Chapter 11
Healthcare Data Standards and Exchange

Timothy D. Imler, Daniel J. Vreeman, and Joseph Kannry

Application Exercise

Based on the clinical vignette:

- Describe how you would identify patients from disparate health care systems to uniquely identify them. What standards would you use?
- Select three standards that you would utilize for sharing data and explain what issues might arise by utilizing these. Describe how the standards are structured.
- Determine how a message would be sent for sharing and why you would prefer a specific mechanism for delivery.
- Based on standards you selected, explain how you might advocate for a change in the standard based on your clinical needs.
- Describe the four different ways that changes are made to healthcare standards.

T.D. Imler, MD (✉)
Division of Gastroenterology and Hepatology and Division of Biomedical Informatics, Indiana University School of Medicine and Regenstrief Institute, 702 Rotary Circle, Suite 225, Indianapolis, IN, 46202, USA
e-mail: timler@iu.edu

D.J. Vreeman, PT, DPT, MSc
Regenstrief Institute, Inc., Center for Biomedical Informatics, Division of General Internal Medicine, Indiana University School of Medicine, Indianapolis, IN, USA

J. Kannry, MD
Division of General Internal Medicine, Icahn School of Medicine at Mount Sinai, New York, New York, USA

J.T. Finnell, B.E. Dixon (eds.), *Clinical Informatics Study Guide: Text and Review*, DOI 10.1007/978-3-319-22753-5_11

Biomedical Informatics Core Competencies

1. Acquire professional perspective
2. Implement, evaluate, and refine
3. Produce solutions

Case Vignette

As a clinical informatics champion for your hospital system you are asked to develop a method for sharing the data with other healthcare institutions within your city. These data are important because many patients rotate between the various hospital systems and vital information is often lost or interventions repeated due to the limited access to the data. Are there mechanisms for sharing data? How would you identify the patients across disparate electronic health record systems? Which standards would you select for sharing, and why?

You are asked to setup a new electronic clinical information system for your practice. You will have a mix of inpatient and outpatient encounters for your practice. How do you plan to document these encounters and more importantly bill (get paid) for what you are doing? What standards are in required? How to you make sure that your special clinical needs are covered?

Introduction

Without any conscious awareness, we depend on many kinds of standards every day. Weighing yourself in the morning, plugging your plethora of gadgets into electrical outlets, getting cash from an ATM, connecting your laptop to wireless networks at work, home, and the coffee shop – these ordinary activities are almost effortless, in large part because of accepted standards. Standards are powerful enablers of technological progress anywhere that variation creates inefficiencies. Standards are shared formats or definitions that constrain the possible variations down to a normalized form or common meaning. They make it possible to communicate efficiently, swap out components, and build complicated systems out of many interacting parts [1].

Standards are important to healthcare in many areas, from the most basic scientific measurement standards to standards of clinical practice (e.g. guidelines (See Chap. 5)). Our focus in this chapter is on the role of healthcare standards (both technical and clinical) in facilitating clinical data exchange. We emphasize the standards used uniquely in healthcare, and while important, we consider more general scientific and technical standards (e.g. cryptography, network protocols (See Chap. 10)) out of scope. Specifically, we will focus on standards for specifying persons (e.g., patients, doctors), transactions (e.g., encounters, medication orders), and data (e.g., systolic blood pressure) in health care.

Health information technology has the potential to improve the quality and efficiency of care, but the success of these systems depends in part on the clinical data within their purview. Too often, clinical information systems function like "islands" or "silos" and cannot get the data when and where it is needed. A major reason for this problem is that many systems cannot communicate effectively (i.e. they are not "interoperable") with each other because they lack shared conventions for the syntax (structure) and semantics (meaning) of clinical data. Each one stores patient data elements in a different format and with different codes and names for the same concepts. The only way to efficiently move and aggregate clinical data is by adopting data exchange standards.

How Are Healthcare Data Standards Created and Maintained?

