EHR.

In Version 2.X, HL7 uses a message as vehicle and communicates via the Electronic Data Interchange (EDI). Every medical-related behavior represents a trigger event, and was then transformed into a message.

Health information standards (such as HL7 and DICOM) are key to the U.S. and other countries in quest of creating an aggregated, patient-centric electronic health record, interchanging data among independent sites that were involved in a person’s care, and creating a population database for health or infectious disease surveillance and for bioterrorism defense.

Rogers’ Innovation Diffusion Theory (IDT) is applied extensively to comprehensively explain and predict the adoption and diffusion of a new technology (21). Moreover, Rogers (2003) pointed out that, when an organization intends to adopt an innovation, it has to take its attributes into account, as well as other factors such as top management, organization and environment (17,22).

Characteristics of the environment and technology According to the IDT, there are five major innovation attributes: relative advantage, complexity, compatibility, trialability, and observability (20,22). To accept an innovation is for its potential accepter to recognize these attributes. Previous studies defined these five innovation attributes as the critical indicators of the degree of innovation acceptance.
Moore and Benbasat (1991) enhanced explanation power of the IDT in the context of IS adoption (25). They deleted the trialability, and proposed an addition of 4 variables, namely image, voluntariness, result-demonstration, and visibility. In addition, result-demonstration and visibility were separated from observability.

Therefore, we included three factors related to environmental characteristics, which are governmental, industrial, and vendor/association considerations. Regarding the characteristics of technology (HL7), we listed three factors: complexity, compatibility, and security.

Discussion

Five factors, namely environmental pressure, system integrity, top management attitudes towards HL7, staff's technological capability, and hospital scale, were identified to influence a hospital's HL7 adoption decision. Environmental pressure represented intra-industrial competitions and governmental involvement. When the main competitor or the majority of competitors had implemented or prepared to adopt HL7, it would create an intra-industrial competition that promoted a hospital's decision to adopt HL7.

In general, hospitals seem reluctant to adopt HL7, because implementing it means the devotion of more time and effort in order to modify the system or organize training programs, creating more work and pressure. Moreover, some of the hospitals expressed that they were totally not aware that the government was advocating HL7, and they did not understand related governmental policies and implementation procedures.

Most respondents expressed that the implementation of HL7 was incompatible with their existing IT architectures regarding hardware, software, applications or networks and introducing HL7 into the exiting practice was complex for their IS staff.

Therefore, in the case of HL7 implementation, if it showed higher compatibility and lower complexity, and aided by sufficient IS staff and staff's IS capabilities, a hospital was more likely to accept it.

Moreover, this research also found top management attitudes toward HL7 a critical factor. This echoed prior IDT-related studies (such as (20,22,37,49)). The top managers (decision makers) were more willing to adopt and accept a new technology if they were more knowledgeable about IT.

In general, abundant resources increase the chance of innovative technology being adopted. Therefore, higher hospital scale does increase the likelihood of a hospital to adopt HL7 as evidenced in our study.

---

**Session 5: Data Interchange Standards in Healthcare IT—Computable Semantic Interoperability: Now Possible but Still Difficult, Do We Really Need a Better Mousetrap?**

*Charles N. Mead, MD, MSc*

*Journal of Healthcare Information Management — Vol. 20, No. 1*
The absence of a robust set of standards to resolve data incompatibility issues is becoming increasingly costly to the U.S. healthcare delivery system.

The V3 Toolkit is presented as an instance of a class of information modeling, data definition techniques and tools required to solve the increasingly complex challenges presented by today's clinical information system environment. The V3 Toolkit is designed to enable unambiguous data exchange and to thereby facilitate meaningful solutions to the data incompatibility problem.

**Definition and Terms**

Unambiguous data exchange is described most succinctly as computable semantic interoperability, or CSI.

**Syntax vs. semantics.** Syntax is structure; semantics is meaning. To illustrate the difference of the two concepts consider the following sentences as examples: “The dog eats red meat.” “The dog sings blue trees.” The two sentences have identical syntaxes or structures; they start with an article, then have a noun subject, verb, modifying adjective and direct object noun. Their semantics or meanings are different; the first one makes sense, while the second one is nonsense.

**The semiotic triangle.** The word semiotic means, “pertaining to signs or symbols.” Originally formalized as a framework for understanding human communication, the Semiotic Triangle (See Figure 2) forms the crux of any meaningful understanding of the challenges of CSI by delineating the difference between the thing, the symbol that is used to refer to the thing, and the meaning or semantics of the thing.

One of the cornerstones of CSI is the use of unambiguous, coded concepts rather than arbitrary symbols such as character strings as the *lingua franca* between machines.

**Datatype.** Datatypes are the fundamental building blocks around which the semantics of a given piece of data are built. Formally, a datatype is fully specified when both its semantics (in other words, its formal meaning) and the set of legal computational operations that can be performed on an instance of the datatype are rigorously specified.
Healthcare requires several complex datatypes to support concepts, such as physical quantity and time (including both events and intervals), as well as datatypes describing coded terms within a terminology, such as coding system name, version, primary code, alternate codes, and others.

**Interoperability.** Interoperability is the ability of two parties, either human or machine, to exchange data or information.

First, syntactic interoperability guarantees the exchange of the structure of the data, but carries no assurance that the meaning will be interpreted identically by all parties.

Next, human or semantic interoperability guarantees that the meaning of a structure is unambiguously exchanged between humans.

Finally, computable semantic interoperability requires that the meaning of data be unambiguously exchanged from machine to machine.

**Messages vs. documents.** In general, messages differ from documents because they are trigger-based and transient, although the data within them may be persistent. In contrast, documents are assumed to be persistent, and therefore subject to long-term management, as well as to carry strong notions of global vs. local authorship, authentication, and human readability.

Healthcare data collection is often document-centric—data are collected in the context of a document, such as history and physical, progress note, appointment registration, and so on. By contrast, healthcare data usage often is data-centric, for example, “I need to see the last three weeks of sodium values on this patient.” Healthcare data often must be extracted from documents and integrated with data from other documents or from non-document sources.

Part of the HL7 V3 Toolkit is a specification for documents called the Clinical Document Architecture, or CDA. The strength of CDA lies in the fact that all CDA document instances are derived from the V3 Reference Information Model, or RIM, the same model from which all non-document message structures also are derived.

**Semantic scalability.** A process is scalable when what works for 100 instances also will also work, with possible linear increases in applied resources, for 1 million instances.

One of the core issues with HL7 V2 is that although specific HL7 V2 messages may be semantically scalable, HL7 V2 in general is not.

A lack of semantic scalability also is demonstrated when XML is used as a simple solution to achieve CSI. Thus, XML solutions are not guaranteed to be semantically scalable from a CSI perspective.

A cornerstone of HL7 V3’s use of XML is the rigorous definition and publication of the semantics of the XML tags present in all V3 messages. In particular, the critical structural tags are derived from the V3 Reference Information Model. This approach enables V3 to be semantically scalable because the data carried in a specific V3 message and defined by a given instance of XML metadata can be unambiguously understood by all V3-aware systems. The data are locally defined; the metadata are published in the HL7 V3 standard.

**The Four Pillars of CSI**

The motivation behind the creation of V3 was the growing awareness that V2 could not meet the robust requirements for CSI in a semantically scalable, cost-effective manner. In particular, HL7 V2.x message
implementations were becoming increasingly costly to support and evolve as both the number and scope of messages—as well as the number of systems—increased. The limitations of HL7 V2.x become most obvious when data exchange requirements cross inter-enterprise boundaries, thereby exposing conflicts or ambiguities in locally defined data semantics.

**V2:** Lacks a common information model that spans all domains of interest.
**V3:** Offers a common model, the HL7 V3 Reference Information Model, that can span all domains of interest.

**V2:** Lacks a computationally robust datatype specification.
**V3:** Provides machines with unambiguous semantics for each data element transferred, through the V3 Datatype Specification.

**V2:** Lacks a sufficiently robust infrastructure for specifying and binding concept-based terminology values to specific message elements.
**V3:** The HL7 Vocabulary Technical Committee and Modeling and Methodology Committee manage a formal process for interleaving the RIM with various terminology models, as well as enabling the binding of domain-specific terminologies such as SNOMED, LOINC, DICOM, MeDRA, MIAME/MAGE and others to message specifications.

**V2:** Lacks a formal top-down message development process.
**V3:** Does not allow use of a top-down methodology for defining each data interchange structure using only RIM.

**The Reference Information Model**
Although each of the Four Pillars is critically important in achieving the overall goal of CSI, the first—a common information model spanning all domains of interest—is the most visible. More specifically, the HL7 V3 RIM defines a high-level backbone containing five abstract structural concepts:

- **Entity:** Things in the world, including place, organization, material and living subject, either person and non-person.
- **Role:** Capability, capacity or competency, usually time-based.
- **Participation:** Role in the context of an act.
- **Act:** Clinical, administrative or financial definitions, plans, occurrences, and so forth.
- **Act relationship:** The semantics of links between acts.

The semantics of each of these backbone classes is specified through a number of attributes. In turn, the semantics of each attribute is specified through its binding to an HL7 V3 datatype.

The RIM also introduces the concepts of state and mood as attributes of acts to enable CSI-compliant descriptions of complex healthcare processes.

Mood describes the critical phases of a business process through which instances of a concept may pass; for example, a drug or lab test that may be defined in a master service catalogue, ordered for multiple patients, and administered or performed any number of times based on the order. Mood is orthogonal to the more familiar notion of state, the denotation of the phase in the lifecycle of an instance of a concept. A single act instance may pass through many states in the course of its life. However, it may have only one mood, such as define, order or request, event, goal, or others.

basic example of the process of representing healthcare delivery semantics through common structures bound to domain-specific terms

---

*I—General premise.*
II—The common structures

III—An exemplar set of statements

IV—The binding of RIM structures and domain specific terms

The consortium has recommended that the following standards for the national health information network be written into regulations and entered into the Federal Register:

• Interface definitions for health information service providers and NHIN-provided services.
• Service-level requirements.
• Data exchange standards, including syntactic standards defining data interchange structures and methods; semantic standards, defining data meaning with sufficient robustness so data can be understood by all processing machines.

Context of Interoperability

Posted on April 30, 2011 by Grahame Grieve


Interoperability Law #1 says that Interoperability is all about the people.

Interoperability—it’s all about the people. And there’s not that much difference between interoperability and diplomacy.

Drive By Interoperability

I refer to the case of Alice and Bob as “drive-by interoperability”. They both want a clean encounter, with as little interaction with each other as possible. And the hospital does too—conflict between vendors can be a real problem to resolve. Also, due to variation between institutions etc, what works at one institution probably won’t just work elsewhere.

This is the historical context of HL7. That’s what we do: sharing data between two systems where we can take nothing for granted. We don’t assume any rational sharing of any nature between the systems at all, except for the specific interaction that is under consideration. The more interactions you depend on, the more risky the overall implementation.

So that’s the classic background of HL7. Drive By Interoperability.

In a case like this, the HL7 standard is a framework for solving problems. The players want fixed technology—because technological change is expensive, and easy adjustment to particular and peculiar business practices between institutions. v2 gives this—Minimal Lower layer Protocol exchanges, common syntax, loose semantics in the fields, highly adaptable. Fit for purpose

National Programs

That approach was just fine for vendors and small providers. But the problem was that it didn’t scale.

We solve interoperability
HL7 does solve interoperability. We’ve sorted half adjusted to the SOA/large program community by allowing more points of contact and so forth. But in our hearts, we still design the information models for worst case interoperability. You can see that manifesting, for instance, in HXIT attributes in ISO 21090 (subject of yet another post).

Programs that take up HL7 interoperability solutions need to be aware that the closer they get to Peter and George (and the further from Alice and Bob), the less HL7 is going to seem like a good fit off the shelf. That doesn’t mean it’s not useful – just that you need to know how to correct for these assumptions of badness that HL7 implicitly makes.

HL7: Interoperability in the worst case.

Multimedia: Health Level 7 Version 2
Multimedia: Health Level 7 Version 3
Session 6: Enabling Joint Commission Medication Reconciliation

Objectives with the HL7 / ASTM Continuity of Care Document Standard

Robert H. Dolin, MD, Kaiser Permanente, Anaheim, CA (Robert.H.Dolin@kp.org);
Gay Giannone, MSN, RN, Siemens Medical Solutions, Malvern, PA; Gunther Schadow, MD, PhD, Regenstrief Institute, Indianapolis, IN

In this manuscript, we illustrate this point by describing how the HL7 / ASTM Continuity of Care Document (CCD) standard supports and enables Joint Commission medication reconciliation recommendations.

The Institute for Safe Medication Practice (ISMP) issued a series of medication error scenarios in 2005, noting that “each error is the direct result of failed communication about prescribed medications during vulnerable transition points in the continuum of healthcare”. The Joint Commission's recommendations further emphasize that medication errors “typically occur at the interfaces of care - when a patient is admitted to, transferred within, or discharged from a health care facility”. It has been suggested that standards, deployed at these critical junctures, can decrease the risk for error.

new medications are ordered or existing orders are rewritten. [...] This process comprises five steps: 1) develop a list of current medications; 2) develop a list of medications to be prescribed; 3) compare the medications on the two lists; 4) make clinical decisions based on the comparison; and 5) communicate the new list to appropriate caregivers and to the patient”.

CCD is the product of collaboration between HL7 and ASTM, and represents an implementation of the ASTM Continuity of Care Record (CCR) specification as a set of constraints layered atop the HL7 Clinical Document Architecture (CDA) specification. The ASTM CCR is a core data set of the most relevant administrative, demographic, and clinical information facts about a patient’s healthcare, covering one or more healthcare encounters. It provides a means for one healthcare practitioner, system, or setting to aggregate all of the pertinent data about a patient and forward it to another practitioner, system, or setting to support the continuity of care. The HL7 CDA is a document markup standard that specifies the structure and semantics of clinical documents for the purpose of exchange.
Discussion
We have found that CCD supports the data exchange requirements underlying the Joint Commission medication reconciliation recommendations. Minor gaps were identified (e.g. the representation of compliance). In addition, CCD provides fields not specifically identified by Joint Commission, but useful nonetheless when managing medications across transitions of care.

Session 6: The Clinical Document Architecture and the Continuity of Care Record: A Critical Analysis
JEFFREY M. FERRANTI, MD, R. CLAYTON MUSSER, MD,MS, KENSAKU KAWAMOTO, W. ED HAMMOND, PHD

Journal of the American Medical Informatics Association Volume 13 Number 3 May / Jun 2006


Both the HL7 Clinical Document Architecture (CDA) and the ASTM International Continuity of Care Record (CCR) strive to facilitate the interchange of health care data among care providers. The CDA is based on the HL7 version 3 Reference Information Model (RIM), and the CCR is a clinical framework that was originally developed by health care practitioners to meet the information exchange needs of primary care providers. Both technologies use the World Wide Web Consortium standard of Extensible Markup Language (XML) to facilitate the exchange of structured medical data.

The CDA is a complex standard that can be challenging to implement.

The initial version of the CCR had as its strengths a lightweight, easily implemented approach, and it was intended primarily for the exchange of health summaries

The architecture of the current CCR specification appears to overlap with the CDA in both complexity and scope.

The CCR was created by health care practitioners based on their perceptions of the data they wish to share in a given circumstance. Unlike many other standards, clinicians were actively involved in the creation of the CCR and were integral to defining its form and content. It is patient focused and emphasizes the data directly related to a patient’s current medical problems.

ASTM International defines the CCR as a “summary of the patient’s health status (e.g., problems, medications, allergies) and basic information about insurance, advance directives, care documentation, and care plan recommendations.”

“the CCR represents the patient summary, which for many EHRs is called the ‘overview’ of the patient.” In essence, the CCR pools relevant information from multiple medical documents and creates a “snapshot” of the patient.

A CCR document is represented using XML, making it easy to transport and display. At present, it is specified by an XML schema and an implementation guide. The CCR supports the use of coding systems.
Realizing that the effective communication of medical information requires both messaging standards and common data structures, HL7 developed the CDA (R1 in 200016 and R2 in 200517) to provide a common representation for clinical documents through “a document markup standard that specifies the structure and semantics of clinical documents

A CDA document “can exist outside of a messaging context and/or can be a MIME-encoded payload within an HL7 message. Thus the CDA complements HL7 messaging.”17 In essence, each CDA instantiation represents a distinct clinical document.

The CDA is basically a constrained version of the HL7 RIM, in which RIM object classes have been assigned specific data types and vocabularies.18 In HL7 terms, this constraint of the RIM is called a Refined Message Information Model (RMIM). The CDA document type and Universal Observation Identifier Names are defined with LOINC19 document codes. Like the CCR, the CDA allows for controlled terminologies such as SNOMED CT to enhance semantic interoperability between medical information systems.

Comparison of the CDA Version 2 and the CCR Version 1a

The CCR stresses the important pieces of data required to care for a patient, whereas the CDA serves as a document architecture to format all clinical documents. The scope of the CCR is focused on the primary care “summary record” and does not explicitly support other use cases.

In contrast, the CDA was conceived and specifically built to represent virtually any type of medical document, and it adds complexity to cover this generality. It has been suggested that a complete CCR “summary document” could be expressed as a CDA document template. Since the CCR has clear clinical utility and the CDA provides a strong structural backbone, it seems that harmonization of these technologies is the best solution. Since the CDA is intimately linked to the HL7 RIM, its components are applicable across other HL7 standards, and it could add additional functionality and interoperability to the clinically useful CCR.

