

NISTIR 7263

Healthcare Strategic Focus Area: Clinical Informatics

Conrad Bock
Lisa Carnahan
Steven Fenves
Michael Gruninger
Vipul Kashyap
Bettijoyce Lide
James Nell
Ravi Raman
Ram D. Sriram

NIST

National Institute of Standards and Technology
Technology Administration, U.S. Department of Commerce

NISTIR 7263

Healthcare Strategic Focus Area: Clinical Informatics

Conrad Bock
Lisa Carnahan
Steven Fenves
Michael Gruninger
Vipul Kashyap
Bettijoyce Lide
James Nell
Ravi Raman
Ram D. Sriram

*Manufacturing Systems Integration Division
Manufacturing Engineering Laboratory
National Institute of Standards and Technology
Gaithersburg, MD 20899-8260*

September 2005



U.S. DEPARTMENT OF COMMERCE
Carlos M. Gutierrez, Secretary
TECHNOLOGY ADMINISTRATION
Michelle O'Neill, Acting Under Secretary of Commerce for Technology
NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
William Jeffrey, Director

CONTENTS

ABSTRACT	2
1. INTRODUCTION	3
2. POSITION OF NIST IN ADDRESSING CLINICAL INFORMATICS	5
3. INFORMATION FLOW IN HEALTHCARE SERVICES	6
4. ELECTRONIC MEDICAL RECORDS	9
5. VOCABULARIES	13
6. EVIDENCE-BASED MEDICINE	15
7. METRICS	18
8. POTENTIAL COMPONENTS OF CLINICAL INFORMATICS	18
9. ROLES FOR NIST	21
10. SUMMARY AND CONCLUSIONS	24
DISCLAIMER	25
APPENDIX: SUMMARY OF CLINICAL INFORMATICS STANDARDS	26

Abstract

NIST has conducted several workshops and other information gathering activities, that have identified several key strategic opportunities in healthcare information technology (HIT) for NIST: clinical informatics; bioinformatics; medical devices; pharmaceuticals; biosurveillance; and enterprise modeling. This report addresses the first area, clinical informatics, as being of prime interest to NIST and focuses on roles that are suitable for NIST. The report presents three aspects of clinical informatics of most immediate interest to NIST's Information Technology (ITL) and Manufacturing Engineering Laboratories (MEL):

- electronic medical records, approaches and standards for the integration and interchange of health information;
- vocabularies, controlled collections of concepts that cover medical knowledge for particular purposes; and
- evidence-based medicine (EBM), applying up-to-date healthcare knowledge by filtering and disseminating it in a way that integrates into clinical experience and patient values.

The report then presents an initial set of components needed to realize an interoperable framework for clinical informatics: systems engineering for clarifying the meaning of information; semantic languages for expressing information with increased precision and expressiveness; model-driven architecture for translating semantically developed system specifications into multiple technologies; and interoperability and conformance testing to help ensure that standards-based solutions provide the necessary infrastructure and seamless information integration across applications within and between homogenous healthcare organizations. A number of potential roles that NIST's ITL and MEL could play in the area of clinical informatics are also identified.

Keywords: clinical informatics; bioinformatics; electronic medical records; systems engineering; semantic languages; interoperability; standards

1. Introduction

The healthcare industry is facing major challenges: increasing costs, unacceptable error rates, and dissatisfied patients and providers. Healthcare costs in the United States were about 14.9 % of the GDP - \$1.6 trillion - in 2002¹, estimated to be 1.9 trillion in 2005² and projected to rise to 3.6 trillion by 2014³. These costs are also a major concern for U.S. industry, as escalating healthcare costs are impeding our ability to globally compete. According to a February 11, 2005 issue⁴ of the Washington Post, General Motors spent \$5.2 billion on healthcare in 2004 for its employees, retirees and their families. These healthcare expenses added \$1,500 to the price of each GM car. Other U.S. automobile manufacturers face similar costs.

In addition, medical errors are of great concern. In a much debated Institute of Medicine (IOM) report entitled “To Err is Human: Building a Safer Health System,” it is claimed that “at least 44,000, and perhaps as many as 98,000, Americans die in hospitals each year as a result of medical errors.” The numbers may not be exact, but the key point is that there is considerable loss of life due to medical errors that can be avoided if appropriate safety mechanisms are put in place.⁵

Steps toward addressing the various challenges facing the healthcare system were articulated in a report entitled “Information for Health: A Strategy for Building the National Health Information Infrastructure” by the National Committee on Vital and Health Statistics, a public advisory body to the Secretary of Health and Human Services.⁶ The report emphasized the need for effective sharing and communication of data, information, and knowledge among various stakeholders in the healthcare network. This can be achieved through the effective use of information technology (IT). The need for knowledge mining, molding, and sharing is also underscored by the Institute of Medicine's report entitled “Crossing the Quality Chasm: A New Health System for the 21st Century.”⁷ The Center for Information Technology Leadership recently published a report stating that “standardized health care information exchange among health care IT systems would deliver national savings of \$77.8 billion dollars every year ...”⁸

¹ National Center for Health Statistics. Health, United States, 2004. With Chartbook on Trends in the Health of Americans. Hyattsville, Maryland: 2004. Table 116. Page 326. Available at: <http://www.cdc.gov/nchs/data/hus/04trend.pdf#116> .

² National Healthcare Expenditures Projections: 2004-2014. Office of the Actuary. Centers for Medicare and Medicaid Services.

Available at: <http://www.cms.hhs.gov/statistics/nhe/projections-2004/proj2004.pdf>.

³ *ibid.*

⁴ Ceci Connolly. U.S. Firms losing Healthcare Battle, GM Chairman Says. Washington Post. February 11, 2005. Available at: <http://www.washingtonpost.com/wp-dyn/articles/A15828-2005Feb10.html>

⁵ The report “To Err is Human” is published by the National Academy Press and is available at <http://www.nap.edu/books/0309068371/html/>. Several articles in the JAMA (Journal of the American Medical Association), July, 2000, issue debate the various findings of the report.

⁶ See <http://aspe.hhs.gov/sp/nhi/Documents/NHIIReport2001/default.htm>

⁷ See <http://www.nap.edu/books/0309072808/html/>

⁸ See <http://www.citl.org>

The Office of the National Coordinator for Health Information Technology (ONCHIT) published a report outlining a strategic plan to guide the nationwide implementation of health information technology (HIT).⁹ The report identifies 4 major goals:

- Inform clinical practice by incentivizing the adoption of Electronic Health Records (EHRs), reducing the risk of EHR investment, and promoting its use in rural and underserved areas.
- Interconnect clinicians by fostering regional collaboration, development of a national health information network, and the coordination of federal health information systems.
- Personalize care by encouraging use of Personal Health Records (PHRs), enhancing informed consumer choice, and promoting use of telehealth systems.¹⁰
- Improve population health by unifying public health surveillance architectures, streamlining quality and health status monitoring, and accelerating research and dissemination of evidence.

Of particular interest to NIST are the report's recommendations pertaining to the need for health information technology standards for clinical trial data, personal health records, as and the need for defining minimal product standards for EHR functionality, interoperability and security which will help reduce the risk of EHR investment and encourage adoption by healthcare providers.

There is growing adoption of Internet-based personal health records by patients, frustrated by lack of access to fragmented medical records maintained by each of their healthcare providers. These Web-based systems also enable patients to update their own records and provide health logs. A recent article in the Wall Street Journal describes this trend and how patient data entry improves the reliability and efficiency of health reporting during clinical encounters.¹¹ However, lack of standards in PHRs and lack of interoperability with EHRs remains an issue.

We believe that IT, in general, and the Internet, in particular, provide major opportunities for improved healthcare delivery. IT holds the promise of providing information and knowledge representation and sharing, establishing a seamless continuum of healthcare to all segments of the population. This will result in better quality healthcare provided in a timely manner, reduced care variance, reduced medical errors, and decreased costs. For example, the use of computerized patient order entry systems at Brigham and Women's Hospital in Boston, MA, has resulted in a 55 % reduction in medical errors.¹²

⁹ The Decade of Health Information Technology: Delivering Consumer-centric and Information-rich Health Care. Framework for Strategic Action. July 21, 2004. Tommy G. Thompson, Secretary of Health and Human Services, and David J. Brailer, M.D., Ph.D. National Coordinator for Health Information Technology. Available at: <http://www.hhs.gov/healthit/documents/hitframework.pdf>

¹⁰ While the EHR is maintained by the healthcare service organization (e.g., physician, hospital, HMO, etc.), a PHR is maintained by an individual.