There are four common methods to create standards [2, 3]. *Ad hoc* standards are developed as people and organizations come together and agree to use a common but informally developed specification. For example, in the early days before there was a standard for describing the contents of a digital radiologic image, the American College of Radiology/National Electrical Manufacturers Association created an ad hoc method for picture archiving and communication system (PACS) images to be sent to electronic health records systems. This method eventually became known as DICOM. *De facto* standards emerge as one earns a large enough for a critical mass of adopters to make its system the standard. An example of a de facto standard is harder to find in health care, but a familiar example is the ubiquitous Microsoft Word Binary File Format (*.DOC) for word processing documents. Given the widespread utilization through the years, even outside of Microsoft software, this file format for electronic documents has become de facto standard. Government agencies or other authoritative bodies can also create standards and require (*mandate*) their use in certain contexts by fiat, a formal authorization for usage. An example of mandated standards include the U.S. Standard Certificate of Death that was required by the National Center for Health Statistics, a part of the Centers for Disease Control and Prevention (CDC). Many of the key health data standards are developed by a *consensus* process. The consensus process can be closed or open though an "open standards development policy" is preferred. The American National Standards Institute (ANSI), a private nonprofit organization, has developed a formal consensus process with open balloting and public review and an accreditation program for organizations that develop standards. Not all standards are amenable to ANSI's process. It would not be feasible nor desirable to have every new concept in a medical vocabulary voted on (with an appeal process), incorporated into a draft, etc. Most terminology standards are created by a controlled process (that varies by terminology) with expert review rather than the open ballot process required by ANSI.

The development and ongoing maintenance of a standard is typically stewarded by a Standards Development Organization (SDO) that has overall responsibility and ownership. The SDO employs a multistep process in producing standards to ensure

Table 11.1 Examples of United States Standards Development Organizations (SDOs)

SDO	Standard
ASC X12 (Accredited Standards Committee X12 (ASC X12)	Claim Benefits and Payments
ASTM (American Society for Testing and Materials)	CCR (Continuity of Care Record
National Electrical Manufacturers Association (NEMA)	DICOM
HIBCC (Health Industry Business Communication Council)	Health Industry Bar Code standard (HIBC)
IEEE (Institute of Electrical and Electronics Engineers	Medical Information Bus (IEEE 1073)
NCPDP (National Council for Prescription Drug Programs)	SCRIPT (i.e, e-prescribing standard)

quality, integrity, and input. ANSI itself is not an SDO but is an SDO accreditation organization. An example of an ANSI accredited SDO is HL7, and example of an HL7 developed standard is FHIR (Fast health Care Interoperability Resources. Another SDO that is responsible for many healthcare standards is ASTM that developed the Continuity of Care Record (CCR), a core data set relevant administrative, demographic, and clinical information facts about a patient's healthcare. IEEE (Institute of Electrical and Electronics Engineers) is an SDO and is responsible for biomedical technology in healthcare. Please see Table 11.1 for a list of example national SDOs and their affiliated standards in the United States.

What Are Information Standards Organizations and How Do They Differ from SDO's?

Information standards organizations are organizations that solely exist to foster, promulgate and the authors would argue to coordinate standards. What creates confusion is these organizations may set rules and framework for development with a different approach than SDOs. The organization employs experts to develop the standard. A further source of confusion is that these organizations may house SDOs. For example ASTM is a standards organization but housed the HL7 SDO till it be became an independent entity. Examples of standards organizations are in Table 11.2.

The Standards and Interoperability Framework organization (S& I Framework) is a collaborative community of participants from the public and private sectors who focus on ways to facilitate the functional exchange of health information, to harmonize standards related to interoperability, and to ensure these standards meet the objectives and priorities of healthcare priorities, health outcomes, and meaningful use (www.siframework.org). The S& I Framework works with SDOs as key partners to extend existing standards, or develop new ones as necessary.

Table 11.2 United States Standards Organizations

American National Standards Institute (ANSI)
Integrating the Healthcare Enterprise (IHE)
National Institute of Standards and Technology
S&I Framework
Workgroup for Electronic Data Interchange (WEDI)

International Standards Organizations

CEN (Comité Européen de Normalisation) TC 251 – Technical Committee on Health Informatics – www.cencenelec.eu/standards/Sectors/healthcare/

The European Committee for Standardization (CEN; Comité Européen de Normalisation) is a non-profit standards organization that was founded in 1961 to develop standards and specifications for both healthcare and non-healthcare related services. The work is done through multiple subcommittees. For healthcare related settings, CEN and the European Committee for Electrotechnical Standardization (CENELEC) develop standards for safety, quality and performance requirements for medical devices on the European market as well as providing interoperability of health information systems in Europe (CEN / TC 251). The organization works closely with other global organizations for optimization of the standards.