The CDA is document centric and useful in modeling the complex structure of a multitude of clinical documents. In contrast, the CCR was designed to focus on the data elements rather than the documents.

Both the CDA and the CCR have an unclear relationship with a complete EHR. The CDA was designed to create a wide variety of individual clinical documents; thus, it should be possible to organize and collect these documents in the context of an EHR system. The CCR focuses on the summary record and, as such, it could span multiple encounters and multiple health care providers. Although the initial purpose of the CCR was to represent a patient summary, the CCR can be used to represent a lifelong EHR if the referring physician chooses to include that level of detail. The content of a CCR relies heavily on the opinion of the referring physician. This option may be viewed as either an advantage or a disadvantage.

by design, the CCR does not provide for user-configurable fields. this lack of local extensibility could make it difficult for institutions to tailor the CCR to meet needs beyond the stated purpose of the CCR.

Conclusion
Overall, the CCR is a clinically useful document that has been forged from the ground up to meet a specific need. Its major contribution is capturing the intent of providers and vendors to move data between disparate groups in a human-readable, summary form. Its utility as a clinical tool remains unquestioned, and it promises great advances beyond our current paper-based systems.

the technical implementation of the CCR falls short of the mark because the CCR was designed for a single purpose. Although the CCR meets an important clinical need, users may require the enhanced interoperability offered by the CDA when other needs exist.

If the two standards cannot be directly linked in this way, a second option might be to define a Common Data Element Set to be used in all clinical documents. (In this sense, data elements can be thought of as individual entries, such as fill-in-the-blanks, on a medical intake form.) The elements can be combined and organized to accommodate a wide variety of clinical scenarios. Of course, special care must be taken to assign terminologies to each data element. In this way, CDA documents, CCR records, X12N claim forms, DICOM image reports, and many others might finally be able to communicate together. In the end, the ideal solution is a single standard using content and knowledge from both groups.

Kensaku Kawamoto*,1, Guilherme Del Fiol1, David F. Lobach1 and Robert A. Jenders

The Open Medical Informatics Journal, 2010, 4, 235-244

Despite the potential for CDS to significantly improve health and health care, the reality remains that most clinical care continues to be provided with only minimal CDS support, if at all [1,4]. While there are many reasons for this limited adoption of CDS capabilities, one important reason is the predominant use of non-standard approaches to implementing CDS that are oftentimes specific to a given CDS application and implementation setting [1,5,6]. As a result, CDS capabilities developed at one institution oftentimes cannot be easily transferred to other health care institutions, or even to other types of CDS applications within the same institution [1,7].

One promising approach to enabling the widespread deployment of CDS capabilities is the centralized management of machine-executable knowledge resources, which are then leveraged across multiple care settings by CDS engines interfaced with various health information systems [1,8-11].

2. STANDARDS NEEDED FOR SCALABLE CLINICAL DECISION SUPPORT
At a broad functional level, CDS implementation needs can be classified as consisting of (i) the need to communicate with various systems regarding relevant health care concepts; (ii) the need to create and represent clinical knowledge that can be used to enable automated CDS; and (iii) the need to utilize these clinical knowledge resources to deliver CDS interventions within health information systems such as electronic health record (EHR) systems and computerized provider order entry systems.

A. Need for Standards for Data Representation and Mapping
knowledge sharing has two main forms: knowledge reuse and knowledge transfer. In knowledge reuse, CDS knowledge is employed for a different purpose or application than its original form within its original organization.

Knowledge also may be transferred across organizational boundaries for the same purpose common information models are needed due to the significant challenges associated with mapping CDS knowledge to different, non-standard information models [16-19].

Several types of standards can facilitate the implementation of CDS by providing a common framework for conceptualizing health care. Such enabling standards include standard terminologies and information models for representing health care data, as well as standards for the types of patient data that are expected to be available for CDS within specified contexts such as EHR systems [22]. Moreover, CDS implementations can be aided by the availability of standard approaches for obtaining terminology and ontology inferences, such as the mapping of concepts across terminologies and the identification of concept properties and semantic relationships [23,24].

B. Need for Standards for Knowledge Representation

Thus, the widespread use of one or more common knowledge representation approaches could substantially facilitate the use of the encoded knowledge to provide CDS.

C. Need for Standards for Leveraging Knowledge Resources to Deliver CDS

In providing CDS, one of the most difficult challenges is leveraging machine-executable medical knowledge resources and electronic clinical data to deliver useful CDS interventions within the context of various health information systems.

The heterogeneity of knowledge resources and health information systems makes it difficult to re-use a CDS implementation infrastructure developed for one clinical context within other applications and care settings.

Thus, in order for knowledge resources to be usable in a scalable manner across various applications and institutions, standardized approaches are needed for how machine-executable, implementable knowledge resources are leveraged within health information systems to generate and deliver CDS interventions.

3. CURRENT STANDARDS LANDSCAPE

A. Standard Terminologies. Various standard terminologies are already well-developed [30]. Most of these terminologies are compiled in the National Library of Medicine’s Unified Medical Language System [36], which maps concepts from each of over 100 source terminologies to unique master concepts. Notable standard terminologies included in the Unified Medical Language System include the Systematized Nomenclature of Medicine-Clinical Terms (SNOMED CT) [20], the Logical Observation Identifiers Names and Codes (LOINC) [37], RxNorm [38], and the International Classification of Diseases (ICD) [21].

Thus, the challenge lies less with the lack of relevant standards, but more with the fact that multiple terminologies are in concurrent use, including many non-standard terminologies (e.g., vendor-specific laboratory terminologies). The concurrent use of these various terminologies is a significant challenge, as differences in the granularity of terminologies oftentimes prevent a concept from one terminology from being mapped one-to-one to an equivalent concept in another terminology. As a result, a CDS resource
designed for use in one setting may not be readily usable in another setting that uses a different set of terminologies, even when similar concepts are being captured.

**B. Standard Information Models.** Standard information models are also available for supporting CDS in a scalable manner. Perhaps most prominent among these information models are the models used for HL7 version 2 messages, which are ASCII-encoded messages that communicate various information relevant to the care of a patient [39]. While easy to understand and widely adopted, these models have been the target of substantial criticism. In particular, HL7 version 2 information models have been criticized for their adoption of loose semantics that ultimately compromise interoperability, including the optional use of non-standard coded values and the use of “Z segments” that allows nonstandard data elements to be represented in message instances.

Work has recently begun within the HL7 CDS Work Group to establish a standard for a concept known as the virtual medical record (vMR) [43]. As originally proposed in the literature, a vMR consisted of a standardized definition of both (i) the semantics of information communicated between a CDS engine and a health information system and (ii) the functional capabilities of a health information system available to the CDS engine (e.g., to order a prescription or to deliver an alert) [44].

**Despite these promising standards, important gaps and challenges remain**

First, current information models oftentimes lack sufficiently tight binding to terminologies. For example, a clinical information model may recommend the use of SNOMED CT to represent medical diagnoses but also allow the use of ICD9 or ICD10.

Furthermore, as a second challenge, many of the current generation of information model standards – including many of the HL7 version 3 information model standards – entail much more flexibility and complexity than is required or desirable for CDS.

This difficulty is exacerbated by a lack of tooling and other practical resources to support the widespread use of these information models by health information technology (IT) professionals.

Moreover, much of the information captured in these models is largely irrelevant for CDS applications (e.g., who performed the laboratory test).

Finally, there is little alignment between these standard models and the logical structure or schemata of clinical data repositories in actual use.

**C. Standards for Patient Data Expected to be Available for CDS.**

While it may be unrealistic to have a single standard which is universally adopted, it would be possible to have a small number of standards on expected data availability (e.g., one standard for a full EHR environment, another standard for an environment only with access to claims data, and a third standard for an environment with access to claims data and laboratory data).

**D. Standard Approaches for Terminology Inferencing.**

A critical CDS implementation need is the mapping of clinically relevant concepts to terms used within various health information systems. These terminology needs for CDS typically manifest in two forms: (i) the need to identify which concepts are subsumed by a parent concept, and (ii) the need to translate concepts across vocabularies.
A terminology service can provide these types of terminology inferencing capabilities through a service interface. Standardization of such service interfaces can enable CDS implementations to leverage different terminology services based on their capabilities, rather than being locked into a single solution because of investments in proprietary system-to-system interfaces.

The HL7 Common Terminology Services version 1 [46] and 2 [47] standards provide standard specifications for how various terminology inferencing capabilities can be obtained from a software service. Perhaps more importantly, significant tooling and resources are being developed to support the use of this standard.

E. Standard Representation of Non-Executable Clinical Knowledge

Since 2002, the Guideline Elements Model (GEM) has been available as an ASTM International standard for the representation of the contents of clinical practice guidelines in a structured, non-executable format that is suitable for translation into an executable format [27].

F. Standard Representation of Executable Clinical Knowledge

Several standards are available for representing machine-executable clinical guidelines. These standards include two standards that can be used to represent machine-executable clinical rules. First, the HL7 Arden Syntax standard specifies various aspects of Medical Logic Modules,

Despite these available standards, no standard is widely agreed upon for representing clinical practice guidelines in a machine-executable format.

G. Standard Approaches to Utilizing Executable Clinical Knowledge

Two HL7 standards are currently available for accessing and utilizing machine-executable CDS resources through the calling of an external software service. First, the HL7 Decision Support Service draft standard [33] provides a standard approach to providing structured patient data and receiving structured, patient-specific inferences based on the use of Decision Support Service knowledge modules.

The HL7 Context-Aware Knowledge Retrieval (“Infobutton”) standard [56] provides a specification of how data on a patient, a clinical question, and the relevant clinical context can be passed to an external software service to retrieve back knowledge resources relevant for that specific situation as an attempt to help clinicians and patients fulfill their knowledge needs.

H. Standard Approaches to CDS Delivery

Despite these important standards, significant gaps and challenges still remain for the standard delivery of CDS within health information systems. First, there is a lack of standards for the EHR services upon which many CDS implementations depend.

A further challenge to scalable CDS delivery is the lack of use in health care of standard business process modeling approaches [61]. If widely adopted within health care, standard business process modeling approaches would allow CDS implementers to specify, share, and adapt workflow specifications of human and computer actors relevant in the delivery of CDS.

Finally, a significant challenge to the standardized delivery of CDS is the enormous heterogeneity that currently exists, and is likely to continue to exist, in the health information system infrastructure available across different care settings.

4. DISCUSSION AND RECOMMENDATIONS
First and foremost, we recommend that standards development organizations focus on supporting the operational use of standards relevant for CDS. Currently, many health IT standards suffer from barriers that limit their widespread adoption, including excessive complexity, limited tooling, and poor documentation on how the standards should be used in operational clinical settings.

Measures that can be taken to improve the usability of standards include:

- A focus on implementability and simplification where possible;
- Supporting the development and dissemination of improved tooling;
- And improved documentation on how to make use of the standards.

Second, we recommend that efforts should be made to harmonize and recommend the use of existing and future standards for CDS.

Third, we recommend that health care stakeholders, and in particular the federal government, provide much greater support for the creation and adoption of health IT standards relevant to CDS.

Fourth, we recommend that authors of clinical practice guidelines and other knowledge resources that could be used for CDS publish that knowledge in at least a computable, and if possible a fully implementable, format that incorporates standards and is unambiguously represented.

As a final recommendation for enabling scalable and high-quality CDS, we recommend that the federal government put into place a robust process and funding for developing standards-compliant CDS knowledge resources that are easily accessible to the wider health care community.

---

**Session 6: System-Agnostic Clinical Decision Support Services: Benefits and Challenges for Scalable Decision Support**

Kensaku Kawamoto*, Guilherme Del Fiol, Charles Orton and David F. Lobach

*The Open Medical Informatics Journal, 2010, 4, 245-254*

While health informaticists have demonstrated repeatedly how CDS systems can improve health and health care [4], most patient care continues to be conducted with minimal CDS, if any [3]. Although many factors contribute to this limited availability of CDS capabilities [3], one important reason is the difficulty of sharing machine-executable medical knowledge across applications and care settings.

As an alternative and complementary strategy for knowledge sharing, CDS capabilities – and in particular the analysis of patient data to generate patient-specific inferences – can be encapsulated within independent, system-agnostic software services and then leveraged by various downstream applications.

In this manuscript, such CDS services that provide patient evaluation capabilities independent of specific CDS systems or system deployment settings are referred to as system-agnostic or system independent CDS services. A software architecture based on the coordinated use of such independent business services is known as a service-oriented architecture (SOA).
we have developed a system-agnostic, standards-based CDS service for supporting this approach to CDS implementation across any medical domain. This system-agnostic CDS service is known as SEBASTIAN, and it allows machine-executable medical knowledge to be encoded in discrete modules and then leveraged across applications and care settings [5]. SEBASTIAN is implemented as a Web service, in which the client and server communicate over the Internet using extensible markup language (XML) messages.

the SEBASTIAN service interface has served as the basis of the HL7 Decision Support Service specification [19], which specifies a standard service interface for system-agnostic CDS services and was adopted as an international draft standard in 2006.

3. BENEFITS

Relative Ease and Flexibility of Use. A primary potential benefit of using system-agnostic CDS services is the relative ease and flexibility with which such services can be used to facilitate the implementation of CDS capabilities across information systems and clinical care settings. Moreover, because system-agnostic CDS services are self-contained, designed for external integration, and make minimal assumptions about their deployment context, they can be integrated into various applications with relative ease. Rapid development was possible because SEBASTIAN was able to fulfill a core functional requirement of the CDS systems – namely, the analysis of patient data to generate patient-specific inferences.

Facilitation of Centralized Knowledge Management and Sharing. Another potential benefit of using system-agnostic CDS services is that this architectural pattern can facilitate the centralized management and sharing of machine-executable medical knowledge. Each knowledge resource is potentially applicable across a wide range of CDS deployment settings, thereby making the knowledge resources more conducive to centralized management and sharing across multiple deployment contexts.

Centralized knowledge management has many inherent benefits. These benefits include:

- a reduction in redundant knowledge management efforts, as well as an increased capacity for developing and supporting effective knowledge management processes and tools.
- Furthermore, centralized knowledge management enables an increase in the effort and resources that can be dedicated towards the development of the knowledge base.
- Also, centralized knowledge management can improve quality control due to increased vetting of the knowledge by both the creators of the knowledge as well as by the many users of the knowledge.
- Moreover, centralized knowledge management enables the development of knowledge components that can be reused across many different decision support modules.

Potential to Support Multiple Knowledge Representations and Resources through Common Service Interface. As another important benefit, system-agnostic CDS services have the potential to support a variety of knowledge representation formalisms and knowledge resources through a common service interface.
Improved Simplicity and Componentization (Separation of Concerns).

By explicitly separating out patient data analysis and inferencing as a separate and independent functional component of an overall CDS system, the use of system agnostic CDS services can help to simplify the CDS deployment architecture because a system-agnostic CDS service is a functionally independent system component, it can be loosely coupled with other functionally independent system components and services to fulfill a given CDS need such as point-of-care disease management [16], while at the same time being coupled with a different set of system components to fulfill a different CDS need such as population health management [17].

Easier Testing and Validation. As a corollary to the improved simplicity and componentization just discussed, the modular and functionally independent nature of systemagnostic CDS services can reduce the effort required for testing and validation.

Enabling of Distributed CDS Development. Finally, an additional potential benefit of using system-agnostic CDS services is that this approach can enable the distributed development of application-level CDS capabilities. Whereas the knowledge itself in a CDS service must undergo central coordination and quality control, the development of downstream CDS applications can be conducted in a highly decentralized and distributed manner.

4. CHALLENGES AND POTENTIAL SOLUTIONS

Increased Effort Required for Developing and Supporting Knowledge Resources. Because knowledge resources within a context-independent CDS service are designed to support a variety of CDS applications across multiple clinical contexts, developing and supporting such knowledge resources can require substantially more effort than a similar resource designed to support a specific CDS application in a specific implementation setting.

the terminologies supported by a CDS service, as well as the manner in which terminology services are leveraged, are both important considerations when designing and implementing system-agnostic CDS services.

a reasonable approach would be to support major standard terminologies in conjunction with any non-standard terminologies that are in use by existing clients.

Need for Standardized Service Interface. Another important challenge to the use of system-agnostic CDS services is that the full potential of such services may only be realized if different services use a common interface for communicating with clients. Without such standardization, clients may need to manage multiple, non-compatible interfaces to various CDS services.

The HL7 and OMG Decision Support Service standards [19,20] begin to address this need for standardization by defining a standard interface for system-agnostic CDS services, including a mechanism for defining common service semantics through the use of profiles. The widespread adoption of these standards would significantly facilitate the ability of system-agnostic CDS services to enable CDS on a much larger scale.

Need for Customization. As another potential challenge, the outputs of CDS services may need to be customized to meet the unique needs of client systems in various contexts of use.
In addressing this need for customization, one potential solution is to incorporate customization parameters as inputs into the service.

**Need for Other System Components.** By design, a system-agnostic CDS service fulfills only one of the tasks required for delivering CDS – namely, the evaluation of patient data to generate patient-specific inferences. Consequently, other system components must be established for delivering CDS, including components for interacting with users, retrieving required clinical data, and parsing and presenting CDS service results.

When possible, such system components should be implemented in a manner that allows for efficient maintenance and re-use. One recommended approach for achieving such reuse is to implement needed functionality as independent software services that are leveraged in conjunction with system-agnostic CDS services within a standards-based SOA [6].