¹¹ Sarah Rubenstein. Next Step Toward Digitized Health Records. Wall Street Journal. May 9, 2005. pp. B1, B5.

¹² As reported by Dr. Blackford Middleton in the June 2003 eHealth workshop held at NIST.

A 1993 report entitled “Health Security: The President's Report to the American People” underscored the fact that the medical field lagged behind other industries in the utilization of computers.¹³ Ten years later, this is still true. Other sectors of society have had better success in implementing IT. For example, we have witnessed considerable improvement in productivity, reduction in costs, and timely delivery of products in the manufacturing industry in the last decade due to the effective use of IT. Hence, based on NIST’s experience with the application of IT in other sectors such as manufacturing and electronic commerce, we believe that NIST can play an active role in addressing the information management problems of the healthcare industry.

A major area of the healthcare complex, clinical informatics, addresses the efficient and accurate use of medical knowledge and information in patient care settings. It addresses how data, information, and knowledge are captured, represented, used, stored, and transmitted for clinical applications. Subfields include: electronic health records (EHR); clinical decision support systems (including expert systems); medical imaging and image processing; and standards for vocabularies for representing and disseminating medical knowledge. Based on the above reports and the workshops held at NIST (one in August 2002 and the other in June 2003)¹⁴ and the information gathering activities we have conducted, we believe that clinical informatics should be the area of prime interest to NIST. In this report we focus on this area and the role within it that is suitable for NIST.

Organization of report

The report is organized as follows. We first describe the position of NIST in addressing the issue of clinical informatics. Next, in order to set the stage, we sketch the information flow among the various participants in the healthcare industry. Then we present three aspects of clinical informatics of most immediate interest to NIST in terms of the current state of the art and the gaps and problems we identified within each aspect. Next, we present metrics for the evaluation of the efficacy of IT intervention and an initial set of potential solutions applicable to the aspects presented. We conclude the coverage of clinical informatics by describing potential roles for NIST.

2. Position of NIST in addressing clinical informatics

The mission of NIST is to develop and promote measurements, standards, and technology to enhance productivity, facilitate trade, and improve the quality of life. This mission includes research addressed to the improvement of enterprise efficiency in all sectors of the economy, including healthcare. NIST’s Information Technology Laboratory (ITL), provides measurement expertise and technology to the IT sector. NIST’s Manufacturing Engineering Laboratory (MEL) contributes concepts and expertise from the manufacturing sector, including advanced work on the interoperability of manufacturing information systems.

¹³ Health Security: The President’s Report to the American People, DIANE Publishing Company, ISBN: 0-7881-1106-X, 1993.

¹⁴ A CD-ROM of the June 2003 eHealth workshop can be obtained by sending e-mail to bettijoyce.lide@nist.gov.

NIST's mission complements that of the National Institutes of Health (NIH) and major federal agencies dealing with healthcare delivery, including the full Department of Health and Human Services, the Department of Defense, the Veterans Administration, the Centers for Medicare and Medicaid Services, the Indian Health Service and others.

The institutes and centers that comprise NIH focus on research, education and dissemination of information geared towards developing new treatments and drugs for healthcare. NIH contains an institute doing research on bioinformatics that addresses computational issues in biology and a department for Clinical Research Information Systems that provides computational and data warehousing support for the storage and analysis of clinical research and protocol data. However, there is no unit within NIH, or in the other federal agencies, that addresses the critical need to efficiently and effectively administer and manage healthcare delivery and the particular role that may be played by IT in optimizing the healthcare delivery processes and reducing costs.

By contrast, NIST has developed expertise in the application of IT to manufacturing and the integration of IT standards and emerging technologies into various industry domains. This positions NIST well among the various federal research agencies to investigate and provide roadmaps and solutions for efficient healthcare delivery. Interoperability of information systems is one of the central problems in the healthcare sector. Current NIST projects addressing this problem include conformance and interoperability test development for the exchange of clinical information, as well as efforts at integrating multiple standards for healthcare and IT to provide use-case based solutions.

3. Information flow in healthcare services

The healthcare services industry generates and processes large amounts of complex information relating to the diagnosis, testing, monitoring, treatment and health management of patients, billing for healthcare services and asset-management of healthcare resources. Healthcare delivery is a collaborative process with many physicians, healthcare specialists, nursing staff and healthcare technicians from multiple healthcare organizations participating in the treatment of patients. In addition, a number of external organizations utilize healthcare information including government, insurance companies and employers who pay for healthcare services, medical researchers, life insurance companies, pharmacists, and even lawyers in malpractice suits. The major participants in the information flow process are shown in Figure 1.

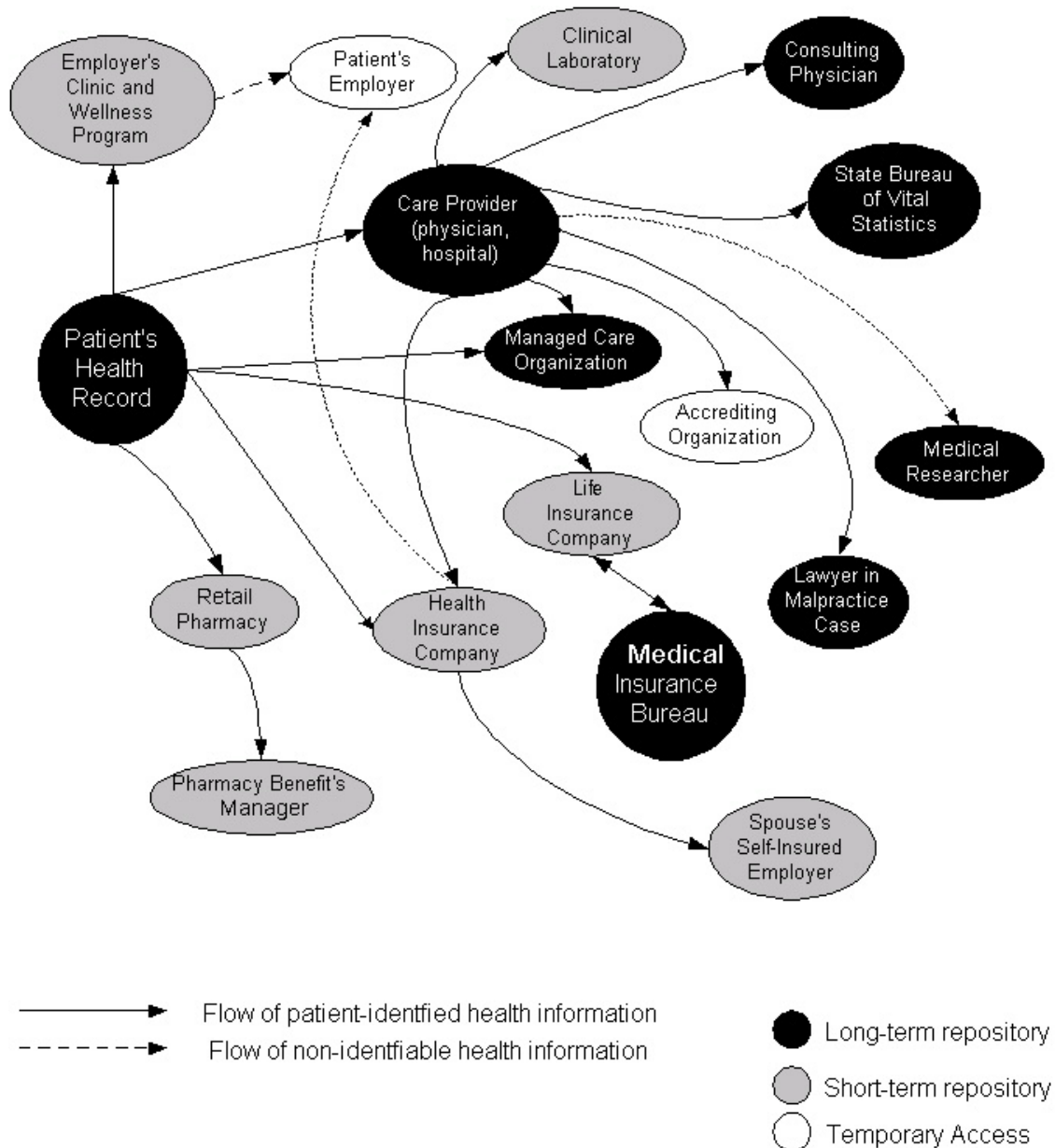


Figure 1 Healthcare Information Flows¹⁵

The classes of information created and used by the healthcare industry include:

- detailed medical records of each patient for every episode of illness or type of healthcare delivered;
- workflow for referral of patients to specialists, physician orders for diagnostic tests or procedures, and admission/discharge from hospitals;

¹⁵ Source: Figure 3.1 in For the Record: Protecting Electronic Health Information, Computer Science and Telecommunications Board, National Research Council, National Academy Press, 1997, page 73.