ISO – Technical Committee 215 on Health Informatics – www.iso.org/iso/iso_technical_committee?commid=54960

The International Organization for Standardization (ISO) is an international standard-setting group that has members from throughout the world and was founded in 1947. Within healthcare the ISO has a technical committee 215 that deals with health informatics. The ISO/TC 215 seeks to facilitate coherent and consistent interchange and use of health-related data. The ISO/TC 215 was started in 1998 and includes 33 countries that actively participate with 26 "observing" countries (as of 2015). They have released more than 100 reports including on personal health device communication and point-of-care medical device communication. From within the United States, the American National Standards Institute (ANSI) is the representative for the ISO.

What Is Certification?

Certification is a process in which a "neutral body" certifies that a vendor conforms and complies with the standard [3]. Neither Standard Development Organizations nor Information Standards Organizations certify that vendors are compliant with a standard. For example, HL7 does not certify vendors as compliant and this is one reason why HL7 version 2.x had such variability in vendor implementation. The Office of the National Coordinator (ONC) through the ONC-Authorized Testing and Certification Bodies (ONC-ATCBs) coordinates testing and certification of

systems to assure the required technological capability, functionality, and security to maintain consistency across certified products. In addition, the certification process is designed to give confidence to providers and patients that certified products are secure and can work well with other systems. – *http://www.healthit.gov/policy-researchers-implementers/certification-bodies-testing-laboratories*

How Are Standards Selected and Adopted for a Particular Purpose?

Use of standards for a particular purpose may be driven by either, or both, "bottom up" and "top down" approaches. Standards may achieve significant market penetration by organic (bottom up) adoption across an industry. Many standards demonstrate the "network effect", which means they become more valuable as more people use them. If you are the only person in the world with a phone, it will not be very useful to you. But, if everyone on the planet has one, then a phone becomes a much more helpful technology. In the same way, the more data producers that can output their results in a particular standard format, the more valuable that standard becomes to data receivers. It is a virtuous cycle.

At the same time, regulations from government agencies that require use of certain standards can significantly accelerate their use. A prominent recent example of this approach is the Meaningful Use regulations in the United States, which provides incentives to hospitals and providers who use certified EHR technology. The EHR certification process requires use of designated standards for enabling technical and semantic interoperability so that health information can be efficiently and securely exchanged across care settings [4]. There is evidence that this approach is beginning to bear fruit. A recent ONC report to Congress [5] noted that more than six in ten hospitals electronically exchanged patients' health data with providers outside their organization, an increase of more than 50% since 2008. However, even with this improvement in transportability of patient data, there is substantial work to be done as a study showed only 14 % of providers sharing data with providers outside their organization from 2009 to 13 [6].

Many factors can contribute to the selection of a standard for a particular use case. Cost, fitness for the intended purpose, ease of implementation, and many other factors could all be important determinants. In the U.S., the HIT Standards Committee is a federal advisory committee charged with making recommendations to the National Coordinator for Health IT on standards, implementation specifications, and certification criteria for the electronic exchange and use of health information. As one example of a formal selection process, the HIT Standards Committee developed a set of six criteria [7] based on maturity and ease of implementation and adoption:

1. Maturity of the specification
2. Maturity of the underlying technology components
3. Market adoption

4. Ease of implementation and deployment
5. Ease of operations
6. Intellectual property

Patient Identifier Standards

Anyone who has worked in a clinical setting recognizes the importance of the medical record number (MRN). It is required to uniquely identify a patient and align the clinical documentation to the correct person. Patient identifiers such as the MRN ("for whom" the service was done) and National Provider ID (NPI) ("by whom" the service was rendered) are essential components to allow for uniquely tracking information within an electronic health record system.

Within the U.S. there are many challenges with MRNs, but the biggest is the inability to match these unique identifiers across the different health record systems that assign them. Health Information Exchanges (HIE) and hospital networks have sought to connect disparate systems with unrelated MRNs by algorithms that incorporate surname, social security number (something that may or may not be present), telephone number, and other demographic features. Within an HIE, the ability to match patient records in a single center showed a sensitivity, specificity and positive predictive value of 95.4, 98.8 and 99.9 % respectively [8]. However, a single incorrectly linked piece of clinical data could have devastating outcomes.