**Black-box” Evaluation may be Unacceptable.** Given the potentially high stakes of CDS in many clinical settings, the black-box nature of a CDS service may be unacceptable in certain cases.

One potential solution to this challenge is to provide detailed meta-data on how a clinical decision is reached by a CDS service, including a human-readable description of the clinical guideline or algorithm used and a detailed explanation of how input data are processed to generate the service output.

**Desire for Local Service Installation.** As a practical matter, health care organizations and health information technology vendors may insist on having a local service instantiation. To address such concerns, CDS services can be designed for distributed deployment. In such a distributed deployment model, it is important that processes be established to ensure synchronization of the client instances with the centrally managed CDS service.

**Different Data Availability and Data Models.**

A significant challenge to the use of system-agnostic CDS services is that there can be significant heterogeneity in the information models and terminologies used at different care settings, as well as significant heterogeneity in the types of data available across these settings.

In addressing this challenge, the most desirable long-term solution is to standardize the data expected to be available for CDS, including which information models and terminologies are used to capture such data. Such an effort is currently underway within the HL7 virtual medical record project [31].

**Limited Content Availability.** Finally, an important potential challenge to the use of system-agnostic CDS services is the limited availability of content within such services. Beyond knowledge resources for basic medication inferencing, there are currently only limited medical knowledge resources available through system-agnostic CDS services.

One way to address this challenge would be to create an interoperable, standards-based market for such knowledge, in which knowledge resources from various organizations are available through a common Decision Support Service interface.

5. **BALANCE OF BENEFITS AND CHALLENGES**

6. **CONCLUSION**
Session 6: HITSP Summary Documents Using HL7 Continuity of Care Document (CCD) Component

HITSP Summary Documents Using HL7 Continuity of Care Document (CCD) Component Released for Implementation 20090708 V2.5

1.0 INTRODUCTION
1.1 OVERVIEW

The HITSP Summary Documents Using HL7 Continuity of Care Document (CCD) Component describes the document content summarizing a consumer's medical status for the purpose of information exchange. The content may include administrative (e.g., registration, demographics, insurance, etc.) and clinical (problem list, medication list, allergies, test results, etc) information. Any specific use of this Component by another HITSP specification may constrain the content further based upon the requirements and context of the document exchange. This specification defines content in order to promote interoperability between participating systems. Any given system creating or consuming the document may contain much more information than conveyed by this specification.

This Component is essentially a subset of the healthcare data that has been developed for specific business Use Cases. This subset contains the minimum critical or pertinent medical information sections as specified by the business case. Information conveyed according to the Component Construct is a representative extract of the information available on the creating system. The information in the HITSP Summary Documents Using HL7 Continuity of Care Document (CCD) Component and the creating systems must be consistent. Furthermore there should be no data elsewhere in the creating system that would contradict the meaning of any data in this construct.

2.0 COMPONENT DEFINITION
2.1 CONTEXT OVERVIEW

HITSP Summary Documents Using HL7 Continuity of Care Document (CCD) Component is essentially a subset of the data in an EHR or PHR system that has been developed for interoperability purposes for specific business Use Cases. This subset contains the minimum critical or pertinent medical information of sections and data elements used in those business cases. The information in the subset of data to be exchanged and the EHR or PHR system must be consistent. Furthermore there should be no data elsewhere in the EHR or PHR system that would contradict the meaning of any data in the summary.

2.1.1 COMPONENT CONSTRAINTS

An Interoperability Specification (IS) may constrain this Component to satisfy the requirements of an information exchange scenario. An Interoperability Specification describes the specific context of exchange and may declare additional constraints, e.g., by requiring the presence of information modules that are otherwise described in this Component as optional.
This Component should not be used outside the context of an Interoperability Specification as this may result in loss of interoperability.

However, the modules defined in this Component are intended for use and reuse whenever summary information such as a consumer's patient registration, medical history, immunizations, or other information modules defined in this Component are needed. Individual modules defined in this specification may be reused in other defined document types, such as operative notes, to convey similar information in other Use Cases.

### 2.2 RULES FOR IMPLEMENTING

#### 2.2.1.2 CDA DOCUMENT

At the clinical document level, template identifiers are employed to assert which template(s) the document conforms to. A document may assert conformance to more than one template. Template identifiers for context specific documents are declared in the Interoperability Specifications where the context is defined.

C32-[CT1-19] A CDA Document **SHALL** declare conformance to this specification by including a `<templateID>` element with the root attribute set to the value 2.16.840.1.113883.3.88.11.32.1.

Asserting conformance to this specification via the inclusion of the Summary Document templateID indicates that additional constraints from this specification are followed when applicable.

- Required modules from this specification shall be present and follow the associated constraints
- Modules that have been explicitly prohibited shall not be included
- Optional modules, when present, will follow the associated constraints if that module also asserts conformance to this document, i.e., includes the associated templates
- Additional CCD entry elements (the equivalent to modules in this specification) may be present. The consumer of the document may choose to accept or exclude the additional content, but shall not reject the document solely based upon the presence of the additional content

#### 2.3.2 SELECTED STANDARDS

<table>
<thead>
<tr>
<th>Table 2-5 Selected Standards Standard</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health Level Seven (HL7) Implementation Guide: CDA Release 2 – Continuity of Care Document (CCD), April 01, 2007</td>
<td>The Continuity of Care Document implementation guide describes constraints on the HL7 Clinical Document Architecture, Release 2 (CDA) specification in accordance with requirements set forward in ASTM E2369-05 Standard Specification for Continuity of Care Record (CCR). The resulting specification, known as the Continuity of Care Document (CCD), is developed as a collaborative effort between ASTM and HL7. It is intended as an alternate implementation to the one specified in ASTM ADJE2369 for those institutions or organizations committed to implementation of the HL7 Clinical Document Architecture. For more information visit <a href="http://www.hl7.org">www.hl7.org</a></td>
</tr>
</tbody>
</table>
### Integrating the Healthcare Enterprise (IHE) Exchange of Personal Health Record Content (XPHR)

The Exchange of Personal Health Record Content (XPHR) integration profile describes the content and format of summary information extracted from a PHR system used by a patient for import into healthcare provider information systems, and visa versa. The purpose of this profile is to support interoperability between PHR systems used by patients and the information systems used by healthcare providers. This profile does not address all the data exchange requirements of PHR systems. For more information visit [www.ihe.org](http://www.ihe.org).

### 2.3.3 INFORMATIVE REFERENCE STANDARDS

<table>
<thead>
<tr>
<th>Table 2-6 Informative Reference Standards Standard</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health Level Seven (HL7) HL7 Version 3 Standard: Clinical Document Architecture (CDA), Release 2</td>
<td>The HL7 Clinical Document Architecture is an XML-based document markup standard that specifies the structure and semantics of clinical documents for the purpose of exchange. CDA is one instantiation of HL7’s Version 3.0 Reference Information Model (RIM) into a specific message format. Of particular focus for HITSP Interoperability Specifications are message formats for Laboratory Results and Continuity of Care (CCD) documents. Release 2 of the HL7 Clinical Document Architecture (CDA) is an extension to the original CDA document markup standard that specifies the structure and semantics of clinical documents for the purpose of exchange. CDA R2 includes a prose document in HTML, XML schemas, data dictionary, and sample CDA documents. CDA R2 further builds upon other HL7 standards beyond just the Version 3.0 Reference Information Model (RIM) and incorporates Version 3.0 Data Structures, Vocabulary, and the XML Implementation Technology Specifications for Data Types and Structures. For more information visit <a href="http://www.hl7.org">www.hl7.org</a>.</td>
</tr>
</tbody>
</table>

---

**Session 6: HHIITSSPP CCDDAA CCoonntteenntt MMoodduulleess CCoommmmpoonneenntt**

Optionality/Repeatability
This column identifies the conditions under which the data element is sent, and whether it may be repeated in the exchange. The column contains two fields separate by a slash (/). The first field indicates when the data element is to be sent and the list of values used in that column is described below in Table 2-3.

<table>
<thead>
<tr>
<th>Optionality</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>REQUIRED - Required data elements must always be sent. Data elements that are required may under exceptional circumstances have an unknown value (e.g., the name of an unconscious patient). In these cases the sending application is required to indicate the reason that the data are not available where the standard permits. Some standards may not permit an unknown value at all</td>
</tr>
<tr>
<td>R2</td>
<td>Required if known - If the sending application has data for the data element, it is REQUIRED to populate the data element. If the value is not known, the data element need not be sent</td>
</tr>
<tr>
<td>O</td>
<td>OPTIONAL - Data elements that are marked optional may be sent at the choice of the sending application. An optional element need not be sent, but when it is sent, the meaning of that data element and a receiver can always be assured of what represents when it is present. Senders should not send an optional data value. If the value is not known, simply do not send the data element</td>
</tr>
<tr>
<td>C</td>
<td>Conditional - Data elements that are marked conditional (C) are REQUIRED to be sent when the conditions specified in the HITSP additional specifications column are true. The conditions under which the data element is to be exchanged will be specified as a constraint on the data element in the last column</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Repeatability</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>No. The data element SHALL NOT be repeated</td>
</tr>
<tr>
<td>Y</td>
<td>Yes. The data element MAY be repeated</td>
</tr>
</tbody>
</table>

2.2.1 CDA SECTIONS
Two types of content components are specified in this section, they are:
• CDA Entries – a collection of Data Elements pertaining to a single instance of the specified concept. For example, the Allergy/Drug Sensitivity Entry Module describes all the Data Elements for one allergy
• CDA Sections – a collection of Entries pertaining to a single specified concept. For example, the Allergies and Other Adverse Reactions Section can contain a list of allergies (multiple Entry Content Modules)
The Health Information Technology for Economic and Clinical Health Act (HITECH) authorized incentive payments through Medicare and Medicaid to clinicians and hospitals when they use EHRs privately and securely to achieve specified improvements in care delivery.

It will make available incentive payments totaling up to $27 billion over 10 years, or as much as $44,000 (through Medicare) and $63,750 (through Medicaid) per clinician.

Equally important, HITECH’s goal is not adoption alone but “meaningful use” of EHRs — that is, their use by providers to achieve significant improvements in care.

The department published proposed meaningful use requirements on January 16, 2010.

The most important part of this regulation is what it says hospitals and clinicians must do with eHRs to be considered meaningful users in 2011 and 2012. In the original proposal, we identified a broad set of objectives, all of which would need to be met. This included 23 objectives for hospitals and 25 for clinicians. The DHHS received many comments that this approach was too demanding and inflexible, an all-or-nothing test that too few providers would be likely to pass.

In the final regulation, we have divided these elements into two groups: a set of core objectives that constitute an essential starting point for meaningful use of EHRs and a separate menu of additional important activities from which providers will choose several to implement in the first 2 years (see table

Core objectives comprise basic functions that enable EHRs to support improved health care. As a start, these include the tasks essential to creating any medical record, including the entry of basic data: patients’ vital signs and demographics, active medications and allergies, up-to-date problem lists of current and active diagnoses, and smoking status.

Other core objectives include using several software applications that begin to realize the true potential of EHRs to improve the safety, quality, and efficiency of care. These features help clinicians to make better clinical decisions — and avoid preventable errors. To qualify for incentive payments, clinicians must start employing such clinical decision support tools. They must also start using the capability that undergirds much of the value of EHRs: using records to enter clinical orders and, in particular, medication prescriptions. Only when providers enter orders electronically can the computer help improve decisions by applying clinical logic to those choices in light of all the recorded patient data. And to begin extending the benefits of EHRs to patients themselves, the meaningful use requirements will include providing patients with electronic versions of their health information.

In addition to the core elements, the rule creates a second group: a menu of 10 additional tasks, from which providers can choose any 5 to implement in 2011–2012. This gives providers latitude to pick their own path toward full EHR implementation and meaningful use.

For most of the core and menu items, the regulation also specifies the rates at which providers will have to use particular functions to be considered meaningful users.

The HITECH legislation further requires that meaningful use include electronic reporting of data on the quality of care. In the final regulation, we have simplified the January proposals for quality reporting, while
still building toward a robust reporting capability that will inform providers about their own performance and will eventually inform the public as well. Clinicians will have to report data on three core quality measures in 2011 and 2012: blood-pressure level, tobacco status, and adult weight screening and follow-up (or alternates if these do not apply). Clinicians must also choose three other measures from lists of metrics that are ready for incorporation into electronic records.

On June 18, 2010, the DHHS issued a rule that laid out a process for the certification of electronic health records, so that providers can be assured they are capable of meaningful use. The department has also issued still another regulation that lays out the standards and certification criteria that EHRs must meet in order to be certified. Finally, realizing that the privacy and security of EHRs are vital, the DHHS has been working hard to safeguard privacy and security by implementing new protections contained in the HITECH legislation.
<table>
<thead>
<tr>
<th>Objective</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core set of objectives to be achieved by all eligible professionals,</td>
<td>Over 50% of patients’ demographic data recorded as structured data</td>
</tr>
<tr>
<td>hospitals, and critical access hospitals to qualify for incentive</td>
<td></td>
</tr>
<tr>
<td>payments</td>
<td></td>
</tr>
<tr>
<td>Record patient demographics (sex, race, ethnicity, date of birth,</td>
<td></td>
</tr>
<tr>
<td>preferred language, and in the case of hospitals, date and</td>
<td></td>
</tr>
<tr>
<td>preliminary cause in the event of death)</td>
<td></td>
</tr>
<tr>
<td>Record vital signs and chart changes (height, weight, blood pressure,</td>
<td>Over 50% of patients 2 years of age or older have height, weight, and</td>
</tr>
<tr>
<td>body-mass index, growth charts for children)</td>
<td>blood pressure recorded as structured data</td>
</tr>
<tr>
<td>Maintain up-to-date problem list of current and active diagnoses</td>
<td>Over 80% of patients have at least one entry recorded as structured</td>
</tr>
<tr>
<td>Maintain active medication list</td>
<td>data</td>
</tr>
<tr>
<td>Maintain active medication allergy list</td>
<td>Over 80% of patients have at least one entry recorded as structured</td>
</tr>
<tr>
<td>Record smoking status for patients 13 years of age or older</td>
<td>Over 50% of patients 13 years of age or older have smoking status</td>
</tr>
<tr>
<td>For individual professionals, provide patients with clinical summaries</td>
<td>Clinical summaries provided to patients for over 50% of all office</td>
</tr>
<tr>
<td>for each office visit; for hospitals, provide an electronic copy of</td>
<td>visits within 3 business days; over 50% of all patients who are</td>
</tr>
<tr>
<td>hospital discharge instructions on request</td>
<td>discharged from the inpatient department or emergency department of</td>
</tr>
<tr>
<td>On request, provide patients with an electronic copy of their health</td>
<td>an eligible hospital or critical access hospital and who request an</td>
</tr>
<tr>
<td>information (including diagnostic-test results, problem list, medication</td>
<td>electronic copy of their discharge instructions are provided with it</td>
</tr>
<tr>
<td>lists, medication allergies, and for hospitals, discharge summary and</td>
<td></td>
</tr>
<tr>
<td>procedures)</td>
<td></td>
</tr>
<tr>
<td>Generate and transmit permissible prescriptions electronically (does not</td>
<td>Over 40% are transmitted electronically using certified EHR technology</td>
</tr>
<tr>
<td>apply to hospitals)</td>
<td></td>
</tr>
<tr>
<td>Computer provider order entry (CPOE) for medication orders</td>
<td>Over 30% of patients with at least one medication in their medication</td>
</tr>
<tr>
<td>Implement drug–drug and drug–allergy interaction checks</td>
<td>list have at least one medication ordered through CPOE</td>
</tr>
<tr>
<td>Implement capability to electronically exchange key clinical</td>
<td>Functionality is enabled for these checks for the entire reporting</td>
</tr>
<tr>
<td>information among providers and patient-authorized entities</td>
<td>period</td>
</tr>
<tr>
<td>Implement one clinical decision support rule and ability to track</td>
<td>Perform at least one test of EHR’s capacity to electronically exchange</td>
</tr>
<tr>
<td>compliance with the rule</td>
<td>information</td>
</tr>
<tr>
<td>Implement systems to protect privacy and security of patient data in</td>
<td>One clinical decision support rule implemented</td>
</tr>
<tr>
<td>the EHR</td>
<td></td>
</tr>
<tr>
<td>Report clinical quality measures to CMS or states</td>
<td>For 2011, provide aggregate numerator and denominator through</td>
</tr>
<tr>
<td></td>
<td>attestation; for 2012, electronically submit measures</td>
</tr>
</tbody>
</table>
The American Reinvestment & Recovery Act (ARRA) was enacted on February 17, 2009. ARRA includes many measures to modernize our nation's infrastructure, one of which is the "Health Information Technology for Economic and Clinical Health (HITECH) Act". The HITECH Act supports the concept of electronic health records - meaningful use [EHR-MU], an effort led by Centers for Medicare & Medicaid Services (CMS) and the Office of the National Coordinator for Health IT (ONC). HITECH proposes the meaningful use of interoperable electronic health records throughout the United States health care delivery system as a critical national goal.

Meaningful Use is defined by

- the use of certified EHR technology in a meaningful manner (for example electronic prescribing);
- ensuring that the certified EHR technology is connected in a manner that provides for the electronic exchange of health information to improve the quality of care; and
that in using certified EHR technology the provider must submit to the Secretary of Health & Human Services (HHS) information on quality of care and other measures.