- detailed administrative records for managing healthcare resources (ranging from scheduling patient appointments, tracking hospital bed utilization to inventory management of pharmaceutical supplies);
- billing for healthcare services, healthcare cost control procedures and coordination of benefits; and
- research reports, clinical observations and results of clinical trials of new pharmaceuticals and new guidelines.

The major information flows within the healthcare services industry fall into the following categories:

a) *Information flows within a healthcare facility:*

- An administrator obtains patient medical history, employment and social data and health insurance information and enters these into the patient chart;
- A nurse records the patient's vital signs, medications and chief complaints for a particular visit; and
- A physician conducts an examination and writes or dictates an "encounter note" for subsequent transcription and signoff that would be included in the chart. The physician also submits billing information identifying services rendered (CPT code) and the diagnosis code (DRG/ICD) for use with insurance claims.

Larger healthcare institutions may have internal advisory groups that recommend treatment guidelines (evidence based medicine) to be followed by the healthcare staff to improve healthcare quality.

b) *Information flows between healthcare facilities:*

- The physician may prescribe an order for laboratory or diagnostic imaging test or procedure, or she/he may refer the patient to a specialist or have the patient admitted to a hospital; in some of these cases the physician would include the patient's relevant clinical history;
- The results of a laboratory test or the report and images of a diagnostic imaging would subsequently be sent to the physician; and
- Upon discharge from the hospital, the patient's discharge summary would be sent to the admitting physician.

c) *Information flows between a healthcare facility and other agencies:*

- The clinic Office Administrator submits claims to the insurance agency using the physician-supplied codes for claims processing, and sometimes may seek benefits authorization for specialized treatment for the patient;
- Health insurance agencies may seek additional justification for treatment provided to the patient or for recommended course of tests or treatment;

- Health insurance agencies may provide a list of preferred medications (formularies) and other cost containment measures to healthcare facilities;
- Managed care organizations and other payor agencies may also provide treatment guidelines to be used in the treatment of patients;
- Life insurance companies may seek a patient's medical record to evaluate the risk of a policy applicant, or determine fraud due to a known, but undisclosed pre-existing medical condition;
- Clinic nursing staff may, on occasion, have to report incidents of certain diseases to public health agencies and record pediatric immunizations with the appropriate state's vital statistics bureau;
- Medical researchers may seek medical records of patients with certain profiles for investigations; the clinic may provide the information (with patient consent) after removing patient-identifiable data;
- Malpractice lawsuits may require a healthcare facility to submit medical records of patients (with profile similar to the litigant's) to determine adherence to standards of practice; and
- Accrediting organizations may review patient records to review operational and quality standards.

4. Electronic medical records

4.1 State of the art

Penetration of information technology and automation into healthcare settings varies widely. David Kibbe, Director of Health Information Technology of the American Academy of Family Physicians stated at the NIST 2003 eHealth Workshop that only 5 % of the family physicians in the U.S. use an Electronic Medical Record (EMR) system in their daily practice.¹⁶ Predictably, large healthcare institutions such as hospitals have significant investments in computer systems, but in a majority of healthcare settings (especially small clinics) paper-based records and fax-based communications is still the norm, with computers used primarily for billing and administrative functions.

In a recent report, the Institute of Medicine identifies six major healthcare networks (including the Veterans Health Administration and the New England Healthcare Electronic Data Interchange Network) as noteworthy examples of implementations of electronic health records deployed on secure platforms.¹⁷ Yet the same report notes that these are exceptional cases, with the norm being that at most hospitals, much patient

¹⁶ David Kibbe, "From the Provider's Perspective," Designing for Quality: eHealth Development Workshop, NIST, Gaithersburg, MD, June 24-25, 2003.

¹⁷ Key Capabilities of an Electronic Health Record System, Letter Report, Institute of Medicine, The National Academies Press, Washington, DC, July 31, 2003. Available at: <http://www.nap.edu/html/ehr/NI000427.pdf>

information was written on paper and many hospitals even lacked the facilities for automated delivery of laboratory results.

In some of the leading healthcare institutions, there is a high degree of use of IT systems and of integration of information resources leading to improved flow of patient information. For example, the scheduling of the order for diagnostic imaging would have been performed online along with the submission of relevant patient clinical history.

A variety of messaging and information exchange standards (e. g., HL7, DICOM, E1460, E1467, E1381, E1394, IEEE 1073 – see Appendix) permits an enterprise to integrate the various health information systems and archive the data as an electronic medical record. Specifically, HL7 messaging standards allow disparate healthcare information systems to communicate with each other. Independent healthcare institutions can submit orders and referrals via HL7 for healthcare services for their patients. HL7 has recently released a first version of an EHR functionality specification and ASTM has recently released a specification for the exchange of patient information for continuity of care purposes such as transfer and referrals. DICOM standards enable the interchange of information between imaging systems and facilitate remote access for physicians at their clinic. With standards-based integration of information systems and authenticated remote access to reports and images, physicians can have access to the radiologist's report as well as the diagnostic images for review and patient counseling.

Mapping and correlation of patient identifiers and standards-based information interchange facilities can enable a large healthcare enterprise to provide a virtual patient record that overlies the fragmented records of the patient in the different healthcare facilities.

Figure 2 illustrates the classes of clinical information and some of the standards that link these classes to a full electronic medical record.

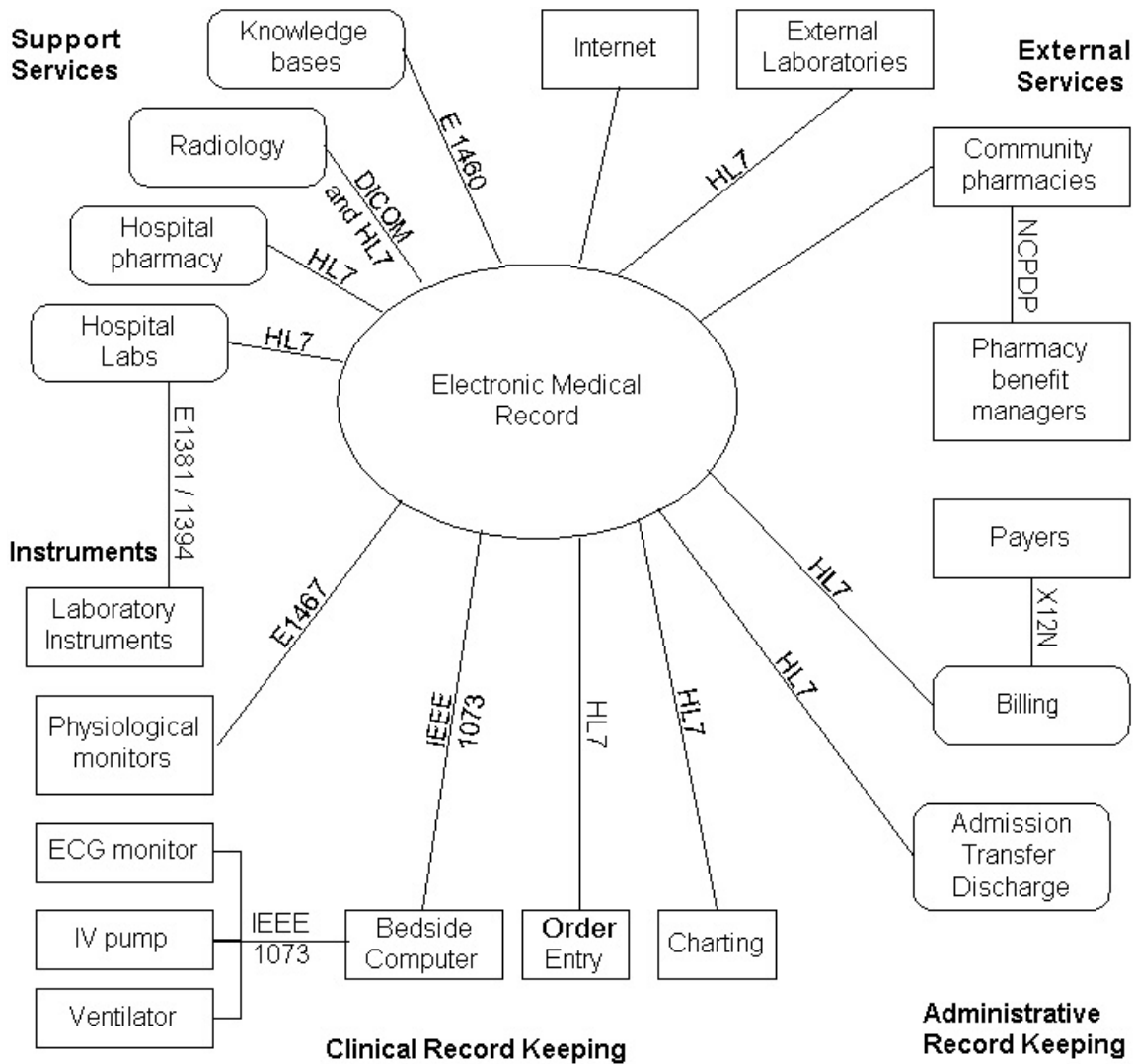


Figure 2. Electronic Medical Record¹⁸

4.2 Gaps and problems

As Table 1 in the Appendix shows, there are both significant gaps and overlaps in the standards covering electronic medical records. Both the gaps and the overlaps contribute to the major problem pertaining to clinical informatics: the lack of interoperability among the standards.

There are several standards activities working on improving the effectiveness of healthcare informatics exchange. However, for the most part there has been no systems view tying the various parts of the healthcare informatics domain together. Many of the

¹⁸ Source: U.S. Congress, Office of Technology Assessment, "Bringing Healthcare Online: The Role of Information Technologies," OTA-ITC-624, Washington, DC: U.S. Government Printing Office, September 1995.

groups that create standards have recognized the problem of achieving a fully consistent and unambiguous terminology among the standards. Potentially, there may be several standards-writing activities or organizations working in the same, or in related areas. The domain of healthcare informatics is large enough so that several activities are warranted, but these groups should work, from the start, with a strategy to connect the end results. Often, however, funds and time are limited, and competitive forces are great, so that the terminology-defining activity covers the scope of a single domain and there is no effort directed at any harmonization strategy.

Another challenge for healthcare informatics standards is that healthcare information has existed for a long time and providers have used several vehicles to enter, store, and relay information about patients. Practice tends to rely on the fixed methods of information gathering used in the past. Who has time to learn a new system, and how forgiving will those affected be when some of their data gets lost as the new system is proceeding down the learning curve? Electronic exchange of healthcare information in general, and EHR in particular, needs to recognize the diversity of techniques in use and be aware that unless the proposed new system is user-friendly, mnemonic, inexpensive, and easy to use, the new system will likely go unused. Switching over from a paper-based patient chart to an EHR is a difficult task. Healthcare facilities have large archives of longitudinal patient records. These must be accessible and incorporated into any new system. Importing this information is expensive and time-consuming, but essential for medical and legal reasons and for providing quality care to patients. Otherwise healthcare providers will have to learn a new system and incorporate new records within the EHR, but keep having to refer to the paper charts for historical patient data.

Proper timing of the standardization process is important also. Ideally, a standard should be ready for use when the technology to permit its use is available at a price that makes widespread implementation productive. If the standard is too early it becomes an academic exercise. A further drain on standards-development costs occurs when the domain being standardized requires technology development at the same time as standards development. On the other hand, developing a standard too late may be a loser if the demand for the technology covered by the standard is high, because ad hoc implementations will occur making backwards standard application unfeasible. The solution will likely be an open environment that would allow extensions to the information system.

There are several EHR system vendors; however their products are too expensive for many healthcare facilities (a large majority of which are small clinics). Despite the growing acceptance of the HL7 standard, interoperability and interchange of healthcare information remains a barrier due to different degrees of adoption of the revisions of the HL7 standard. While vendors may have developed products consistent with HL7 version 2.5, most EHR implementations are still at versions 2.1 to 2.3. The high cost of upgrading and interfacing each of the hospital's information systems often results in legacy systems that prevent adoption of newer standards. The high cost of integration of health information systems is a barrier that prevents hospitals from switching to vendors with lower costs, newer products.

The adoption of standards for information interchange will facilitate integration of disparate healthcare systems. However, implementations of healthcare data integration should not be simply geared to support human readability of medical reports, but should incorporate the formalism and details necessary for proper computer interpretability of healthcare information. Such measures would prevent the loss of information during data interchange that may otherwise occur due to differences in terms and codes and their semantics in the various healthcare vocabularies. Healthcare institutions would then be able to deal transparently with information obtained from external agencies as well as that generated by in-house healthcare information systems. Their applications could perform data mining of patient medical records for healthcare quality metrics, identify patients across populations for timely medical interventions, and check for compliance with preventive-service protocols.¹⁹

5. Vocabularies

Vocabularies are controlled collections of concepts that seek to capture medical knowledge from a particular perspective or point of view. A concept may be defined as a cluster of terms or strings that have the same and unique meaning. These vocabularies are developed in the context of various assumptions and targeted towards multiple uses:

- vocabularies such as MeSH have been developed for annotation and indexing of biomedical research articles, as in the MEDLINE database; concepts from MeSH can then be used for retrieving research articles related to those concepts; and
- vocabularies such as Diagnosis Related Groups (DRGs) have been used for cataloging and categorizing various patient populations according to criteria such as severity of illness; this finds application in epidemiological health studies and analyses.

An important role of a medical vocabulary is *knowledge representation*, i.e., the representation of medical knowledge in a standardized manner, for example in an electronic patient record. This enables data integration and interoperability across multiple healthcare information systems.

5.1 Current state of the art

There has been a plethora of vocabularies in the field of healthcare and medical informatics targeted towards different applications (e.g., billing, statistics, and epidemiology) and sub-fields (e.g., mental health, nursing, diseases). Primarily, they have been used to abstract and structure the information in clinical patient records. New emerging fields such as genomics and bio-informatics have developed their own vocabularies (e. g., gene ontology) targeted to their specific needs.

¹⁹ Balas, E. A., Weingarten, S., Garb, C. T., Blumenthal, D., Boren, S. A., Brown, G. D., Improving preventive care by prompting physicians, Arch Intern Med, 2000, Feb 14;160(3):301-8.

A list of the principal vocabularies and their target applications is as follows (see Appendix for details):

- *International Classification of Diseases (ICD) Family*, one of the oldest and archetypical coding systems for patient record abstraction. The basic ICD is meant to be used for coding diagnostic terms, but ICD-9 as well as ICD-10 also contains a set of expansions for other families of medical terms. ICD-9 has generally been perceived as inadequate for the level of detail desired for statistical reporting in the United States
- *Diagnosis Related Groups (DRGs)*, developed in the US for use in prospective payments in the Medicare program. It is an abstraction of an abstraction and applies to lists of ICD9-CM codes that are themselves derived from medical records. The purpose of DRG coding is to provide a relatively small number of codes for classifying patient hospitalizations while also providing some separation of cases based on the severity of illness.
- *Current Procedural Terminology (CPT)*, a pre-coordinated²⁰ coding scheme for diagnostic and therapeutic procedures that has been adopted in the US for billing and reimbursement. CPT codes specify information that differentiates the codes based on cost.
- *Systematized Nomenclature of Medicine (SNOMED)*, multi-axial coding system with 11 axes, each of which serves as a taxonomy for a specific set of concepts (e. g., organisms, diseases, procedures). Coding of patient information is accomplished through the post-coordination²¹ of terms from multiple axes to represent complex terms that might be desired but do not exist in SNOMED. Its goal is to provide the codes needed for electronic medical records.
- *Read Clinical Codes (RCC)*, a set of codes for electronic medical records developed to support medical record summarization and patient care applications. Both pre-coordination and post-coordination of terms are used.
- *Logical Observations, Identifiers, Names and Codes (LOINC)*, a naming system for laboratory tests and observations, being extended to include vital signs, electrocardiograph, and so on.
- *Medical Subject Headings (MeSH)*, a vocabulary, maintained by the National Library of Medicine (NLM), by which the world medical literature is indexed and searched. It is not generally used as a direct coding scheme for patient information.