The concept of meaningful use rested on the '5 pillars' of health outcomes policy priorities, namely:

1. Improving quality, safety, efficiency, and reducing health disparities
2. Engage patients and families in their health
3. Improve care coordination
4. Improve population and public health
5. Ensure adequate privacy and security protection for personal health information

CMS grants an incentive payment to Eligible Professionals (EPs) or Eligible Hospitals (EHs), who can demonstrate that they have engaged in efforts to adopt, implement or upgrade certified EHR technology. In order to encourage widespread EHR adoption, promote innovation and to avoid imposing excessive burden on healthcare providers, meaningful use was showcased as a phased approach, which is divided into three stages which span 2011 (data capture and sharing), 2013 (advanced clinical processes) and 2015 (improved outcomes). The incentive payments range from $44,000 over 5 years for the Medicare providers and $63,750 over 6 years for Medicaid providers (starting in 2011). Participation in the CMS EHR incentive program is totally voluntary, however if EPs or EHs fail to join in by 2015, there will be negative adjustments to their Medicare/Medicaid fees starting at 1% reduction and escalating to 3% reduction by 2017 and beyond.

**Meaningful Use Stage 1**

On July 13, 2010, CMS/ONC displayed the final rules as related to meaningful use in the context of objectives and measures and standards, implementation and vocabulary respectively. The final rules were published in the Federal Register on July 28, 2010, and became effective on September 26, 2010. The CMS final rule requirements have been divided into 15 core set objectives, and 10 menu set objectives (where there is an option to pick 5 of 10. However, it is mandatory to include at least one population/public health measure). In the future, ONC and CMS intend to propose expansion on the stage 1 criterion where it is likely that the currently proposed menu set of measures will be transitioned into the core set for stage 2.

For 2011, the public health community is working to assess and ensure readiness in Immunization Information Systems (IIS), Electronic Laboratory Reporting (ELR) and Syndromic Surveillance (SS). Public health will need to both test capability of systems to report and actually receive data where required and accepted. Public health will seek to expand current case reporting between hospitals/providers and public health and increase capacity for data management and analysis. There will be a need to coordinate across programs, state health information technology (HIT) coordinators, state Health Information Exchange (HIE) plans, and CMS.

**Meaningful Use Stage 2**

On August 23, 2012 the HHS’ Centers for Medicare & Medicaid Services and HHS’ Office of the National Coordinator for Health IT released 1) final requirements for Stage 2 Meaningful Use (MU) that hospitals and health care providers must meet in order to qualify for incentives during the second stage of the program, and 2) the criteria that electronic health records must meet to achieve certification. In Stage 2 MU, the Eligible Professionals (EPs) must meet or qualify for an exclusion to 17 core objectives and 3 of 6 menu set objectives. Similarly, the Eligible Hospitals (EHs) and Critical Access Hospitals (CAHs) must meet or qualify for an exclusion to 16 core objectives and 3 of 6 menu objectives. This final rule delays
the onset of Stage 2 MU criteria until 2014; the start date for EHs will be October 1st, 2013 and for EPs it will be January 1st, 2014. Specific to the Stage 2 MU Public Health objectives, the capability to submit electronic data for Immunizations is in the core set for EPs, and the capability to submit electronic data for Immunizations, Reportable Laboratory Results and Syndromic Surveillance are all in the core set for EHs. In addition, two new public health objectives for EPs have been added to the menu set, they include the capability to identify and report 1) cancer cases to a cancer registry and 2) specific cases to a specialized registry (other than a cancer registry).

Session 7: Semantic integration in healthcare networks

Richard Lenz, Mario Beyera, Klaus A Kuha

a Philipps-Universitaet Marburg, Institut fuer Medizinische Informatik, Bunsenstrasse 3, 35037 Marburg, Germany
b Technische Universitaet Muenchen, Lehrstuhl fuer Medizinische Informatik, Munich, Germany

International Journal of Medical Informatics 76 (2007) 201–207

IT applications should guide data acquisition in a way that data are placed in a meaningful context from the beginning, so that they are ready for reuse in different contexts without the need to manually index or transform the data.

Semantic heterogeneity remains as a major barrier to seamless integration of autonomously developed software components (cf. [3]). Semantic heterogeneity occurs when there is disagreement about the meaning, interpretation or intended use of the same or related data [4]. It occurs in different contexts, like database schema integration, ontology mapping, or integration of different terminologies.

Stonebraker characterizes disparate systems as “islands of information” and points out two major factors which aggravate systems integration [5]: 1. Each island (i.e. application) will have its own meaning of enterprise objects.

2. Each island will have data that overlaps data in other islands. This partial redundancy generates a serious data integrity problem.

Based on this statement, data integration can be led back to a mapping problem (how to map different conceptualizations in a semantically correct way) and a synchronization problem (how to ensure mutual consistency of redundant data which are stored in different databases under the control of autonomous applications).

In order to reduce the integration efforts caused by semantic heterogeneity standards for systems integration are needed.

At a conceptual level, information systems are designed around three layers: presentation, application logic, and resource management.

• Data integration The goal of data integration is to create a unique semantic reference for commonly used data and to ensure data consistency. To create such a semantic reference different facets of data semantics have to be considered. In this article three facets are roughly distinguished:
The *instance level*, referring to the semantics of individual data objects, which corresponds to the meaning of entries in a database.

The *type level*, designating the semantic classification of data objects, which roughly corresponds to the database schema.

The *context*, which refers to the semantic relationships that associate an object with other objects. To illustrate the difference of these aspects we may consider a concept “diagnosis” on the type level, and a particular instance, say “encephalitis”, and the context of this instance, which is determined by the patient, the physician who made the diagnosis, and other objects that contribute to a particular statement (information).

*Functional integration* refers to the meaningful cooperation of functions of different software components.

In our characterization of integration aspects data integration is concerned with the consolidation of declarative knowledge, while functional integration is concerned with the consolidation of procedural knowledge on which applications are based. Both aspects have to be considered for application integration.

*Desktop integration* or *presentation integration* refers to the user interface of a distributed system. Desktop integration is aimed at user transparency, meaning that the user would not know what application was being used or what database was being queried [14]. This requires more than a unified layout and uniform interaction mechanisms. Examples for functions needed to achieve desktop integration are “single sign-on” and “desktop synchronization”. Desktop synchronization is needed when a user has multiple windows to different applications on her desktop that share a common context. Synchronization is required when the context is changed in one of the interlinked applications.

By “technical integration” we refer to the technical infrastructure which supports application integration. “Semantic integration”, in contrast, refers to the meaning of data and functions.

**4. Results**

XML and RDF are examples for standard *syntactic frameworks* supporting data integration [15]. Standards for semantic integration in healthcare are increasingly based on XML in order to improve syntactical compatibility with commonly accepted data processing formats.

*Middleware* standards typically provide a common infrastructure for interconnecting distributed software components. Such standards are primarily intended to provide programming abstractions, which help a programmer to easily bridge different hardware, operating systems, and programming languages. Examples for standardization efforts in this area are CORBA, .net, EJB, or Web Services.

*Ontologies and vocabulary* standards support semantic data integration, as they serve as a semantic reference for system programmers and users.

Despite well accepted standards for data integration like HL7 V2 and DICOM, healthcare applications are still far from plug and play compatibility. One reason for this is that the existing standards do not address functional integration issues sufficiently.

In order to avoid these difficulties common *application frameworks* are required which serve as an additional semantic reference for programmers to create functionally compatible software components.

The best example for such a standard in the healthcare domain is the IHE initiative (“Integrating the Healthcare Enterprise”) [24]. IHE does not develop new standards for data interchange but specifies integration profiles on the basis of HL7 V2 and DICOM. Thereby “actors” and “transactions” are defined.
independently from any specific software product. An integration profile specifies how different actors interact via IHE transactions in order to perform a special task.

It turns out that there is a gap in the lower right corner where standardized medical processes could have been expected (such as IHE addresses organizational processes). Medical pathways and guidelines fall into this category. This is essentially medical knowledge which has to be consented by medical experts and which evolves over time.

Yet, despite of many attempts, a unique and comprehensive ontology of the medical domain is not within sight. A closer look at the given examples would reveal that medical terminologies continuously evolve over time (cf. [39]), and that there is no stable unique reference for system programmers. Thus, semantic integration of heterogeneous systems in healthcare will have to deal with volatile medical concepts.

5. Discussion and conclusions

Different kinds of standards are necessary to ease systems integration. In particular, both reference ontologies and application frameworks are needed to support semantic integration. Yet, standards should not try to comprehensively model an application domain, because systems must be capable to rapidly adapt to an evolving application domain.

Thus, the evolution of information systems should be a demand-driven process under the control of healthcare professionals. Process integration is concerned with the alignment of IT systems to actual business processes in a concrete setting. This is not addressed by standards, but by appropriate models for demand-driven software development (e.g. [41]).

Desiderata for such a demand-driven process are

• Rapid application development: In order to be able to flexibly react to newly arising demands, tools and techniques for rapid application development (RAD) are desirable.

• Robust and stable integrated domain-specific IT infrastructure: An IT infrastructure for a healthcare network should provide a robust and stable basis for application development. Thus, the framework should be based on generic but stable domain models instead of comprehensive but volatile domain models.

• Separation of domain concepts and system implementation: In order to cope with domain evolution, modeling of domain concepts should be separated from IT system implementation. IT systems should be implemented by IT experts and medical knowledge should be modeled and maintained by domain experts.

• Multi-level software engineering approach: To bring application development as close to the end user as possible, a multi-layered software engineering approach is proposed. An idealized abstract model for such a multi-level approach for software engineering is shown in Fig. 2. The basic idea is to distinguish different competencies and different responsibilities on the different layers of system design. The aim is to reduce complexity within each layer and to provide reusable services for higher layers.

• Layered ontologies: To support semantic integration within such a layered approach, layered ontologies are needed, which may serve as semantic references on different levels of software development. The layered approach of the clinical document architecture (CDA), and the generic HL7 V3 reference information model (RIM) are emerging standards which are already built on this fundamental principle.
Layered approaches have proven to be a successful technique for separating concerns and reducing system complexity (cf. [42,43]). Transferring this principle to the development of information systems in complex application domains is aimed at allowing application developers and end users to build well-integrated healthcare applications without the need to do low level coding and debugging.

A layered approach, as sketched above, fosters a system evolution process that follows the principle of “deferred systems design” [44], which is aimed at closing the gap between systems design and healthcare process reality [45].

Multi-layered service-oriented architectures are expected to provide a suitable technical platform for IT support in heterogeneous healthcare networks, as they provide the necessary technical infrastructure for loosely coupled interoperating components [1,46].

In addition to such a basic healthcare-specific service infrastructure there is a need for concepts that help to separate different layers of software engineering, as different responsibilities and competencies are to be addressed for technological evolution, domain evolution, and site-specific adaptation. A first step towards such a separation of concerns is provided by the “archetype” approach, which has been developed in the context of the GEHR project [52,53]. This concept is aimed at separating IT systems from domain knowledge, in order to enable medical knowledge to be modeled by domain experts rather than IT specialists. Archetypes are focused on the specification of declarative medical knowledge.
Fig. 1 – Contribution of different standards to application integration.

Fig. 2 – A layered approach for system evolution.
The Person Centered Coordination Plan (PCCP) framework addresses and solves many issues of health care reform. PCCP enables technology-assisted longitudinal health care collaboration that is personalized, flexible and focused on measurable outcomes. PCCP reorients the incentives, encouraging levels of care appropriate for the individual and providing the resources required. PCCP is designed to enable improved analytics and reliable automation of care designed and executed by experts working at the top of their licenses. PCCP enables new and improved reimbursement scenarios.

PERSON CHARACTERISTICS DOMAIN
Person characteristics include demographics, traditional health data, and an extended set of environmental and personal features. The extended set allows optimal configuration of plan designs and the setting of expectations that are realistic and achievable for each individual. Plans must be understandable by the person, requiring consideration of their education, language, culture, health and numeric literacy and learning styles. Beliefs and desires will shape the intentions, which in the PCCP are called tasks (next domain). The care tasks for someone who believes smoking is not harmful are different from tasks for a person who knows smoking is harmful and desires to quit. Change is generally a journey with incremental stops along the way to an ultimate destination. People and care collaborators should be rewarded for moving to measurable and meaningful states along the path to an ultimate outcome.

Personal characteristic are acquired with active participation of the individual using questionnaires, surveys, interviews, aggregating existing information, and observations. Many members of the health team and community can contribute insights to the person characteristics domain.

**TASKS DOMAIN**
Tasks are the primary “currency” of the PCCP. NQF defined key components of the task: an accountable entity, task content and the anticipated outcome1. The NQF Quality Data Model (QDM)2 allows us to represent tasks in a computable format. The accountable-entity can be the person themself, an electronic system, physician, nurse, pharmacist, home health worker, transportation service, or other collaborating care team member.

Tasks are often sequential or hierarchical reflecting the complex nature of most workflows in health care. Different steps in the sequence or hierarchy require different skills and, if using the top-of-license strategy, many different team members. Getting a medication into a person is an example of a sequential set of tasks. It involves formulary options, a prescribing event, medication dispensing, delivery to the person and their taking it.

**TASK MANAGER DOMAIN**
PCCP tasks are computable. The implementing systems should automate the management of tasks. Collaborating team members should have confidence that they can create and assign a task, knowing that it will stay on schedule and be satisfactorily completed or returned to them for adjustments. It is the task manager’s role to notify accountable-entities and provide sufficient information for them to decide whether they will accept the role. The task manager then tracks accepted task milestones and deliverables, reporting only when there are exceptions and then using the appropriate strategy. Exceptions might be managed by the accountable-entity through reminders or escalated to a “top of license” exception manager.

**INCENTIVE MANAGER DOMAIN**
The PCCP can fundamentally change the behavior of the health care enterprise by supporting new analytics that compute motivating incentives for the collaboration team. The incentives can be inserted throughout the care continuum to assure desirable entry paths, rewards for outcomes and data along the way to drive performance improvement.

The entry path or “front end” of PCCP creates a contemporary view of the person’s characteristics and tasks for the collaborative care team. This enables analytics that are dramatically different from traditional actuarial methods in predicting the costs of care. Specific tasks can be monetized based on the resources used (costs) and the benefits obtained by the anticipated outcome (savings). Part of the monetized value can be paid at the front end as a risk adjusted regular payment which supports the deployment of the required resources. Part of the monetized value can be paid when the specific anticipated outcomes are attained. The PCCP task manager acquires specific workflow data which could be utilized in at least two ways to provide additional value. First, time driven activity based costing.
(TDABC)6 provides granular data for practice improvement and enables retail reimbursement models which eliminate inefficient claims workflows. Second, PCCP workflows can be measured and coordination metrics rewarded with specific payments.

CONNECTING CAPABILITIES
The four domains of the PCCP are each first class citizens operated as independent services managed by experts in their implementation. But they must interact with each other in predictable and computable ways. This functionality is accomplished by a series of “connecting capabilities” that were defined in the NQF and IHE frameworks and can be supplemented with additional capabilities.

Beliefs, desires and intentions (BDI) are linked concepts. In the PCCP, beliefs and desires are part of the personal characteristics and these must be translated into intentions which are PCCP tasks. This is a complex process which has been successfully implemented in other industries using technology.

The service agreement is a key PCCP feature that does not have an identified industry solution yet. Its two main functions are to assure 1) access to care and 2) appropriate care.

To implement service agreements, NQF created the concept of a transparency agreement. A PCP has to understand the collaborating providers and a transparency agreement authorizes them to see the necessary information. Together, the transparency and service agreements allow any PCP or other entity to create a virtual network of accountable-entities who can implement a PCCP. This framework can support models such as accountable care organizations (ACO) or health care homes. But it has flexibility and agility to accommodate many other models and the entrepreneurial spirit of innovative health care stakeholders.

ECOSYSTEM REQUIREMENTS
The PCCP is a longitudinal and dynamic document which must be readily accessible by all collaboration team members. Implementing PCCP can begin by utilizing health information exchange (HIE) capabilities. The PCCP could reside on the HIE platform or be accessed on an HIE node through the exchange. The modular nature of PCCP lends itself to incremental implementation steps that may differ in various venues depending on their existing capabilities and priorities. The modular nature also encourages division of labor in developing components which can then be shared with others.

Many capabilities must come together for the health care system to function optimally. Interoperability is the ability of components to interact using computable methods. Maximizing the integration of these two concepts depends on these components and capabilities sharing a common ontology. Stated more technically, semantic interoperability of the components and capabilities would use a common foundational framework. This lingua franca or semantic role might be served by the Quality Data Model if it becomes a component of many different capabilities.

The PCCP is a good use case because it has capabilities, illustrated in Figure 1, which can be linked ontologically using QDM. A common ontology enables interoperability.

---

**Session 7: Health Information Technology Automation of Quality Measurement: Quality Data Set and Data Flow**
In this, its second report, *Health Information Technology Automation of Quality Measurement: Quality Data Set and Data Flow*, HITEP addresses the issue that quality measure specifications do not leverage EHR systems because the clinical information required for quality measurement is not adequately captured in EHRs. To resolve these gaps, HITEP drafted a quality data set (QDS) to empower automated, patient-centric, longitudinal quality measurement. This report describes the QDS framework and necessary “connectors” to electronic information called data flow attributes.