5.2 Gaps and problems

The vocabularies introduced above each model pieces of information at different levels of granularity and cover different domains of healthcare information. For every vocabulary,

²⁰ Pre-coordinated coding implies that all possible concepts in a vocabulary are exhaustively enumerated

²¹ Post-coordinated coding implies that there exist a set of “atomic” or primitive concepts that can be combined to construct new concepts as required.

there exists a set of terms that are likely to overlap with, but are inconsistent with, terms in another vocabulary. This will obviously lead to conflict if information needs to be integrated or exchanged across systems using different coding schemes and vocabularies. Thus, to enable the interoperation of clinical information, it is crucial to develop tools and techniques to translate concepts across multiple vocabularies.

Sometimes codes for multiple information systems and vocabularies are mixed together in a way that might not make sense, for example, “transportation to remove abscess.” Semantics need to be modeled more clearly to disallow such combinations. The relationship of the various medical vocabularies to the *HL 7 Reference Information Model (RIM)* needs to be investigated (see Appendix).

Sometimes combinations of codes might be ambiguous, for example, combining the codes for (larynx, tube) and for a removal procedure does not specify whether it is the tube or the larynx that needs to be removed. A more precise semantic specification can help resolve such ambiguities.

In the case of vocabularies supporting post-coordination, there might be multiple expressions referring to the same concept. Inference mechanisms are required to determine the equivalence of these concepts.

6. Evidence-based medicine

Medical knowledge is increasing, but exceeds our ability to disseminate it for use by healthcare workers. Starting in the second half of the 20th century, we have been witnessing a rapid growth of medical knowledge. According to Janet Zipser, MEDLARS Management Section, NLM, “approximately 10 million records are available in MEDLINE back to 1966; 120 million searches are conducted just on the NLM computer, and 400 000 new citations are added to MEDLINE each year. To keep up with the 400 000 articles, a physician could read 2 articles each day, every day of the year and by the end of the year fall 550 years behind.”²² There is a major gap in translating such advances in medical knowledge into useful clinical guidelines and knowledge.

6.1 Current state of the art

An important approach to applying the most up-to-date healthcare knowledge is *Evidence-based medicine* (EBM). EBM is the continual dissemination of verified medical knowledge in a way that integrates into clinical experience and patient values, and is applied to individual cases:

“EBM's ultimate application is at the level of the individual clinician's decisions about managing patients. It is an explicit approach to problem solving and

²² See <http://www.nlm.nih.gov/bsd/historypresentation.html>.

continual professional learning which requires the use of current best evidence in making medical decisions about individual patients.”²³

EBM addresses the large amount of medical information available and the decreasing amount of time that physicians have to track it. In EBM, the literature is filtered for information that has passed rigorously-controlled testing and is applicable to individual patient situations, not just general knowledge of disease. Organizations such as the EBM Resource Center and the Institute for Clinical Systems Improvement (ICSI)²⁴ distill medical knowledge into a form usable by physicians.

ICSI, in particular, defines guidelines for handling cases with procedures that physicians can easily follow. ICSI’s clinical practice guidelines are systematically developed statements intended to assist practitioner and patient decisions about appropriate healthcare for specific clinical circumstances. Guidelines synthesize medical knowledge and offer formal recommendations for patient care. Many experts believe that widespread adherence to clinical practice guidelines could significantly improve patient safety, quality of care, improve outcomes, and reduce healthcare costs.

To bring the guideline aspect of EBM to fruition, NIST’s Advanced Technology Program (ATP) funded a project called the Standards-based Sharable Active Guideline Environment (SAGE).²⁵ Its aim is to support the authoring of clinical guidelines in a computable environment accessible by clinical information systems (CIS). It will present guidelines to physicians in a way that is applicable to individual patients, and will be accessible through the local CIS.

6.2 Gaps and problems

One of the key hurdles for EBM is to integrate distilled medical knowledge into clinical practice.²⁶ These problems begin when the knowledge is first written down. It is usually expressed in text-based documents that must be found and read by physicians. This is less onerous than sifting through the literature, but still very time-consuming in comparison to the average physician’s schedule. In the case of clinical guidelines, studies have shown that physician compliance improves dramatically when patient-appropriate guidelines are presented when care is being delivered. Practically, this is only possible when guidelines are embedded electronically in a CIS. Few institutions have developed in-house technologies to deploy guidelines through CISs. These efforts usually require significant resources for the implementation of even a few guidelines. Further, implemented guidelines cannot generally be shared with other institutions or systems.

One approach to addressing the integration problem is to record distilled medical information in a computable format that can be integrated into the CISs that practitioners

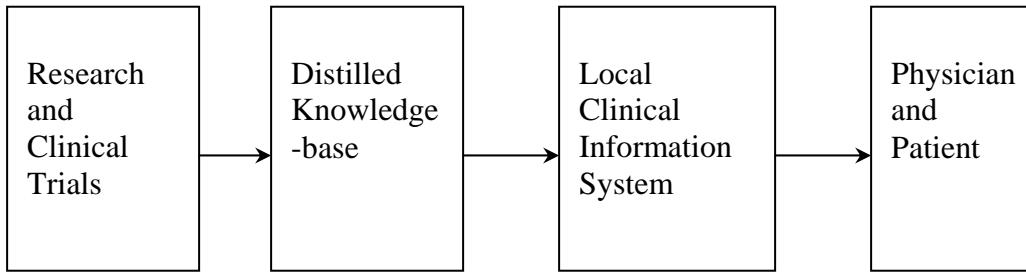
²³ Evidence-based Medicine Resource Center, <http://www.ebmny.org>.

²⁴ Institute for Clinical Systems Improvement, <http://www.icsi.org>.

²⁵ The SAGE Project, <http://www.sageproject.net>.

²⁶ See <http://www.sageproject.net/guidelines/guidelines.htm>.

already use. This folds the latest medical information into the clinical workflow. The flow of information in an ideal EBM scenario would be:



Issues arise at each of the transition points in the above process:

- Medical information might be expressed in a way that is suitable for physicians but not necessarily for computation or even implementers of information systems. It is necessary to bridge the gap between subject matter experts and information systems;
- Local CISs have their own ways of recording information, including workflow information. These are usually not compatible with each other or with any centralized medical knowledge base; and
- Physicians and patients have their own experience and opinions. In the case of guidelines, existing procedures and expectations from patients must be integrated with established guidelines.

The above problems revolve around the differing styles in which information is expressed at each stage. These styles may have the same conceptual content, but differ in the particulars of language and implementation. This causes disconnects in communication that are not necessarily due to differences of opinion, but to differences in terminology and computational implementation. It is a significant barrier to the adoption of EBM, and will impact the success of projects such as SAGE that attempt to create more uniform dissemination and integration of evidence-based medical information.

In the particular area of guidelines, medical information content describes processes and workflows. There are a number of computational languages for processes and workflows, such as BPEL, WSCI, and BPML²⁷, backed by major corporations. The healthcare standards community constructed process languages separately from the mainstream, such as the one used in SAGE²⁸ and the HL 7 Act model.²⁹ It is unlikely that process languages developed in the medical community will become widespread given that the implementations will most likely be provided by major corporate vendors.

²⁷ See <http://www-106.ibm.com/developerworks/library/ws-bpel>, <http://www.w3.org/TR/wsci>, and <http://www.bpmi.org/specifications.esp>.

²⁸ Campbell, J., et. al., "The SAGE Guideline Model," Available at: <http://www.sageproject.net/links/Sage%20Guideline%20Model%20Paper.pdf>.

²⁹ See http://www.hl7.org/library/data-model/RIM/modelpage_mem.htm.

In general, the insulation of healthcare information systems from the mainstream will lead to chronic incompatibility with other systems and increased cost due to unnecessarily specialized implementations.

7. Metrics

Metrics are important to determine the efficacy of information technology for increasing healthcare quality while reducing costs. The Agency for Healthcare Research and Quality (AHRQ) has published indicators for inpatient quality and patient safety.³⁰ At present, these metrics are not well characterized for the application of IT in the healthcare domain. Hence, we need to conduct further research to provide metrics for evaluating the ability of information systems to improve healthcare quality.

8. Potential components of clinical informatics

Data becomes information when it is interpreted in its proper context. Metadata provides a descriptive framework for the data and is the overhead for effective information interchange. Metadata consists of: 1) machine interpretable representational notation, 2) data descriptors based on terms in a domain vocabulary, and 3) contextual semantics. Communication facilities enable information to be manipulated, transmitted, shared and integrated.

In well-understood (homogeneous, highly integrated) application scenarios, data can be transmitted in the raw without the overhead of an explicit metadata framework. For example, a medical specialist can return results of a test without sending descriptions of all the equipment that produced the test results. More often, information exchange can cause interoperability problems in healthcare (just as in other domains), because raw data without the metadata framework can be misinterpreted or overlooked. For example, the format for describing test results used by a laboratory (see Figure 1) might be so different from the format used by a clinic where the patient is being examined that results may be confused or misinterpreted.

Semantics and information models help improve information interchange. These models are developed by subject matter experts to define the terms and relationships of concepts and behaviors in a particular domain. Incorporation of these terms and semantics in the metadata of the information transferred between systems enables automation to interpret the data and translate it into site and system-specific data representations.

Effective clinical information interchange requires the following four components which are described in the following sections:

1. Systems engineering techniques to develop semantic models
2. Languages for expressing semantic information

³⁰ See <http://www.qualityindicators.ahrq.gov>.

3. Model-centered architecture for platform-neutral system design
4. Testbeds for assessing interoperability and conformance

NIST's expertise in these areas can be applied to improve the approaches for interoperability of healthcare information systems.

8.1 Systems engineering

Semantic and information models should be developed using a systems engineering³¹ process that ensures that the models reflect the requirements and perspectives, vocabularies and transactions of all relevant players in that domain.

Systems engineering ties each element in a system to stakeholder requirements and enables the evaluation of alternatives on how to address those requirements. It begins with the goals and requirements of a desired system independently of which parts of it are assigned to human, hardware, or information components, and then incrementally translates these goals into system designs that satisfy the goals. This is important for healthcare applications, which use many kinds of elements to achieve their purpose.

For example, a model for a patient referral process may include terminologies and relationships reflecting the requirements and viewpoints of:

- healthcare providers (who initiate the referral request),
- benefits eligibility coordinators or payers (who review the request and authorize the payment for services),
- clinical information system vendors (whose system will export relevant data from the patient's electronic medical record associated with the referral), and
- internal, industry and regulatory privacy guidelines (which define the scope of patient information relevant for processing referral requests).

Given the complexity and range of medical vocabularies, a systems engineering approach would help reduce the size and scope of information and semantic models for specific healthcare transactions and healthcare information interchanges.

8.2 Semantic languages

Semantic approaches to expressing information raise the bar on precision and expressiveness, and in particular support more accurate reflections of clinical reality in vocabularies and guidelines. Advances in modeling, ontology, and knowledge representation can generally be applied to clinical systems. Leading industry standard development organizations have been developing tools and languages that help define information models and semantic concepts. For example, OMG defines the Unified

³¹ Journal of The International Council on Systems Engineering, Wiley, www.incose.org.

Modeling Language (UML), with improved support for semantic expressiveness.³² W3C is developing the Ontology Web Language,³³ for recording concepts independent of implementations, such as XML Schema. ISO is defining an ontology that focuses on unambiguous description of processes, the Process Specification Language.³⁴ All of these can be extended and applied to the healthcare domain.

8.3 Model-centered architecture

Disciplines concerned with software quality are recognizing the benefits of semantic-based development in addressing implementations that span platforms and organizations. Model-driven architecture,³⁵ (MDA) is a platform-neutral approach for software development that emphasizes the development of models that capture application and domain concepts, relationships and behaviors and isolates functional specification from platform-specific implementation considerations. Also known as repository-centered development,³⁶ this form of development can be used as the basis for systems that operate on multiple technologies and that are accessible through multiple user interfaces styles. Most importantly for clinical informatics, these techniques provide a common way to describe vocabularies and guidelines, from authoring to deployment in clinical information systems. A generalized form of model-centered architecture is computation-independence, which covers those aspects of a total system that will be handled manually or by hardware, as well as software, and elements in the environment of the system.

The use of standard models and service interfaces will make medical knowledge accessible to a wider set of tools and user bases, than is possible in today's proprietary systems. Functional specifications could be automatically translated to various platforms facilitating integration with existing clinical information systems. Hospitals could customize information for their local systems while maintaining compatibility across organizations.³⁷ Moreover, authors of evidence-based medical information would find their products have wider use.

8.4 Interoperability and Conformance Testing

The creation of a viable national healthcare information infrastructure in the U.S. depends on all parties involved (consumers, healthcare professionals, researchers, and insurers) having systems, tools, and information that are complete, correct, secure and

³² Object Management Group, "UML 2.0 Superstructure Specification," <http://www.omg.org/cgi-bin/doc?ptc/03-08-02>, August 2003.

³³ W3C, "Web Ontology Language (OWL) Guide Version 1.0," <http://www.w3.org/TR/owl-guide>, August 2003.

³⁴ Gruninger, M., "PSL 2.0 Ontology -- Current Theories and Extensions," <http://www.mel.nist.gov/psl/pslontology>, 2003.

³⁵ Object Management Group, "Model-driven Architecture", <http://www.omg.org/mda>.

³⁶ Bock, C., "UML without Pictures," IEEE Software Special Issue on Model-driven Development, <http://csdl.computer.org/comp/mags/so/2003/05/s5toc.htm>, September/October 2003.

³⁷ Denno, P., et. al., "Model-driven Integration Using Existing Models," IEEE Software Special Issue on Model-driven Development, op. cit.

interoperable. Much depends on the availability of development, adoption and interoperability of healthcare information standards that are complete and implementable.

NIST can assist the healthcare industry by developing tools and techniques to ensure that clinical informatics standards are complete and testable and by providing conformance interoperability tests by which vendors may validate their systems. Testbeds that utilize semantic information models would be well positioned to validate interoperability and conformance to standards of vendor products regardless of their design, platform or implementation.

9. Roles for NIST

The potential roles that NIST could play in the area of clinical informatics are described below, grouped into four major categories:

Clinical informatics research activities:

- Investigate semantic consistency issues pertaining to HL7's Functional Electronic Health Record Model (EHR);³⁸
- Develop reference architecture to confirm semantic consistency for healthcare information interchange;
- Investigate information models and standards such as the Reference Information Model (RIM) and their relationship to various medical vocabularies;
- Investigate approaches to enable semantic translations of concepts across multiple vocabularies; design and develop semantic distance metrics in the context of healthcare information and use them to identify semantically closest translations of a concept into different vocabularies;
- Apply and extend semantic concepts and languages to representative examples of guidelines, in particular for evidence-based medicine, aiming to increase the precision and expressiveness, especially as used in the SAGE project;
- Use results of above to compare and link representative examples of existing healthcare guidelines;
- Investigate standards for representing clinical trial data, including the incorporation of image data;
- Develop extensions to the current DICOM family of imaging standards to support semantics-based interoperability linking all images associated with a single patient into an integrated master patient record across specialty domains (e.g., radiology, pathology, ophthalmology, dermatology, gastroenterology, etc.);

³⁸ DHHS had asked the Institute of Medicine (IOM) to provide guidance for developing an EHR. IOM drew up the Core Functional Specifications in 2003. HL7 collaborated in this endeavor and developed an EHR System Functional Model in 2004 that is being reviewed for eventual submission as a standard.

- Develop conformance tests for relevant messaging and information exchange standards; and
- Participate in interoperability testing activities that seek to provide system-wide standards-based solutions.

Standards-related activities:

- Increase participation in healthcare standards development organizations;
- Provide aid in the harmonization of different healthcare standards;
- Engage stakeholders in the healthcare standards development process;
- Develop strategies for enabling adoption of healthcare standards;
- Help define standards that are correct and testable;
- Develop tests and operations for running healthcare standard conformance tests; and
- Establish a leadership role in developing accrediting mechanisms for healthcare standard interfaces.