The American Recovery and Reinvestment Act of 2009 and the Health Information Technology for Economic and Clinical Health Act have significantly raised the bar and shortened timelines for implementation by providing funding to support the adoption of qualified EHRs and the alignment of related timelines. The acts specifically define the meaningful use of HIT systems as the use of electronic prescribing (e-prescribing), the electronic exchange of health information to improve the quality of healthcare, such as promoting care coordination, and the submission of information on clinical quality measures.

Efficient measurement must automatically gather reliable, high-quality, clinical information from numerous electronic sources. Based on the results of an environmental scan, HITEP had a clear mandate that measures must be more clearly and consistently defined, that there needs to be recognition that structured data and the reuse of data elements that exist in EHRs or other electronic formats are essential, and that workflows are complex.

HITEP therefore created two workgroups—the QDS Workgroup and the Data Flow Workgroup. The QDS Workgroup standardized data elements and developed a framework to consistently use standard code sets and code lists. The Data Flow Workgroup addressed how to determine from its use within the clinical workflow that any given data element is the authoritative source for the information required.

**The QDS framework contains three levels of information: standard elements, quality data elements, and data flow attributes.**

Standard elements represent the atomic unit of data identified by a data element name, a code set, and a code list composed of one or more enumerated values. Examples include diabetes and all pertinent ICD-9-CM codes or diabetes medications and all representative medications coded in the code set RxNorm. Standard data elements can be reused within other quality data elements.

Quality data elements are pieces of information that are used in quality measures to describe part of the clinical care process. Examples include active diabetes diagnosis, diabetes family history, diabetes medication dispensed, and diabetes medication administered. Quality data elements can be reused by other measures, clinical guidelines, and clinical decision support (CDS) developers. The quality data type is a grouping of information that indicates the circumstance of use for any individual standard data type. Examples include active diagnosis, family history of diagnosis, and medication prescribed.

Data flow attributes describe the authoritative source for the information that is required to represent any given quality data element. Data flow attributes include the data source, recorder, setting, and health record field.

The *source* is the originator of the quality data element and may be an individual or a device.
The **recorder** is the individual or device that enters the data element into a health record field and also may be the source of the data, but that is not necessarily true.

The **setting** is the physical location where the data element is captured, defining the encounter location where the data are expected to originate.

The **health record field** is the location within an electronic record where the data should be found.

HITEP offers six recommendations for further work to enhance the development and use of the QDS and electronic data sources.

**RECOMMENDATION 1:** NQF should develop and maintain the QDS with the involvement of all stakeholders. Specific recognized standards and taxonomies should be used. The QDS should be hosted in a publicly available, centrally located, web-based repository.

**RECOMMENDATION 2:** Develop measures that use the richness of all available electronic data, focusing on clinical, patient-centered outcomes. Quality measures should leverage clinical data captured in the EHR as a byproduct of routine clinical care.

**RECOMMENDATION 3:** Communicate with all stakeholders and seek their buy-in, and educate and train the quality measure supply chain (e.g., study designers, guideline developers, quality measure developers, performance reporting consumers, EHR vendors, and CDS developers) regarding the QDS and its associated authoring tool.

**RECOMMENDATION 4:** Set a timeline for QDS implementation, including demonstrated functionality and workflow assessment, and enumerate the essential activities and stakeholders. Perform comparative testing to assess the validity and reliability of performance measures derived from EHR clinical data.

**RECOMMENDATION 5:** NQF should move swiftly to incorporate the QDS into the Consensus Development Process. Requesting that measure developers incorporate the QDS model into their measure submissions will ease the process of incorporating endorsed measures into EHR systems.

**RECOMMENDATION 6:** Future quality measure development should use the National Priorities and Goals as a guide. The QDS maintenance activity should track and assign data quality scores for the data requirements for emerging measures using the QDS.

It would be timely and appropriate for NQF to offer an approach to the measurement of meaningful use. Next steps included the development and approval of a set of HIT-sensitive criteria that can be used to identify clinical performance measures that highlight the effect of meaningful use of HIT.

The HIT-sensitive criteria can be used to emphasize measures that demonstrate the effect of the use of core HIT functions on clinical quality:
- e-prescribing;
- preventive services reminders;
- health information exchange; and
- CDS.

The HIT-sensitive criteria can be used to systematically review the NQF portfolio of endorsed and pipeline measures to identify a starter set of HIT-sensitive measures that focus on meaningful HIT use in topical areas related to the National Priorities and high-impact conditions. Measure developers can work with NQF to further retool HIT-sensitive measures to conform to EHR-based specifications.
Future work includes the ongoing maintenance of the QDS, the maintenance of reusable code lists, and the development of a measure authoring tool to enable more facile incorporation of the QDS into the quality measurement development process. Additionally, further coordination with standard-development organizations and EHR certification bodies is required to encourage increased quality data type migration into EHRs.

Introduction

The Health Information Technology for Economic and Clinical Health Act (HITECH) (part of the American Recovery and Reinvestment Act of 2009 [ARRA])5 has significantly raised the bar and shortened timelines for implementation by providing funding to support the adoption of qualified EHRs and alignment of related timelines. The act specifically addresses three areas of critical importance for quality measurement: a. using EHR technology in a meaningful manner including electronic prescribing; b. exchanging health information electronically to improve healthcare quality, with the promotion of care coordination cited as an example; and c. submitting clinical quality measure-related information in a form and manner specified by the Secretary of the Department of Health and Human Services (DHHS).

The goal of this effort has been to represent quality data requirements (concepts, data types, data elements, and sets of values or codes) unambiguously and specifically.

The structure of the QDS must be simple to understand and sufficiently robust to incorporate information about each element such that it can be reused without ambiguity with respect to meaning.

Standardizing quality measures will help to automate successful measurement. Yet the volume and variety of measures that exist in paper format slow standardization.

Standardization will help us to speak and understand the same quality language. Currently, those who use quality measurement in their clinical practices are burdened by new or updated measures.

Although standards enable information sharing, the number of participants and the complexity of information in the quality conversation limit the feasibility of measurement.

Environmental Scan

In summary, the environmental scan provided a clear message that measures must be more clearly and consistently defined, that having structured data and reusing data elements that exist in EHRs or other electronic formats are essential, and that workflows are complex. It also is clear from the scan that the authoritative data required to capture the meaning of elements within a measure can be found in specific medical record locations. In conclusion, it is important that HITEP standardize data elements, consistently use standard code types and common code sets, and enable the specification of measures with respect to an EHR-specific authoritative source for the information desired.

Rationale

The QDS was designed in part to address each of the HITEP-I recommendations.

RECOMMENDATION: NQF should evaluate the quality of data types used in measure specifications as a criterion in the endorsement of new measures, as well as in the reassessment of measures for continued endorsement.

RECOMMENDATION: A coded, interdisciplinary clinical problem list in the EHR should be used in place of billing codes to identify patient conditions, inclusion diagnoses, and exclusion diagnoses for quality.
measurement. It is further recommended that this problem list be accessible and utilized across care settings (e.g., inpatient, outpatient, long-term care facilities).

RECOMMENDATION: NQF should work with HITSP to develop a “reader’s digest” version of a data dictionary for use by measure developers that would contain the HITEP data types and their corresponding HITSP-recommended code sets.

RECOMMENDATION: Medication allergies and side effects should be distinguished from each other and entered using standardized codes

RECOMMENDATION: Standardized codes for summary impressions of diagnostic test results should be developed, where feasible. Quantitative results, when available, should accompany qualitative results of diagnostic studies.

RECOMMENDATION: EHR vendors should develop methods of presenting EHR medication data with external medication data from pharmacies and pharmacy networks to help providers assess patients’ adherence to medication treatment plans.

RECOMMENDATION: Quality and information technology stakeholders should work together to define additional EHR functional requirements that support quality measurement.

Goals

GOAL: Describe, unambiguously, the clinical information that is needed for all quality measures.

GOAL: Reuse quality information definitions.

GOAL: Accommodate current and future measure needs

GOAL: Bridge the translation gap between quality measure content experts and EHR vendors and implementers. A common language that both quality measure developers and HIT programmers can speak will facilitate automated measurement using electronic information.

Framework Process

Quality Data Element

A quality data element is a single piece of information that is used in quality measures to describe part of the clinical care process, including both a clinical entity and its context of use.

Quality Data Type

This data type is the context of use for the information required.

Standard Element

To automatically locate quality data elements within an electronic medical record, information must be described explicitly

Standard elements may contain a list of codes (e.g., ICD-9-CM codes), words (e.g., drug names), or concepts described at length in measure specifications (e.g., definition of smoking cessation counseling components) to allow an abstracter to determine whether the concept is in the medical record.
The advantage of using standard elements is that a single code list can be reused in many different quality data elements.

Quality Data Flow

Quality data flow allows a measure developer to clearly define in the specifications where the quality data should be found to achieve the intended meaning of the measure.

The data flow contains four attributes:

1. **Source** - The source is the originator of the quality data element. The source may be an individual or a device.

2. **Recorder** - The recorder is the individual or device that enters the data element into a health record field. The desired recorder also may be, but is not necessarily, the source of the data.

3. **Setting** - The setting is the physical location where the data element is captured. The setting defines the encounter location where the data are expected to originate.

4. **Health Record Field** - The health record field is the location within an electronic record where the data should be found.

Summary of QDS Framework

The QDS framework contains three levels of information: standard elements, quality data elements, and data flow attributes (Figure 7). A standard element is an atomic unit of data that is identified by a data element name, a code type, and a code set composed of one or more enumerated values. Examples include diabetes and all pertinent ICD-9-CM codes, or diabetes medications and all representative medications coded in the code type RxNorm. Standard data elements can be reused within other quality data elements.

A **standard category** is a class or category of information. Examples include medication, problem, laboratory test, and diagnostic test.

A **quality data element** is a single piece of information used in quality measures to describe part of the clinical care process.

A **quality data type** is a grouping of information that indicates the circumstance of use for any individual standard data type.

**Data flow attributes** describe the authoritative source for the information that is required to represent the quality data element.

Preferred Future State

The QDS framework provides a structure within which to apply quality measure policy recommendations. A few HITEP-I summary recommendations can be achieved as a direct effect of the QDS (Figure 8).
1. **Evaluate the quality of data types as the criterion for measure endorsement.** If quality measures utilize the QDS, the data types are explicitly defined by the measure developer.

2. **Use coded, interdisciplinary clinical problem lists in place of billing codes.** To achieve this, two changes must occur. First, the data flow must specify the field as “problem list” rather than as “billing.” However, this would not solve the problem entirely, because some EHRs use billing codes in the problem list as well. Therefore, the code set that is used to describe diagnoses must change from ICD-9-CM to SNOMED-CT, as recommended by HITSP.

3. **Utilize HITEP data types and HITSP recommended code sets.** Figure 8 also demonstrates transitioning code sets from NDC to RxNorm (for medications) and from CPT to LOINC (for laboratory tests).

4. **Distinguish allergies and side effects from each other.** The QDS contains a data type for “medication allergy” but does not have a data type for “medication side effect.”

**Considerations for Meaningful Use**

The effort requires the following steps:
- Develop and approve a set of HIT-sensitive criteria that can be used to identify clinical performance measures that highlight the effect of meaningful use of HIT.
- Use the HIT-sensitive criteria to systematically review the NQF portfolio of endorsed/pipeline measures to identify a starter set of HIT-sensitive measures that highlight meaningful HIT use in topical areas related to the National Priorities and high-impact conditions.
- Encourage measure developers to work with NQF to further retool HIT-sensitive measures to conform to EHR-based specifications.

**Future Work**

**Maintenance of the QDS**

The QDS contains quality data elements for measurement use. As measures are created and continually updated, the QDS will need to reflect these changes. HITEP recommended that maintenance of the QDS content should occur every six months.

**Maintenance of Standard Code Sets**

Standard elements are best housed in a code set repository so the elements also may be reused for routine clinical information system implementations, for guideline compliance, and for CDS functions.

A code set repository, or registry, is required that contains the standard elements from which new quality data elements can be selected. Some of these elements will be represented in more than one code type because, in the near term, certain code systems will be commonly used in electronic transactions. Alternate code system sets are required to enable transition to future use (e.g., ICD-9-CM to ICD-10) and also to enable transition to preferred code systems to provide more clinical context.

**Measure Authoring Tool**

A near-term future step in this process is to develop a measure authoring tool to allow measure developers to select quality data elements, apply mathematical operators and logic, and create a computer-readable measure specification.

**Quality Data Type Migration to EHRs**
A future action also will require enhancing some of the concepts in the EHR Functional Model and creating new concepts and/or a Quality Profile in HL7 that represents the appropriate concepts as requirements in EHRs. It is expected that HITSP and the Health IT Standards Committee will identify standards for interoperability to allow data entered once, in EHRs, to flow, as appropriate, to registries and other receivers of data.

Box 1: Example Measure Description*

Percentage of patients with diabetes (either by administrative claim of diabetes or if a diabetic medication has been dispensed) who have had a Hemoglobin A1c laboratory test result < 8 percent within the past year.

(National Committee for Quality Assurance [NCQA], currently under review for endorsement)

*This measure is intended to show the effectiveness of care for patients with diabetes.

Figure 1: Quality Data Element for Active Diagnosis of Diabetes
Figure 2: Quality Data Elements and Their Associated Quality Data Types

Diabetes Diagnosis
Diabetes Medication
HbA1c Laboratory

Quality data element
Quality data type

To illustrate different standard categories, the figures use color coding: blue shapes are diagnoses, green shapes are medications, and red shapes are laboratory tests.

Figure 3: Standard Element (including code set and code list) as Part of the Quality Data Element

Code set
Code list
Standard element

ICD-9 CM
250.0 ...
(Diabetes)

Quality data element
Quality data type

The standard element (light blue circle) has a code set and specific code list and is part of the quality data element. The color of the circle indicates the standard category—in this example, diagnosis.

Figure 4: Quality Data Elements, Quality Data Types, and Standard Elements

ICD-9 CM
250.0 ...
(Diabetes)

Diabetes Active Diagnosis

Diabetes Medication Dispensed

Laboratory Test Result

Each quality data element (rounded rectangle) contains a standard element (circle) with a code set (top of circle) and code list (middle of circle) to define the clinical information in a computer-readable format. The color of the circles define the standard categories diagnosis (blue), medication (green), and laboratory test (red).
Session 7: Semantic Interoperability for Better Health and Safer Healthcare

Research and Deployment Roadmap for Europe
Veli N. Stroetmann (Ed.), Dipak Kalra, Pierre Lewalle, Alan Rector, Jean M. Rodrigues, Karl A. Stroetmann, Gyorgy Surjan, Bedirhan Ustun, Martti Virtanen, Pieter E. Zanstra
SemanticHEALTH Report
January 2009

Introduction

The Commission identifies four major levels on which the Member States are encouraged to undertake action. These are the political, organisational, technical and semantic levels, with educational and awareness raising mechanisms to underpin initiatives in those main domains.
Semantic interoperability plays a prominent role – it is described as an essential factor in achieving the benefits from electronic health record systems to improve the quality and safety of patient care, public health, clinical research, and health service management.

Four priority areas and related challenges that would benefit most from these developments have been identified:

**Patient care:** patient safety; dissemination of good practice, integration of education and care; connecting multiple locations for collaborative care delivery (at local, regional, national and international levels); empowerment of citizens (patient centred healthcare)

**Public health:** international statistics; comparative outcome assessment; pharmacovigilance; coordination of risk assessment, management and surveillance of large-scale adverse health events, population health research

**Research and translational medicine:** multi-centre studies and trials, health data repositories, bio- and tissue-banks, development of personalised medicine based on genetic and genomic analyses

**Support for diverse markets:** identification of solutions with superior benefit/cost ratios; enabling plug-and-play best of breed, encouraging industry involvement, especially SMEs; stimulating innovations by health service providers and involving clinicians, harmonising legal and regulatory frameworks.

Defining semantic interoperability

The Semantic HEALTH study applies the following overall interoperability (IOp) definition: *Health system interoperability is the ability, facilitated by ICT applications and systems, to exchange, understand and act on citizens/patients and other health-related information and knowledge among linguistically and culturally disparate health professionals, patients and other actors and organisations within and across health system jurisdictions in a collaborative manner.*

In this context, semantic interoperability (SIOp) addresses issues of how to best facilitate the coding, transmission and use of meaning across seamless health services, between providers, patients, citizens and authorities, research and training.

Semantic interoperability (SIOp) has numerous facets:

*For individual patients* SIOp relevant tasks comprise assisted clinical data capture and quick access to the patient record as well as to pertinent background knowledge. It also includes quality assurance, clinical decision support, monitoring and alerts, as well as feedback regarding quality and costs.

*For aggregated population data* SIOp relevant tasks include reporting, health economics, surveillance, quality assurance, epidemiology (hypothesis formulation), bio- and tissue-banking.

SIOp enables the meaningful linkage of research findings and knowledge to patient information, and the discovery of new knowledge from semantically coherent EHR repositories.