Healthcare information infrastructure development activities:

- Create model of interoperability of laboratories and other healthcare facilities for testing healthcare standard interfaces;
- Develop a prototype healthcare informatics knowledge library, including rules, alerts, clinical guidelines, terminology servers, and use it to populate medical information systems; and
- Develop a reference architecture for describing clinical informatics model implementations in a variety of settings (e. g., small physicians' offices, large healthcare delivery organizations, insurance providers), with expandability to current standards.

Social setting related activities:

- Help develop technology insertion environments that allow small practices to adopt/adapt technologies in common use by large organizations;
- Help develop best practice studies (studies that describe measure of information systems effectiveness and efficiency) as a means to establish migration models to go from current state to next state;
- Set up a community process and portal for developing a healthcare information infrastructure:
 - Create and maintain a portal to act as a forum for participants seeking guidelines on how to pursue tasks related to the Healthcare Information

Infrastructure, such as standardizing EHR concepts, vocabularies and formats, integration approaches and methodologies;

- Publish standardized data sets for benchmarking and research purposes; and
- Establish a clearinghouse to promote and standardize emerging and promising approaches for creating an interoperable healthcare infrastructure and to help harness the enthusiasm and creative energy available that needs to be directed towards accelerating the growth of this promising area.
- Create a community process for developing and deploying software:
 - Identify a set of concrete applications that provide a compelling value proposition for emerging technologies;
 - Develop detailed specifications of these applications, along with a reference architecture and a component/service model;
 - Consider initiating an open organization of international software developers whose charter is to develop and revise technology specifications, reference implementations and technology compatibility kits, in the manner of the Java Community Process; and
 - Consider initial support for this process, which over time will evolve to a formalized process overseen by representative from many healthcare provider organizations and software vendors.

Organizations that NIST may potentially work with include:

- Healthcare providers such as Cleveland Clinic, Kaiser Permanente, Mayo Clinic, and Union Hospital, Elkton, MD;
- Healthcare Organizations such as Partners Health Care;
- Government agencies such as the National Institutes of Health (NIH), Department of Veterans Affairs (VA), Department of Defense (e.g., Telemedicine and Advanced Technology Research Center), Food and Drug Administration (FDA), Department of Homeland Security (DHS), and Agency for Healthcare Research and Quality (AHRQ);
- Standards organizations such as American National Standards Institution (ANSI), International Organization for Standards (ISO), Object Management Group (OMG), Health Level Seven (HL7), and Clinical Data Interchange Standards Consortium (CDISC);
- Professional Organizations such as the Radiological Society of North America; and
- Various vendors.

10. Summary and Conclusions

The following areas of healthcare provide major opportunities for NIST:³⁹

1. *Clinical Informatics*, the efficient and accurate use of medical knowledge and information in patient care settings;
2. *Bioinformatics*, the computational tools and approaches for expanding the use of biological (e.g., proteomics, systems biology, microarray analysis, nanobiosensing, etc.) and non-clinical medical data;
3. *Medical Devices*, an industry that encompasses a range of manufacturing engineering issues;
4. *Pharmaceuticals*, an industry with major issues of product design, manufacturing process and supply chain management improvement;
5. *Biosurveillance*, technologies for public health surveillance information in response to disease outbreaks and bioterrorism attacks; and
6. *Enterprise modeling*, the application of simulation technologies developed in NIST/MEL/ITL to healthcare systems.

The report addresses the first area, clinical informatics, as being of prime interest to NIST and we focus on the roles within it that are suitable for NIST. Three aspects are of most immediate interest to NIST in terms of the current state of the art and the gaps and problems identified:

1. *Electronic medical records*, approaches and standards for the authorized integration and interchange of health information;
2. *Vocabularies*, controlled collections of concepts that cover medical knowledge for particular purposes; and
3. *Evidence-based medicine (EBM)*, applying up-to-date healthcare knowledge by filtering and disseminating it in a way that integrates into clinical experience and patient values.

An initial set of potential solutions applicable to the above aspects of clinical informatics includes:

1. *Systems engineering* as a tool for clarifying the meaning of information;
2. *Semantic languages* for expressing information with increased precision and expressiveness;
3. *Model-driven architecture*, a technique for translating system specifications into multiple technologies; and

³⁹ Since the primary authors of this report are members of ITL and MEL, the focus reflects the expertise in these laboratories.

4. *Interoperability and conformance testing* to ensure that standards-based solutions provide the necessary infrastructure and seamless information integration across applications.

The potential roles that NIST could play in the area of clinical informatics are identified in categories pertaining to:

1. *Clinical informatics research;*
2. *Standards development, harmonization and conformance testing;*
3. *Healthcare information infrastructure development; and*
4. *Activities affecting the social setting.*

Disclaimer

Any mention of organizations, agencies, vendors or commercial products in this document is for illustration only. It does not imply sponsorship, contract, recommendation or endorsement by NIST.

Appendix: Summary of Clinical Informatics Standards

Table 1 is the result of a review of clinical informatics standards. The list is meant to be representative of current efforts in clinical informatics standards. It is by no means a comprehensive list. The Consolidated Health Informatics (CHI) Initiative of the Department of Health and Human Services aims to maintain a comprehensive standards list needed for smooth interoperability of the healthcare enterprise. The URL for CHI is: <http://www.hhs.gov/healthit/chiinitiative.html>.

The columns in Table 1 refer to the various aspects of information transfer standardization. The rows refer to the major subdisciplines, activities or processes in the healthcare domain.

Note that not all cells contain entries, because: either no standard exists or could be identified; we have misinterpreted the scope of some standards; or there is no need to standardize that aspect of the healthcare activity.

Following the table, the standards referenced are briefly described and characterized.

Standards domain \ Healthcare domain	Coded Data Elements: Thesaurus, Terminology	Reference Concept Model	Information Transfer, Message Structure	Information Formats, Languages	Architecture	Meta Schema
Laboratory Information	LOINC RELMA E1712		ASTM E1238 HL7 XML	XML E1238		CANON Group
Clinical Documents	MEDLINE				CDA (HL7)	CANON Group
Patient Medical Record Information (PMRI)	HL7 E1633 SNOMED GALEN	HL7 (RIM) E1384 GALEN-Core	HL7	E1633 GRAIL GALEN IR	E1384	E1384 MESH
Prescriptions			SCRIPT	SCRIPT		
Pharmaceuticals	RxNORM (NIH) NDF-RT(VA)					
Reference and Knowledge	OpenGALEN UMLS	GALEN-CORE		E1460 GRAIL GALEN IR		
Diseases	ICDO ICD9CM ICD10					
General	MSDS (ANSI) MEDLINE SNOMED SDM	SNOMED ?	SDM IEEE P1157 IEEE 1073	SDM E1467		
Digital Imaging	DICOM	DICOM	DICOM E1713	DICOM	DICOM	DICOM
Equipment Manufacturing		STEP	STEP X12 XML IGES	STEP IGES		
Inter-Instrument Information Transfers			E1381 E1238	E1394		
Payer-Provider Communication			X12			

Table 1: Standards Relevant to Clinical Informatics

Brief descriptions of standards

ANSI X12 (American National Standards Institute) Electronic Data Interchange (EDI). Provides for the application-to-application exchange of business information in a structured, standard format without human intervention. There are several EDIs required by HIPAA, The Health Insurance and Accountability Act:

270--Eligibility Inquiry

271--Eligibility Response

276--Claim Status Inquiry

277--Claim Status Response

278--Certification Request and Response

820--Premium Payment

834--Enrollment

835--Claim Payment

837--Claim Encounter

ASTM Standards (American Society of Testing and Materials). These standards are developed by ASTM Committee E31 – Health Care Informatics. Only the standards from that committee that are relevant to this work are listed:

E1238--Standard Specification for Transferring Clinical Information between Independent Systems. Defines message structure for all clinical data such as history, consultant notes, and obstetrical ultrasound. Used as a basis for HL7.

E1381--Specification for the Low-Level Protocol to Transfer Messages between Instruments and computer Systems. Companion to E1394, this standard defines the low-level protocol to establish plug-together compatibility between instruments and computers.

E1384--Standard Guide for Description for Content and Structure on an Automated Patient Health record. Identifies common patient information elements at all care sites; such as, demographics, diagnostic-test results, clinical orders, and medications. Defines a structure of these information segments. [3]

E1394--Standard Specification for Transferring Information between Clinical Instruments and Computer Systems. Companion to E1381, this standard deals with the electrical and mechanical connections between computers and the clinical instruments and the methodology for establishing communication.

E1460--Standard Specification for defining and Sharing Modular Health Knowledge Bases. A standard syntax for description and sharing of the healthcare knowledge bases. The Arden Syntax for Medical Logic Modules is now owned by the HL7 activity.

E1633--Specification for the Coded Values Used in the Automated Primary Record of Care. Catalogs the value sets for those data elements in E1384.

- E1712--Specification for Representing Clinical Laboratory Test and Analyte Names. Universal definition of the terms used for clinical laboratory tests and analyses.
- E1713--Specification for Transferring Digital Waveform Data between Independent Computer Systems. Takes digitally recorded electrophysiological waveform data and defines the message to be transmitted among between laboratories and clinics, instrumentation and computer systems.
- E1769--Guide for Properties of Electronic Health Records and Record Systems. Defines the requirements, properties, and attributes of a computer-based patient record; in computer-sensible form.

CANON Group, a group of researchers addressing technical issues about various medical vocabularies seeking coherent conceptual representation across applications and subject domains. Five aspects are important to adequately represent concepts across the medical domain: controlled vocabulary (ontology); typology (organizing terms into semantic domains); concept model (combine simpler concepts into more complex ones); notation (the way the models are represented, such as conceptual graphs); and granularity (dealing with level of definitions and coarseness of semantic classes).

DICOM (Digital Imaging and Communication), developed by the American College of Radiology - National Electronics Manufacturers Association committee, ACR-NEMA, to exchange imaging information for PMRI and within a medical center. Expanded in scope drastically, see SDM below. Consists of a hardware interface, data dictionary, object-oriented data model, and a set of commands about the description of how the image was made including what settings were used.

GALEN or Open GALEN (General Architecture for Languages Encyclopaedias and Nomenclatures in Medicine), an open source continuation of the original European Community project. It consists of nineteen extensible, application-independent concepts that form a framework for building and integrating terminology from different systems. Contains a concept reference model (CORE) and an intermediate representation, and an integration-language kernel. [<http://www.opengalen.org>]

GALEN CT (Clinical Terminology), a language-independent concept representation developed as a foundation for the next-generation, multilingual coding systems. Includes a Master Notation for medical terminology and Coding Reference (CORE) model.

GALEN IR (Intermediate Representation), acts as a high-level language for GRAIL, and it is an abstraction between clinical end users and underlying technology.

GRAIL (GALEN Representation and Integration Language kernel), the description logic used in Open GALEN clinical terminology. Techniques include *particularization* for composing more specific concepts, and *sanctioning*, or type constraints, to indicate whether a particularization is sensible and to point out if it is a redundant concept.

HL7 (Health Level 7), ANSI accredited organization preparing standards that focus on messages that are to be communicated between heterogeneous systems within a medical center relating to patient-medical-record information. [<http://www.hl7.org>]. The "7"

represents the seventh layer in the ISO OSI (International Organization for Standardization, Open System Interconnection Reference Model) ISO 7498 architecture. HL7 is the most widely implemented healthcare data-messaging standard.

HL CDA (Clinical Document Architecture), document markup standard that specifies structure and semantics of clinical documents. The DCA can include text, images, sounds, and other media content. The clinical document can be sent inside an HL7 message. Information comes from the HL7 RIM and is implemented in the Extensible Markup Language, XML.

HL7 encompasses the following Methodology standards:

- *HL7 RIM (Reference Information Model)*, the source of the data content of HL7 messages. Each coded attribute in the RIM requires a vocabulary-domain specification. The RIM consists of items such as: subject areas, scenarios, classes, attributes, use cases, actors, and trigger events.
- *HL7 Vocabulary Domain Specification*, identifies, organizes, and maintains terms used in coded fields of HL7 messages. It defines a vocabulary domain for each coded-entry message field. A domain specification is a formal ontology for a concept, and a set of allowed values for a coded field.

ICD (International Classification of Diseases), a coding standard to promote consistent worldwide reporting of causes of death and incidence of disease. ICD-O covers oncology; ICD-9-CM, clinical modifications, including diagnosis codes; ICD-10-PCS is intended to replace revision 9 by replacing the diagnosis part with a procedure-coding system, the PCS. ICD-10-PCS was developed by 3M Health Information Systems through a project funded by HCFA, the Health Care Financing Administration.

IEEE MEDIX (Institute of Electrical and Electronic Engineers, Medical Data Interchange Standard), a committee formed to draft a standard set of hospital-system interface transactions based on ISO standards for all seven layers of the ISO OSI Reference model (International Organization for Standardization, Open System Interconnection Reference Model, and ISO 7498). IEEE P1157 was prepared to exchange data between hospital computer systems. IEEE 1073, Standard for Medical Device Communications, has produced a family of documents covering the seven-layer communications requirements for the Medical Information Bus, a communications service for bedside devices in major hospital areas, such as patient rooms, emergency rooms, intensive care units, and operating rooms.

LOINC (Laboratory (originally Logical) Observations, Identifiers, Names, and Codes), a public domain set of codes and names intended for reporting laboratory test results and clinical observations. See RELMA. The Regenstrief Institute in Indianapolis IN maintains the LOINC database [<http://www.regenstrief.com>].

MEDLINE, NLM's bibliographic database of citations and abstracts from nearly 4500 biomedical journals published in the United States and worldwide. Coverage extends

back to the mid-1960s. All citations in MEDLINE are assigned Medical Subject Headings (MeSH), from NLM's controlled vocabulary to assist users in their searches [<http://www.nlm.nih.gov/pubs/>].

MeSH (Medical Subject Headings), part of the UMLS activity. Developed to reflect the concepts appearing in medical literature. See *MEDLINE*.

Read Clinical Codes, a set of codes designed to encode information on electronic medical records. The first version was developed by James Read in the 1980s and adopted by the British National Health Service (NHS). Version 2 was developed to meet hospital needs to cross-map their data to ICD-9. The NHS is expanding the content of the Read Codes to ensure that terms needed by practitioners are represented. In 1999 *SNOMED* and *Read Codes* agreed to merge their efforts to have a single terminology suitable for clinical-patient records.

RELMA (Regenstrief LOINC Mapping Assistant), assists a user to map the local laboratory code to the LOINC code. RELMA requires human assistance in choosing among candidate matches for terms. [<http://www.regenstrief.com/loinc>].

SCRIPT (Standard for the Exchange of Prescription Information), provides communication between providers and pharmacies, and for patient-medical-record information. SCRIPT was developed by the National Council for Prescription Drug Programs. DICOM and SCRIPT were recommended by the National Committee on Vital and Health Statistics (NCVHS) to the US Department of Health and Human Services to support the HL7 standard. With HL7 as a core, SCRIPT and DICOM exchange market-specific Patient Medical Record Information (PMRI)

SNOMED International (Systemized Nomenclature of Human and Veterinary Medicine), a worldwide effort to create a machine-readable, standardized medical nomenclature of basic concepts that would encompass the domain of human and veterinary medicine. *SNOMED* lacks semantic constraints and a framework for discourse, therefore bridges such as SDM, below, are constructive. Concepts are grouped according to similar characteristics rather than alphabetically. These are called axes or dimensions and there are eleven of them. Concepts are related hierarchically on the same axis and non-hierarchically on different axes. There are relationships, linkages, between concepts, such as between a disease and its cause. The result is a comprehensive information model called the *SNOMED Data Model*. *SNOMED* was developed by several American medical associations and translated into several languages--hence the word *International* in the title. See *Read Clinical Codes*.

SDM (SNOMED DICOM Micro-glossary), a combination of the DICOM message standard and the *SNOMED* computerized lexicon, making it part protocol and part database. DICOM's scope expanded from radiology to the general medical domain.

UMLS (Unified Medical Language System), developed by the National Library of Medicine. Large repository and meta-thesaurus of biomedical concepts (~300 000) to support retrieval and integration of information from disparate sources.