In addition to precision of meaning, consistancy, understandability and reproducibility are three major desiderata for semantically interoperable systems:

- **Consistancy** means that the receiving system must be able to recognise what has been sent, so it is a prime requirement for machine-machine communications and dictates the need for unambiguous identifiers.
- **Understandability** is essential for human communication

**Reproducibility** addresses the question of inter-individual reliability when data are collected or encoded. This holds both for individual and aggregated data.
To further clarify these issues, the research distinguishes four levels of IOp, two of them relating to semantic interoperability:

Level 0: no interoperability at all  
Level 1: technical and syntactical interoperability (no semantic interoperability)  
Level 2: two orthogonal levels of partial semantic interoperability  
Level 2b: unidirectional semantic interoperability  
Level 2b: bidirectional semantic interoperability of meaningful fragments  
Level 3: full semantic interoperability, sharable context, seamless co-operability

Semantic interoperability, the goal, the vision and the challenges
In the specific case of EHR systems and their semantic interoperability requirements, we need:
- to enable the safe, meaningful sharing and combining of health record data between heterogeneous systems;
- to enable the consistent use of modern terminology systems and medical knowledge resources;
- to enable the integration and safe use of computerised protocols, alerts and care pathways by EHR systems;
- to link EHR data to explanatory and educational materials to support patient and family engagement and professional development
- to ensure the necessary data quality and consistency to enable rigorous secondary uses of longitudinal and heterogeneous data: public health, research, health service management

Interoperability requires agreement on meanings and labels for those meanings – on ontologies and lexicons, which together we label as terminologies. The primary goal of ontologies and terminologies for interoperability is to enable the faithful exchange of meaning between machines and between machines and people

3.2 Vision for the future

3.2.1 TECHNICAL EVOLUTION
Semantic interoperability will be achieved only gradually beginning with applications with high benefit and modest cost.

For terminologies, this will best be achieved by starting with areas where there is a high degree of consensus on both the content and the need. Key areas are likely to be sensitivities and adverse drug reactions, translational medicine, and large scale public health and population research initiatives such as “biobanking

3.3 Key trends

3.3.1 TECHNICAL TRENDS
Statistical text and Web mining technologies will advance rapidly, and Google-like technologies will take over much of the burden of coarse grained search for navigation information discovery

Direct encoding of free text into formal vocabularies and EHR structures will improve radically, partly driven by voice recognition

Personal medical systems will proliferate

Concerns about privacy and confidentiality will continue to be key limiting factors in interoperability, and may impede developments that would be technically feasible and beneficial.
3.4 Challenges

3.4.1 TECHNICAL CHALLENGES

Without greater economic or policy/regulatory incentives to interoperability little will be done and the status quo may be maintained.

Alternatively, enormous resources may be spent on overambitious plans for semantic interoperability which will inevitably fail. In either case communication will take place by going around rather than via the clinical information systems. In countries where it is mandated, large and unwieldy approaches such as SNOMED CT and HL7 V3 will become taxes on healthcare, absorbing significant resources while returning no, or in some cases even negative, benefits.

Terminologies will remain closed or partly closed.

3.4.2 CHALLENGES OF APPLICATION FIELDS

Without active policy interventions and further research, there will be tardy progress in patient safety improvement and reduction in clinical errors will (continue to) be slow and sporadic.

Public health will (continue to) depend on specialist encoders and be limited by the cost and accuracy of capturing information post-hoc. Bio-surveillance will remain a specialist activity divorced from mainstream clinical practice.

Clinical and translational research will continue to be conducted in silos. The cost of mounting multicentre trials will become the dominant barrier in the application of basic biological knowledge to medical care.

The market will (continue to) be dominated by a few large suppliers who supply ‘complete’, one-size-fits-all solutions to entire hospitals or even countries. Innovation will become more difficult.

SemanticHEALTH roadmap and recommendations

4.1.1 THE NEED FOR AND BENEFITS FROM SIOP

Semantic interoperability is most needed when electronic health record (EHR) data are to be shared and combined from different systems (or across diverse modules within a large system). Full semantic interoperability (Level 3) is required across heterogeneous EHR systems in order to gain the benefits of computerised support for reminders, alerts, decision support, workflow management and evidence based healthcare, i.e. to improve effectiveness and reduce clinical risks. The key semantic interoperability requirement identified in support of evidence based and safe clinical care is the ability to search for particular EHR data entries that are of relevance to such functionalities.

Current attempts to standardise the capture, representation and communication of clinical (EHR) data reply upon three layers of artefact to represent meaning:

1. Generic reference models for representing clinical (EHR) data, e.g. ISO/EN 13606 Part 18, HL7 CDA Release 29, the openEHR Reference Model
2. Agreed clinical data structure definitions, e.g. openEHR archetypes, ISO/EN 13606 Part 2, HL7 templates, generic templates and data sets
3. Clinical terminology systems, e.g., LOINC and SNOMED CT

On the one hand, full semantic interoperability cannot be reached without a clear sharing of roles between reference model, archetypes structure and terminology which are all necessary. On the other
hand, when one of the components is claiming its ability to produce full semantic interoperability alone or under the condition that the two other components conform to its needs, as it has been proposed very often in the past and still in the present, then the goal of full semantic interoperability cannot be reached.

**Therefore the goal of semantic interoperability is: to be able to recognise and process semantically equivalent information homogeneously, even if instances are heterogeneously represented,**

From the perspective of the EHR, achieving Level 1 (syntactic interoperability) enables the exchange of health record information to an extent that permits the mapping of corresponding parts of an information structure between systems, so that data for the relevant patient can be imported and can be selected and retrieved according to non-semantic properties such as the date of recording or the originating provider, and also searched by some coarse grained semantic categories such as a document type. This kind of interoperability is achieved, for example, by using a standard EHR reference model (without any semantic structures such as archetypes) or by using the IHE XDS (cross-document sharing) profile.

Level 2 (partial semantic interoperability) can be achieved in one of two ways. Level 2a (unidirectional semantic interoperability) is achieved by using a deeper level of data structure than simple documents and headings, i.e. finer grained entries are structured and labelled, but in ways determined by each system or vendor. A mapping is required in order for the receiving system to correctly match imported data items with the corresponding equivalents in the local repository.

Level 2b (semantic interoperability of meaningful fragments) is attained by agreeing and sharing fine grained data structures between sender and receiver, as has historically been carried out via predefined clinical messages (e.g. for screening and immunisation programmes, claims reimbursement) and is now being adopted for the transfer of prescription information within eHealth programmes.

In Level 3 (full semantic interoperability) the use of an EHR reference model, a rich library of clinical data structures, and the definitions of terminology bindings to value lists for each element of the data structures have all to be agreed within a record sharing community. This permits any arbitrary extracts from EHRs to be imported and combined with locally-held data seamlessly without the need for specific mappings.

**4.1.2 PRIORITY EHR APPLICATION FIELDS AND RECOMMENDATIONS**

It is instead recommended that Level 3 interoperability is sought in specific areas of clinical practice that are known to be of high patient safety risk, and in priority areas for which the evidence is strongest for a gap to be bridged between current and good practice. In effect, these are the cases for which computerised decision support and care pathway support are most needed. These priority areas are:

- New medication prescriptions requiring comprehensive information on concurrent medication and details of known allergies and conditions (not simple ETP – Electronic Transfer of Prescription)
- Reminders and prompts for overdue or overlooked health care actions and interventions
- Evidence-based care, the use of clinical guidelines and other forms of evidence to determine the optimal management strategy and care pathway for a given patient
- Care transfers, referrals and within-team workflow prompts such as the degree of urgency and the expectations of the referring clinician from another team member
- Care coordination ensuring that a high-level view can be taken of distributed (multi-team) care to protect against duplication, delay and incompatible interventions
4.2 Terminologies and ontologies

**Controlled Vocabulary** – a list of specified items to be used for some purpose, usually in an information system to reduce ambiguity, misspellings, etc.

**System of identifiers** ("codes") – Controlled vocabularies, and many lexicons, ontologies, and thesauri, are usually accompanied by systems of identifiers for their units, e.g., identifiers act as the primary unambiguous means of referring to the entities in the system.

**Lexicon** – A list of linguistic units that may be attached to a controlled vocabulary or ontology, in a specific language or sublanguage, often including linguistic information such as synonyms, preferred terms, parts of speech, inflections and other grammatical material.

**Ontology** (sensu information system) – a symbolic logical model of some part of the meanings of the notions used in a field, i.e. those things which are universally true or true by definition.

**Classification** – an organisation of entities into classes for a specific purpose such as international reporting or remuneration. Examples: ICD and Diagnosis Related Groups.

**Thesaurus** – a system of terms organised for navigation with the primary relationship being "broader than"/"narrower than".

**Background knowledge base** – or "Knowledge Representation System" – the common knowledge to be assumed by the system, including both the ontology – what is universally true – and generalisations about what is typically true.

**Terminology** – Any or all of the above in various combinations. Most health terminologies consist, at a minimum, of a controlled vocabulary and a system of identifiers. They may include extended lexicons, ontologies, thesauri or background knowledge base.

**Coding system** – A terminology with attached identifiers or "codes"

4.2.2 THE PROMISE AND PROBLEMS OF ONTOLOGIES

The primary goal of ontologies and terminologies is to enable the faithful exchange of meaning between machines and between machines and people and not to represent the state of the art of domain knowledge.

Here is the key distinction between ontologies and information models: whereas ontologies represent what is always true about the entities of a domain (whether or not it is known to the person that reports), information models (or data structures) represent the artefacts in which information is recorded.

4.2.5 RECOMMENDED ACTIONS

a) **Areas needing adoption or short term action**

The following areas require actions centred on the content of clinical terminologies:
A careful, methodologically sound, unbiased and public evaluation by independent evaluators of what SNOMED CT, in its current state, can and cannot be used for safely. Demonstration of a semantically sound and well quality assured reformulation of one or more suitable subsets of SNOMED CT35, in order to provide evidence for long term decisions on the role of SNOMED CT in Europe.

Harmonisation and cross mapping of major terminologies, including LOINC, DICOM, ICD 0-22; ICPC.
Adoption of a clear policy that endorses the mapping of all terminologies mapped to UMLS CUIs and LUIs, either by their originators or in collaboration with the US National Library of Medicine.

In summary, the key milestones on the road to SIOp include:

- A semantically sound SNOMED CT fragment supported by tools and organisation;
- A social/collaboratively built ICD-10 with widespread support;
- The mapping of ICD-10 to a semantically sound subset of SNOMED CT;
- The establishment of a set of widely used Web-based terminology services for access to, quality assurance, and feedback on clinical ontologies.

4.3 Public health

Recommended actions in the field of public health therefore include the development of common standards that will allow data-exchange of preidentified variables from an individual EHR, and compilation and comparison of that data across regions, time and populations.

Further action should aim at the establishment of:
National Centres for multilingual, multicultural adaptation of international classifications and terminologies, including SNOMED CT, linked in a well-managed European Network of Competence Centres, to be expanded globally. A European and global Network of Terminology Servers. Sustainability and scalability need to be assured and the terminology servers need to be maintained in order to be useful.

Finally, there is a need for action on the legal framework providing the enabling environment for public health action. There needs to be clarity on security and privacy regulations of public health data, clarity on mandate and responsibility of those actually carrying out the data aggregation, and finally rules on liability, should damage to individuals arise in the process.

4.4 Socio-economic issues

The following questions arise: (1) who is responsible for development and implementation, (2) who will pay, and (3) who will accrue the benefits. The costs of healthcare IT programmes now exceed ten billion Euros per annum across the Union. The health policy system cannot currently associate the benefits of interoperability with those who must pay for it. Therefore a convincing demonstration of the benefit of migration from legacy to interoperable systems is required in order to justify public intervention.

SIOp is not a binary variable, but rather a scale reaching from zero to full IOp. Various levels will imply different benefits and costs, and therefore it will be of critical importance to better understand and estimate these relationships to determine optimal levels of IOp.
Summary and outlook

A strong and coordinated effort is recommended to effectively engage with the relevant stakeholders policy makers.

In order to demonstrate the impact of semantic interoperability it is also recommended to create reference sites within hospitals across the EU.

Session 8: Health information exchange and patient safety
David C. Kaelber, David W. Bates

1. Introduction

Among the many potential advantages of health information exchange (HIE), patient safety stands out as one of the most promising.

As the Institute of Medicine’s 2000 To Err is Human report underscored, iatrogenic causes of injuries are frequent, and they represent an important cause of death among patients, with an estimated 44,000–98,000 deaths per year [1]. Patient safety can be eroded by both errors of commission and errors of omission if the right information is not available to the right person at the right time.

HIE can be thought of both in terms of who the exchange is between and what the type health information exchanged is (Fig. 1) [2]. Usually, the more people and more information involved in the HIE, the more valuable the exchange will be for patient safety. Up to 18% of patient safety errors have been estimated to have occurred because the appropriate information was not available at the time the medication decision was made.

We will divide HIE’s impact on medication information processing into five subsections below.

2.1.1. Drug-allergy information processing
One of the most obvious forms of medication information processing for patient safety is drug-allergy processing. This involves checking drugs against known patient-specific drug allergies before the drugs are given to the patient.

### 2.1.2. Drug–dose information processing

At the most basic level, drug–dose information processing involves being sure that the individual dose, dosing frequency, and total duration of medication fall within accepted general standards. At a more advanced level, drug–dose information processing can take into account patient-specific information such as patient age (geriatric dosing), weight (pediatric dosing), and creatinine clearance (renal dosing).

With respect to interoperability, renal dosing will probably be especially important, as the patient’s level of kidney function should be available for dosing decisions in all settings.

### 2.1.3. Drug–drug information processing

The most robust medication information processing currently occurs at the drug–drug level. This support can significantly improve patient safety, but it will be most effective only if the entire list of all of a patient’s medications, including over the counter medications, herbal medications, and supplements, are available at the time of medication prescription and administration.

Drug–drug information processing generally takes three forms. The most commonly thought of involves adding an additional medication(s) to a patient’s other medications.

A second type of drug–drug information processing involves duplicate pharmacological class checking.

A third type of drug–drug information processing for patient safety occurs when one medication is being added that could indicate the addition of another medication for improved patient safety.

### 2.1.4. Drug–diagnosis information processing

Drug–diagnosis (drug–disease) information processing is an expanding area in which enhanced HIE has the potential to improve patient safety.

Drug–diagnosis information processing will optimize patient safety only in an environment in which HIE allows all of a patients’ diagnoses to be available at the time of drug prescribing and administration.

### 2.1.5. Drug–gene information processing

Although not currently a reality, as gene analysis becomes more prolific and pharmacogenomics becomes more developed, the ability to interchange drug information and patient-specific genomic information will become increasing important for patient safety.

### 2.2. Improved laboratory information processing

Patient safety can also be improved by enhanced laboratory information processing enabled by HIE. The two primary areas for this include (1) helping to ensure that indicated lab testing is ordered and (2) helping to guarantee that lab test results (especially abnormal results) are appropriately followed up on.

### 2.3. Improved radiology information processing

### 2.4. Improved communication among providers
when different primary care providers and/or subspecialists are managing different medical issues, effective information sharing is critical, but does not always occur.

2.5. Improved communication between patients and providers

In general, no one should be more invested in their own healthcare safety then patients themselves. In our current healthcare paradigm, however, minimal HIE occurs with patients and healthcare organizations typically do not sufficiently recognize the key role that patients can play in ensuring their own healthcare safety. With the significant interest and impending growth of improved HIE with patients through personal health records (PHRs), many hope this paradigm will change.

2.6. Improved public health information processing

3. Health information exchange and decreased patient safety

Increasing use of HIT alone does not lead to superior patient safety. Rather, for improved patient safety, healthcare systems must be organized to facilitate effective use of their HIT.

Increasing the level of HIE could reduce patient safety in a variety of ways, for example,

- if incorrect patient-specific information were made available to providers,
- if one patient’s information was believed to be that of another,
- if there were errors in translating information between one system and others or
- if implementation of HIE slowed systems to a significant degree, since delays can affect safety.

Those who are evaluating HIE should be alert to these and other unintended consequences of implementation of HIE.

4. Standards for health information exchange and patient safety

For robust, efficient HIE, standards must be developed dictating the type and content of information to be exchanged.

5. Completeness of information for health information exchange and patient safety

Therefore, our ability to process health information in patients who have some health information that is not readily exchanged and/or to deal with patients who do not want their information exchanged may decrease and these patients could potentially see their patient safety eroded.

Session 8: Health information exchange policy and evaluation

Janet M. Marchibroda *


1. Drivers for policy change at the national levels

---

* Indicates authorship.
In a country where healthcare spending is 16% of the gross domestic product, and much higher than other industrialized countries, the United States—according to many leading employers—is losing its competitiveness and ability to compete globally

Concerns about cost and quality are driving policy makers at multiple levels of the system—federal, state and local—to take actions to improve health and healthcare for Americans. Most of these actions fall into three primary areas:

- driving transparency in quality and efficiency;
- aligning incentives with higher quality, more efficient healthcare; and
- using interoperable, standards-based health information technology (IT).

2. Federal policy action on health IT and health information exchange

As noted above, federal policy makers have been taking several actions designed to address healthcare challenges through the use of health IT and health information exchange

Since August 2006, DHHS Secretary Leavitt took several actions to implement the Executive Order, under a set of “Four Cornerstones” which are detailed in DHHS’ Prescription for a Value-Driven Health System [5]

- Connecting the system: Every medical provider has some system for health records. Increasingly, those systems are electronic. Standards need to be set so all health information systems can quickly and securely communicate and exchange data.
- Measure and publish quality: Every case, and every procedure has an outcome. Some are better than others. To measure quality, we must work with doctors and hospitals to define benchmarks for what constitutes quality care.
- Measure and publish price: Price information is useless unless cost is calculated for identical services. Agreement is needed on what procedures and services are covered in each “episode of care”.
- Create positive incentives: All parties—providers, patients, insurance plans, and payers—should participate in arrangements that reward both those who offer and those who purchase high-quality, competitively priced health care [5].

it became clear that in order to realize much of the value of health IT, it needed to be interoperable and data needed to flow across the institutions, organizations and practices that both deliver and pay for healthcare

As a result, policies began to shift slightly. There began to be more emphasis on standards for interoperability as opposed to incentives for health IT adoption, with policy makers recognizing that before large investments could be made, the federal government—and large private sector purchasers—needed to be assured that systems could “talk with one another”

it began to be clear that a large “national health information network” was likely not “the answer”.

Given differences in market-level characteristics, the need for “social capital” to enable the sharing of information among diverse, and in some cases—competing organizations, concerns about the privacy and security of a nationwide network, and sheer technical feasibility, policy makers at multiple levels of the system began to recognize that a set of state, regional and local health information exchanges—a “network of networks” was likely the best route forward for digitizing the U.S. healthcare system.
3. State-level policy action on health IT and health information exchange

State legislative mandates or executive orders are calling for state-level planning and coordination bodies to conduct activities such as:

- Gathering information about ongoing local, regional or statewide efforts.
- Determining the extent to which health IT is currently utilized within the state.
- Determining how health IT can be effectively deployed in the future within the state.
- Obtaining expert advice and information regarding the establishment of health information networks to facilitate the communication of clinical information.
- Assuring privacy and confidentiality of patient information through development and implementation of policies for information sharing.
- Investigating ways to coordinate health information exchange activities within the state. States’ increase in focus on and funding for health information exchange efforts at the state and local levels, furtheremphasizes the need for evaluation. As more and more health information exchange networks develop, it is critical that they learn from the experiences of others.

4. New York State as an example

5. Taking a look at health information exchange efforts across the US

Survey results also indicate that the most common functionalities of such efforts are those related to care delivery.

The most significant challenges for health information exchange initiatives, based on the survey results, are those related to securing initial capital and sustaining their operations.

The most difficult challenge facing these initiatives and organizations today, is that related to assessing the value of services that emerge from the health information exchange to various stakeholders groups (or customers) such as providers, payers, and employers, and converting those value assessments to business plans that promote and assure sustainability for these initiatives.

6. The need for evaluation

The most critical evaluation questions that need answers include the following:

1. Does effective implementation of health IT and/or health information exchange improve the quality of healthcare? Under what circumstances?

2. Does effective implementation of health IT and/or health information exchange improve the safety of healthcare? Under what circumstances?

3. Does effective implementation of health IT and/or health information exchange improve the efficiency of healthcare? Under what circumstances?

4. What value does health IT and/or health information exchange provide for various stakeholders in the system, such as practicing clinicians, employers and other healthcare purchasers, health plans, hospitals, and patients?

5. How can such value be converted into a sustainable business model for health information exchange at the state, regional and local levels?

6. If the goal of seed funding through grants and contracts is to create sustainable health information exchange efforts, how can funding and selection criteria be developed to assure such sustainability is achieved?
7. What key barriers (in addition to value assessment and the development of sustainable models) exist for health information exchange initiatives? Can any of these barriers be addressed through policy change?

Session 8: Health-information exchange: why are we doing it, and what are we doing?
Gilad J Kuperman

J Am Med Inform Assoc 2011;18:678e682. doi:10.1136/amiajn-2010-000021

NEED FOR INTEROPERABILITY AND EARLY WORK BY THE OFFICE OF THE NATIONAL COORDINATOR FOR HEALTH INFORMATION TECHNOLOGY

Key hurdles to solving the interoperability problem recognized at that time included

- the need for standards to represent clinical data,
- the need to identify a patient consistently as they moved among different providers and
- a framework to assure the patient’s privacy.

There were also questions of who should play a leadership role to address these issues and the kinds of organizational models that could best support interoperability.

A key component of many of the initiatives is a ‘record locator service,’ which serves to identify all the locations where a patient has received care. Increasingly, the term ‘health-information exchange’ (as a noun) is being used to represent an organization that addresses the business issues of interoperability, though the term RHIO also continues to be used

EARLY NATIONWIDE HEALTH INFORMATION NETWORK PROJECTS

To demonstrate that creating regional exchanges would not result in simply larger silos, ONC funded the $18.6 million Nationwide Health Information Network (NHIN) Prototype Architecture initiative, which ran from late 2005 to early 2007. The four projects in the program demonstrated that health information-exchange initiatives could be successfully connected to one another using a ‘network of networks’ approach that did not require a national patient identifier or a large-scale centralized operation. These lessons affirmed that regional health-information exchanges that ‘played by the rules’ would be able to participate in a nationwide network. ‘Playing by the rules’ meant using standard data services, consent services, security services (such as user authorization and auditing documentation), and other network management services.

Following the completion of the NHIN prototype projects, in 2007 ONC chartered the NHIN Trial Implementations project. The purpose of this project was to demonstrate data exchange among
operational health-information exchanges. Nine communities were the initial participants; eventually a total of 20 organizations participated. The specific goals included demonstrating the ability to

1. Identify a patient across disparate health-information exchanges,
2. Retrieve the patient’s clinical data from the other exchanges and display the aggregated data, and
3. Incorporate patient permissions

Each interoperability specification described in detail the software services and data structures that the participants in the Trial Implementations project needed to adhere to.

ADVENT OF HITECH

The final rules regarding meaningful use and EHR certification allow a fair amount of flexibility about how providers and hospitals can meet the interoperability-related meaningful use objectives as long as criteria related to vocabularies and data structures are met.

In response to the concerns, based on guidance from the NHIN Workgroup of the ONC Health IT Policy Committee, the NHIN Direct project was initiated in the beginning of 2010. In October 2010, the project was renamed the ‘Direct’ project. The goal of Direct is to create specifications to enable the secure exchange of health information between authorized healthcare providers to support stage 1 meaningful use.

Direct is an example of a ‘push’ model of health-information exchange. The Connect protocols support push-based transactions but also support a ‘pull’ of data, that is, the retrieval of data from multiple sources. The Connect protocols enable aggregation of a patient’s data across a community, whereas Direct services would not.

A push model, such as Direct, avoids problems that arise when trying to integrate a patient’s data across a community. Most notably, it is not necessary to link a patient’s identifiers across systems before data can be transferred. The cost and complexity of developing a record locator service, as well as developing privacy policies to support the retrieval of data from multiple sources, can be avoided. An inbound message is linked to a particular patient file by the message recipient, and the linking may be done manually.

If Direct services are to provide clinically useful health-information exchange capabilities, there are important related challenges that will need to be addressed. Directories will be needed to represent healthcare organizations as well as individual providers. Methods will be needed to allow an authorized sender to look up an authorized recipient.

The Information Exchange Work Group of the ONC Health IT Policy Committee is developing recommendations to promote the creation of such provider directories. An overall trust fabric that includes business, policy, and legal requirements; transparent oversight; enforcement and accountability; identity assurance; and minimum technical requirements will need to be established.

A governance model will be needed to develop the relevant policies and a process to assure compliance. At a technical level, the Direct project has grappled with coordinating the addressing and transport of messages between healthcare organizations. These issues need to be handled in a way that supports security and does not disrupt the way organizations manage messages internally.

Implementation of Direct services may present some pragmatic challenges as well.
An undesirable outcome would be to have provider offices overwhelmed by messages appearing in the EHR's inbox, akin to email overload.

From an informatics perspective, the Direct project offers a model of health-information exchange that is more constrained than models that involve record locator services. Whereas the Direct project supports the transmission of messages between providers, more complex protocols are needed to support the retrieval of data across an entire community.

ONC has noted that the Direct project is intended to be a first, perhaps easier, step that allows eligible providers to meet the objectives of stage 1 meaningful use and allows both providers and EHR vendors to participate in interoperability activities. ONC notes that the Direct project and the retrieval capabilities of Connect are complementary components of a complete ultimate interoperability vision.

As RHIOs grapple to support interoperability-based services that improve the quality and efficiency of care, they will have the opportunity to understand how best to combine pull- and push-oriented capabilities. A state that is developing a health-information exchange strategy as part of its response to the ONC’s State HIE Cooperative Agreement Program will have to determine how Direct-based health-information exchange will fit in with its plans.

Session 8: Protection Detail Protecting Against Breach of Electronic Protected Health Information

Today, most HIPAA/HITECH covered entities that we have worked with or talked to have a much higher vulnerability for unauthorized access to confidential electronic data (both ePHI and business confidential data) due to the various and numerous locations where the data resides.

It should be no surprise that the HITECH Act requires encryption of vulnerable ePHI. It was what was left out of the HIPAA Security Rule!

Do you encrypt all confidential data in vulnerable locations, or do you reduce the locations of your ePHI. We believe you do both. You need to encrypt vulnerable ePHI where it will continue to exist, such as on authorized laptops and pen drives. You also need to lock down user workstations for a large percentage of your users, and centralize user data on network servers that are technically and physically secure.

You also need to assess your software vendors that remotely host any your vulnerable ePHI. Are they SAS70 certified?

Finally, don’t forget what is sometimes considered to be the hardest part – documenting your compliance activities in order to demonstrate evidence of due diligence in and avoid major $$$$ penalties for negligence under the HITECH Act of 2009.
Session 8: Qualitative evaluation of health information exchange efforts
Joan S. Ash *, Kenneth P. Guappone


2. What are qualitative methods?

Qualitative research is an approach to scientific inquiry that relies on more naturalistic, humanistic and interactive processes. The methods are primarily language based, with data in the form of words rather than numbers.

The most common qualitative data gathering strategies include interviews, observation and document analysis.

4. Studying health information exchanges longitudinally

The eHealth Initiative Foundation’s Second Annual Survey of State, Regional and Community-Based Health Information Exchange Initiative and Organizations (August 2005) identified six “stages of development” of HIEs (Table 1) [4]

<table>
<thead>
<tr>
<th>eHealth initiative foundation stages of HIE development [4]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1 Recognition of the need for health information exchange among multiple stakeholders in your state, region or community (public declaration by a coalition or political leader)</td>
</tr>
<tr>
<td>Stage 2 Getting organized; defining shared vision, goals, and objectives; identifying funding sources, setting up legal and governance structure (multiple, inclusive meetings to address needs and frameworks)</td>
</tr>
<tr>
<td>Stage 3 Transferring vision, goals and objectives to tactics and business plan; defining your needs and requirements; securing funding (funding organizational efforts under sponsorship)</td>
</tr>
<tr>
<td>Stage 4 Implementing technical, financial and legal (pilot project or implementation with multi-year budget identified and tagged for a specific need)</td>
</tr>
<tr>
<td>Stage 5 Fully operational health information organization; transmitting data that is being used by healthcare stakeholders (ongoing revenue stream and sustainable business model)</td>
</tr>
<tr>
<td>Stage 6 Demonstration of expansion of the organization to encompass a broader coalition of stakeholders than present in the initial model</td>
</tr>
</tbody>
</table>

For some, an initial assessment done early in the development process could yield critical information about whether different entities are ready for this effort (Stages 1–3).

The plan and planning process

Cultural foundations

Clinician involvement

Workflow assessment

The stakeholders

Some ideas for research questions that might be explored on an ongoing basis include:
Clinician satisfaction

Other stakeholder perceptions

issue are usually about (1) problem identification (the system does not seem to fit clinicians’ workflow, so what are the problems?), (2) description (what are the concerns about confidentiality?) or (3) explanation (why is the formal decision making structure not working?). The qualitative evaluation design will be iterative. Each step, from idea generation, to design, to data gathering, to analysis, to interpretation, will likely be done more than once. During each step, it is not only permissible, but suggested, that you revisit one or more of the prior steps

5. What are qualitative research methods?

Qualitative and quantitative methods are simply different and equally valuable ways of seeking the truth. Qualitative methods can be described as inductive, subjective, and contextual, where quantitative techniques are deductive, objective, and generalized.

Qualitative methods are subjective in that they can assess how people make sense of things—how they view the world.

Qualitative methods are inductive. They are excellent choices if the evaluator wants to generate theory from observation, as they are oriented to discovery and exploration.

On the other hand, quantitative methods are deductive. They are the methods of choice when one wants to test theory after empirical observations have already been acquired, and they are oriented to cause and effect. The evaluation design is generally predetermined within a fixed timeframe.

Qualitative research can be considered subjective in that it emphasizes meanings and interpretation and tries to describe the perspectives of others so those perspectives can be understood. This kind of research relies on the researcher as the research instrument.

Quantitative research, though, emphasizes measurement and uses an outsider’s perspective, and it is vitally important that the observer remains isolated from the data.

Qualitative research is also characterized as being contextual. This is because of its naturalistic approach and ability to analyze systems holistically. It emphasizes depth and detail of findings by providing “thick” descriptions of relatively few cases.

Quantitative methods can be experimental, using a “laboratory” to isolate variables so that one can statistically analyze those selected variables to make inferences beyond the subjects under observation. They emphasize what might be called “thinner” data on a large number of cases.

Qualitative research cannot and makes no attempts to infer generalities to other populations, but hopes to understand the context itself that is being evaluated.

A recent and more pragmatic approach has been to combine the two methods as they can quite effectively complement one another.

6. Strategies for rigor

One common concern about qualitative methods is their presumed lack of validity—the data are “soft.”
A more descriptive term for this kind of validity is “trustworthiness.” Trustworthiness might be defined as the confidence that a second person, presented with the same data, would arrive at the same interpretation.

Strategies for the scientific rigor or trustworthiness of qualitative results include reflexivity, triangulation, member checking, saturation in the field, and an audit trail.

Reflexivity means that those gathering and analyzing the data recognize their preconceived biases and world views and take these into account as they proceed

Triangulation is a term from surveying, meaning that one can pinpoint a location along different axes and it is accomplished in qualitative research by using different methods, researchers, sites, times, or kinds of subjects for learning the truth.

Member checking implies that researchers check back with informants to make sure the results and interpretations seem reasonable to them

Saturation in the field means that there is a sense that enough data have been gathered, that the same patterns are seen with no new ones being identified.

The audit trail is a step-by-step record that details how the research has been conducted.

7. Two useful methods: interviewing and observation

Interviewing is a more active process that consists of engaging in some verbal discourse of varying complexity with the informant, whereas observation is a more passive activity with little to no interaction with the observed

Qualitative interviews encompass communication somewhere in between those two ends of that continuum. Usually the researcher develops a handful of questions and asks them in a way that opens the door for the interviewee to talk, tell stories, or reminisce

Informants, or interviewees, are usually carefully selected in a purposive manner

All informants should be selected for a reason, primarily based on the type of information needed. This is called “purposive” selection

The questions in a qualitative interview usually follow a “funnel” style. To “open the door,” and help the interviewee feel comfortable, a broad open ended question is asked early in the interview.

8. Focus groups

Focus groups are group interviews (the term originated as “focused group interviews”) with an added benefit that interviewees develop synergy by feeding off one another, developing or expounding on ideas from others.

The synergy that develops in a group, the so called “sharing and comparing” generates a different type of information than that which a single individual can provide. This lively interaction that must be created and sustained is critical to allowing the informants to voice unselfconscious ideas that can only come to the surface when conversing with others

9. Observation
Observation has several advantages. It can be relatively unobtrusive and non-invasive, so that busy subjects are not inconvenienced.

Often called “participant observation,” observation in the field is rarely just fly-on-the-wall shadowing. There is a spectrum of degrees of participation.

There are some rules about observation which, if followed, should allow any intelligent and interested person to succeed in doing it well. First, one must be able to focus and pay attention.

You must use rigorous methods to establish the trustworthiness of your observations. Finally, you must be introspective enough to understand your own biases.

10. What to do with the data

11. Presenting the results

The venue and format for presenting qualitative results depend on the audience and purpose of the report.

There are few standards for the format of qualitative results.

When evaluating qualitative work, experts generally look for a clearly stated research or evaluation question, a description of the context within which the work was done, articulation of the research or evaluation design, strategies used for enhancing rigor, and a clear and reasonable presentation of the results.

Session 8: The financial impact of health information exchange on emergency department care

Mark E Frisse,1,2 Kevin B Johnson,1,3 Hui Nian,4 Coda L Davison,1 Cynthia S Gadd,1 Kim M Unertl,1 Pat A Turri,5 Qingxia Chen4


Discussion Applied only to the study population, HIE access was associated with an annual cost savings of $1.9 million. Net of annual operating costs, HIE access reduced overall costs by $1.07 million. Hospital admission reductions accounted for 97.6% of total cost reductions. Conclusion Access to additional clinical data through HIE in emergency department settings is associated with net societal saving.

Clinical information was available for almost every patient whose data were accessed through the web HIE interface; the amount and type of data varied among patients based on the frequency and nature of their care in participating hospitals and clinics. Patients were offered the chance to ‘opt out’ from HIE participation at the time of every encounter at participating hospitals and clinics. The percentage of patients ‘opting out’ when consent was sought ranged from 1% to 3% across sites over the study period.

Access to HIE data was relatively low (6.8%) because the data were not integrated into and presented through the many different electronic health record (EHR) systems used within the region.
Our observations support the findings of others who suggest that providers find HIE helpful in managing non-urgent chronic medical conditions particularly for indigent populations seeking care in more than one setting. As is the case across the country, many ED visits were for chronic or nonurgent medical conditions. Our HIE system was used more frequently for ‘repeat visits'.

---

**Session 8: The Next Step in Health Data Exchanges Trust and Privacy in Exchange Networks**

By Steve D. Gravely, JD, MHA, and Erin S. Whaley, JD, MA

www.himss.org volume 23 / number 2 n SPRING 2009 n jhim 33

These new relationships between HIEs present a number of technical, policy, legal and regulatory challenges. From a technical standpoint, each HIE must have a reliable and functional interoperable architecture that meets standard protocols necessary to enable exchange with other HIEs.

What has proved to be more difficult are the legal, regulatory and policy barriers to a fully interoperable model that allows for a free, open and secure exchange of health data between HIEs.

Structuring the exchange of data between HIEs so that each HIE can comply with all of its legal requirements regarding privacy and security of data is a difficult, but important task. It requires establishing trust between all HIE participants through a legal framework for the exchange of data which is memorialized in an agreement between the parties.

It is far more efficient for all of the interoperable HIEs that wish to exchange data with each other to enter into one agreement pursuant to which all of the participating HIEs agree to the same terms, conditions and standards of performance.

**Those that will participate in the NHIN will be required to execute a Data Use and Reciprocal Support Agreement (DURSA), which provides the legal framework pursuant to which participants will exchange data**

This emerging model can best be described as a “one-to-many” model in which health data is made available through a network of HIEs to a potentially infinite number of users who may at

Unlike the point-to-point model where the HIEs exchanging data know each other and have an established relationship, the HIEs in a one-to-many model do not necessarily know each other. Their connection is the simple fact that they are all participating in the one-to-many network of HIEs

**Trust Agreement Characteristics**

- The agreement is built upon the participants’ common desire to improve the quality, effectiveness and efficiency of clinical patient care.
• The agreement reflects a multi-stakeholder process in which stakeholder representatives have had the opportunity to express their preferences and requirements and to negotiate the parameters of the relationship that are important to them.

• The agreement contains some mechanism to establish that each participant is a trusted party.

• The agreement appropriately memorializes the participants’ commitments to protecting the privacy and security of the health information that is exchanged between the networks.

• The agreement should memorialize a consensus reached by the stakeholders on the allocation of liability risk associated with the exchange of health data.

Common goal to improve clinical patient care. To be successful in connecting a variety of disparate HIEs, all HIE participants must be working toward the same basic goal. For HIEs looking to exchange health data, that goal should be improving the quality, effectiveness and efficiency of clinical patient care.

Stakeholder input. When thinking about connecting interoperable HIEs to enable the one-to-many exchange of health data, we can identify numerous stakeholders. These stakeholders include healthcare providers, patients, pharmacies, payors and the HIEs themselves, to name a few. As we discussed above, all of these stakeholders should share the same primary reason for participating – improving clinical patient care.

It is important to involve all of these different stakeholders in the development of the trust agreement for numerous reasons. First, by involving the stakeholders in the development of the agreement, you can help assure their buy-in and participation in the exchange network.

Second, each stakeholder will bring a unique viewpoint, background and knowledge to the process. These unique characteristics will help ensure that the trust agreement considers and addresses a wide variety of issues and circumstances.

At a certain point, including all stakeholders becomes a hindrance instead of a help. In such cases, it is important to include representatives of each stakeholder group in the trust agreement creation process.

Determining how to group stakeholders and identify representatives of those stakeholder groups is challenging and requires an intimate knowledge of the stakeholders’ characteristics, business operations and primary and secondary goals.

Mechanism to establish trust. In any trust agreement, the key factor is obviously trust.

The trust model, memorialized in the one-to-many agreement between all HIE participants, must establish the framework and foundation upon which trust can be given and grow. The model must be built upon clearly articulated roles for participants and their users, privacy and security safeguards, and appropriate allocation of risk and liability.

In addition, the trust agreement must establish criteria and requirements for HIE participations.

Protecting exchanged information

First and foremost, the trust agreement should acknowledge and affirm the participants’ obligations under HIPAA.

In addition to complying with the requirements of HIPAA, participants will also have to comply with any applicable state privacy laws.
For networks that cross state lines, participants will have to decide how to reconcile these varying state laws there must be some discussion of how to reconcile different state laws

The reconciliation of state laws comes into play most prominently when discussing consents and authorizations for release of data.

When the exchange network involves more than a few states, however, it is unrealistic to think that the participants in one state will agree to comply with the more stringent laws of another state. Instead, the participants may decide to require each participant to comply with the participant’s own applicable state laws. Essentially, each participant is only required to know, understand and comply with its own state law. It does not need to be familiar with or comply with the laws of another state. The participants must understand, however, that if they request data from a participant in a state with more stringent privacy laws, the requestor may be asked to provide the documentation that the responder needs to comply with its state law.

Part of compliance with HIPAA and state privacy laws is determining who will have access to the exchanged data and for what purposes. Each participant will have its own system access policies. These policies will describe what types of persons can access the system and with what permissions

In addition to deciding which HIE participant’s access policies will govern, the HIE participants will also have to determine what constitutes an appropriate reason to exchange data through the exchange network

Once the HIEs and other stakeholders are able to agree on the reasons that will support the exchange of data through the exchange network, they will need to reach consensus on what the recipient of the data is able to do with the data in the future

Allocation of liability risk. After privacy and security, allocating liability risk may be the most difficult issue on which to get the stakeholders to reach consensus

---

Session 8: The United Hospital Fund Meeting on Evaluating Health Information Exchange


George Hripcsak, MD, MS,
*Columbia University*

Rainu Kaushal, MD, MPH,
*Cornell University*

Kevin B. Johnson, MD, MS,
*Vanderbilt University*

Joan S. Ash, PhD,
*Oregon Health & Science University*

David W. Bates, MD, MSc,
*Harvard Medical School*

Rachel Block,
ISSUES IN EVALUATING HEALTH INFORMATION EXCHANGE

One challenge to agreeing on a set of methodologies is the realization that there is no single model for HIE. Focusing too quickly on a narrow set of methodologies may limit what can be learned.

In practice, HITEC will begin with six projects that will be selected via an environmental scan. The scan includes market characteristics, organization and governance, perceived barriers to success, technologic aspects, strategic vision, and capacity for research.

A framework may be organized by the subject being evaluated. For example, an HIE program may be broken into three levels: the platform infrastructure and the operations of that platform, the interventions and services that use the platform, and the program as a whole. The platform evaluation includes usage (who, how often), breadth, completeness, and confidentiality. The intervention evaluation includes quantitative and qualitative iterative assessment of expected effects on outcome, such as mortality, quality, safety, efficiency, usage, and satisfaction. The program evaluation addresses policy and regional outcomes by aggregating the evaluations of the interventions and by looking at project and program characterizations.

Prioritizing evaluation methods requires understanding the important goals of HIE and the important challenges facing HIE.

The financial impact of an HIE project can be seen from different points of view: the immediate business case differs from the full return on investment from a societal point of view.

A major driving force for HIE is improving the quality of care, although it is difficult to measure.

One challenge in demonstrating an effect on quality is that traditional health information technology interventions such as automated decision support may create a bigger impact than HIE itself. HIE projects that include the new deployment of health information technology interventions should ideally separate the cost of HIE and the incremental benefit of HIE if they are to assess the true value of the information exchange, though this will be challenging.

Evaluation of the HIE platform is important to steer the design of the project, to monitor the ongoing operation of the project, and to offer lessons to other projects. The evaluation includes issues around transferring data: completeness, timeliness, and accuracy of data transfer, success in matching patients, and confidentiality breaches. Strictly speaking, usage and impact studies do not apply to the HIE platform itself, but to the intervention (service) that uses the platform. Most of the HIE evaluation will be specific to the use cases.

A rule of thumb for evaluation is that you must spend 10% of the budget on evaluation in order to learn from what you are doing. Not all evaluations need to be done on every project, however.

Evaluation takes time.
Process measures will be available faster than outcomes

Gathering stories about how HIE has helped the health care process may create enough incentive for investors until hard evidence about return on investment arrives

Carrying out a good evaluation is a balancing act.

In summary, HIE has no effect until the platform is put to some concrete use. Every project should carry out some platform-level monitoring, such as measuring data movement by user and data type. Projects can tailor their intervention-specific evaluation to the likely impact of the intervention, the timeframe, and the evaluation budget. The AQA measures are a starting point for some quality interventions, but more sensitive indicators should be used if possible. Not every HIE project should attempt to carry out the same detailed quality study; as a rough guide about three similar evaluations need to be published. The business case for each project should be evaluated. At the policy level, the impact of broad programs like the HEAL NY Program should be assessed.

EVALUATION PRIORITIES

1. Platform evaluation

The information exchange infrastructure must undergo initial evaluation to verify its success, and it must be monitored over the course of the project

2. Usage studies

Given the primary use case, usage can be measured. Important features include who is using the system and how much

3. Immediate business case

For the project to stay in operation, the immediate business case must be assessed.

4. Assessment of clinical and administrative impact

Actual demonstration that the HIE project has caused improvement in clinical outcomes is most difficult and requires a large controlled trial. Such studies need not be duplicated in every HIE project; three or so may be enough. After that, confirmation that process variables are improving may serve as sufficient evidence that the HIE project is clinically successful

5. Unintended consequences

6. Comprehensive return on investment

The full return on investment can be calculated only after the direct and indirect costs are estimated and the clinical and administrative impact is measured and assigned a value. Unintended consequences must be included.

7. Program evaluation

Assuming the project is part of a larger program, the program itself can be assessed, and the results for the individual projects can be aggregated
Session 8: Using Health Information Exchange to Improve Public Health

Jason S. Shapiro, MD, MA, Farzad Mostashari, MD, MSc, MSPH, George Hripcsak, MD, MS, Nicholas Soulakis, MS, and Gilad Kuperman, MD, PhD


American Journal of Public Health | April 2011, Vol 101, No. 4

The electronic transfer of data for public health reporting requires each health care partner to translate data from its proprietary structures—its vocabulary or format for storing data, and its protocols for sending the data as messages—into standards defined specifically by and for various public health authorities so the data are represented consistently and can be analyzed in a uniform fashion.1,2

As part of this work, HIOs provide the organizational infrastructure, legal underpinnings, and technical expertise to enable HIE. This includes building physical data interfaces between the stakeholders and the HIO, and mapping proprietary database codes from each stakeholder to widely accepted standard vocabularies.

POTENTIAL USES OF HEALTH INFORMATION EXCHANGE

Mandated Reporting of Laboratory Diagnoses

the electronic transmission of laboratory reports can increase the efficiency of public health surveillance for high-volume diseases and the timeliness of reporting for cases requiring immediate public health action

the HIO could ensure that (1) all the data necessary for notifiable disease reporting would be integrated and mapped to standard vocabularies, (2) notifiable conditions would be identified according to a standard rule set (e.g., what constitutes positive syphilis serology), (3) a standards-based secure message could be sent to public health, and (4) an electronic log of transmissions would be maintained for audit purposes.

Nonmandated Reporting of Laboratory Data

These cases will not require public health action on an individual basis, but knowledge of the disease patterns in the community can help stakeholders guide public health messages and rule out other, less innocuous outbreaks

Another example of nonreportable laboratory data that could be very useful for public health monitoring is antimicrobial resistance patterns.

Mandated Reporting of Physician-Based Diagnoses

A counterpart to electronic laboratory reporting would therefore be enhanced reporting of suspect or confirmed clinical cases on the basis of diagnoses, procedures, or medications entered into clinical information systems and available to the local HIO infrastructure

Nonmandatory Reporting of Clinical Data

Examples include monitoring of emergency department chief complaints or discharge diagnoses and Websearch engine hits for topics such as flu-related illnesses or sales of over-the-counter flu remedies
Because identifiers are not required or mandated, 2 general approaches could be employed. As with laboratory data, individual clinical cases of interest could be found according to a standardized rules engine and then deidentified or pseudonymized before reporting. Alternatively, events of interest could be aggregated and reported to public health as counts.

**Public Health Investigation**

In this use case, public health queries the HIO for clinical data relating to a particular case that has already been identified as requiring investigation through other means (e.g., laboratory report, contact tracing).

**Clinical Care in Public Health Clinics**

In different jurisdictions, public health departments are directly responsible for providing health care for certain conditions (e.g., tuberculosis, sexually transmitted diseases) or in certain settings (e.g., schools, jails, homeless shelters). Under this scenario, public health would already be in possession of full patient identifiers used to query the HIO and would access data needed for clinical purposes, much as any other user of the HIO network.

**Population-Level Quality Monitoring**

To the extent that regional HIE can penetrate across systems of care, it offers the possibility of measuring the quality of care delivered to members of a community across health plans and providers. The HIO infrastructure could be leveraged to provide core data elements (e.g., medications, procedures, diagnoses) needed for a focused, high-quality data warehouse.

**Mass-Casualty Events**

If a record locator service (RLS) architecture was designed to receive ongoing admission-discharge-transfer messages from clinical registration systems (with updated dates of service), then the HIO would be well-suited to fulfill this patient locator function. The HIE hub would merely need to allow the designated call center(s) to query the RLS in the event of an emergency.

**Disaster Medical Response**

Electronic availability of clinical data through the Department of Veterans Affairs and other systems allowed some large health care providers to have their patients’ data available very soon after Hurricane Katrina destroyed much of their physical infrastructure.

**Public Health Alerting: Patient Level**

One could conceive of the HIO infrastructure also enabling targeted communication from public health to HIO network users. Health departments are typically constrained in their ability to disclose identifiable surveillance data for clinical purposes, but there are some examples where the duty to warn outweighs this prerogative to confidentiality.

A similar use case for patient level alerting involves antibiotic resistant organism (ARO) surveillance for early isolation of infected patients when they present to a hospital.

An ARO surveillance use case could be designed around HIE from the hospital’s local health department or neighboring facilities, so that the patient could be flagged on the HIO enterprise master patient index,
and the admitting facility could be alerted at registration that the patient should be placed in appropriate isolation.

**Public Health Alerting: Population Level**

The HIE user interface could also be a gateway to relevant epidemiologic information that the provider might be interested in.

**DISCUSSION**

Different technological approaches may be required to support public health use cases, depending on the system or network architectures for a given HIO. Examples of the varying architectures include:

1. Centralized repositories, as in the case of large hospital networks with enterprise wide electronic health record implementation (e.g., the Veterans Affairs or Kaiser Permanente health systems);

2. Hybrid peer-to-peer file-sharing models, in which all clinical information is stored at the participant organization on edge servers that sit behind their firewalls but with patient demographics stored centrally to allow patient matching and retrieval of relevant clinical information; and

3. Patient-controlled health records, in which patients determine which data to deposit into their account and who has permission to view or change them (e.g., Google Health or Microsoft Health Vault).

Having a single point of contact on the clinical side for establishing, testing, and maintaining data flows would be invaluable to public health partners.

---

**Session 8: Case Study: Patient Matching Sharp HealthCare’s Journey**

Lesli Ann Davis PM Portfolio & IS Outreach  
Debbie Coates Director, IS Revenue Management Systems  
Sarah Harrington Manager, MPI  
Elizabeth Renfree Director, IS Interoperability & Strategic Support  
William A. Spooner Sr. Vice President & CIO  
Sharp HealthCare 8695 Spectrum Center Boulevard San Diego, CA 92123-1489

With the integration of new facilities came challenges in linking healthcare information, and the patient medical record number (MRN) was key.
the new entity's patient list was matched against the cumulative Sharp database using common factors such as name, social security number, or date of birth (DOB) and others. Thus, a single patient would have multiple entity medical record numbers, in addition to their assigned overall Sharp Master Patient Index (MPI) number. During this period, when patient record anomalies were identified, each entity would manually manage the chart merge.

Additionally, we installed GE’s Enterprise Index, an application which incorporates the Initiate probabilistic matching algorithms and seamlessly integrates registration for Practice Management and Hospital Admitting. It also identifies possible duplicate registrations in the MPI. Enterprise Index checks for potential duplications when a new registration is entered. It detects a potential duplicate based on when a user enters any of the following data elements: Name, SSN, Sex, DOB, and phone number. The user is notified that the entry is a potential duplicate. The module similarly has the ability to report duplicates created and merge the records

**PatientSecure** Uses near infrared light to map an image of the vein pattern in your palm. The blood in your hand absorbs the light making your vein appear as a black pattern. Your individual pattern is then assigned an identification number which is correlated to your medical record. A vein pattern in the hand is unique to each person. For initial authentication and enrollment, the patient’s hand is scanned into a database and the scan is indexed to her MPI record. When the patient subsequently presents for care, the agent enters the patient’s DOB while she places her hand on the palm reader. If the patient is in the system, the MRN displays in the name field of Centricity Business. If patient is not in the system, the agent can then enroll the patient.

**Multi- Factor Lookup** Multi-factor lookup directs the user to a custom screen to enter 3 required data elements DOB, Age, Gender if none of the standard lookups (e.g. MRN #, Visit #) are utilized. All patients matching the data elements appear for selection and the correct patient can then be selected

The Master Patient Index is key to ensuring patients’ clinical information stay with them, providing clinicians with the most complete data for clinical decisions as well as the opportunity to reduce unnecessary duplicate testing. Both enterprise and community Health Information Exchanges (HIE) strive to provide the most complete view of a patient’s care over multiple visits and settings.

Without a reliable MPI, there is no reliable Health Information Exchange.

Our experience has led us to one conclusion; we need a National Patient Identifier. Without this, we will continue to make mistakes, invest dollars in inaccurate data sharing, and risk putting our patients in jeopardy