In 1997 Health and Human Services released a statement saying "The need for unique patient identifiers has become urgent and critical. The widespread implementation of information technology and the emergence of computer-based patient records have paved the way for its potential success". Unfortunately the desire for a *Unique Patient Identifier* (UPI) has yet to come to fruition within the United States despite many years and a rapid proliferation of electronic health care data. This means the "for whom" portion of identification standards is yet to be resolved and is unlikely to in the near future in the United States. We are therefore left to use more cumbersome methods that require intermittent manual review and ongoing optimization.

The "by whom" portion of identification was resolved in the United States by the Centers for Medicare and Medicaid Services with the *National Provider Identifier* (NPI). This 10-digit numeric identifier replaced the previous unique provider identification number (UPIN) in 2007 and is required for medical billing to both government and private insurance. The NPI's ten digits allow for the first nine to be uniquely identifying with the 10th digit a check digit based on the Luhn algorithm for validation. The NPI can be utilized as the identification within electronic health record systems, for prescriptions, and for many other utilizations. This provider identification does not replace either the Drug Enforcement Administration (DEA) or the local state license number.

Recently there has been an increased push towards having specific medical devices uniquely identified through a *Unique Device Identification (UDI)* (www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/UniqueDeviceIdentification).

The UDI would allow a "by whom" mechanism for non-human services that are provided (e.g. pacemaker). This trend has been pushed to allow more accurate reporting of adverse events, allow for specific identification of a device by providers, and to prevent counterfeit devices from being on the market. The Food and Drug Administration (FDA), who mandates utilization of the UDI, will phase in the requirement starting in September 2014 and ending in September 2020.

The key to identification standards is for wide-spread utilization. Government mandates for both NPI and for UDI have allowed for widespread (NPI) and future (UDI) utilization across health systems. The hope is that in the near future a UPI can be adopted to complete the "for whom" and "by whom" loop. However, there is still a need to add additional identification standards such as "where" or "in what setting" the care was delivered [9].

Terminology Standards

Controlled terminologies or controlled vocabularies (a way to organize knowledge for subsequent retrieval) are essential for capturing, storing, and processing electronic patient data. The prose with which clinicians describe health is inherently nuanced and ambiguous. While humans revel in the rich expressiveness of language, computers falter. Computers need controlled terminologies because they enable reusability of the data within and among system. As we discuss many different healthcare coding systems, there are many labels for these things that are often used interchangeably. You might be wondering what the difference is between a terminology, classification, and nomenclature. For our purposes we will use definitions based on ISO Standard 1087 (Terminology work – Vocabulary) and Giannangelo [10]:

- *term: designation of a defined concept in a special language by a linguistic expression*
- *nomenclature: system of terms elaborated according to established naming rules*
- *vocabulary: a collection of words or phrases with their meaning (i.e. a dictionary)*
- *terminology: a set of terms representing the system of concepts of a particular subject field*
- *classification: a system that organizes like or related entities*
- *semantics: the insertion of meaning via relationships (see SNOMED CT below)*

Another somewhat subtle distinction is the term *ontology*. In the context of information science (ontology also refers to the philosophical study of the nature of being), an ontology is a formal representation of some pre-existing domain of reality in a way that allows it to support automatic information processing [11–13]. In the context of healthcare terminologies, an ontology could be thought of as a terminology that contains some formal representation of definitional information. Ontologies serve to represent a truth (e.g. body temperature) and do not reflect the presence or absence of this knowledge.

In this chapter, we cannot provide an all-inclusive list of healthcare terminologies. There are, in fact, whole textbooks devoted to the subject [10]. Our intent is to highlight those that are in most widespread use and of greatest importance the field of clinical informatics.

Healthcare "Billing" Terminologies and Classifications

International Classification of Diseases (ICD) – www.who.int/classifications/icd/en/

The International Statistical Classification of Diseases and Related Health Problems, more commonly known as ICD is an international standard that is published by the World Health Organization (WHO) and was put in place for morbidity reporting. Originally known as the International List of Causes of Death to International Statistical Classification of Diseases the first usable version was ICD-6 that was published in 1949. This has undergone revisions through the years with the most notable ones with ICD-9 in 1978 and ICD-10 in 1990.

The ninth revision has been used extensively as a billing mechanism in the United States under the International Classification of Diseases, Clinical Modification (ICD-9-CM) and is the requirement for Medicare and Medicaid claims along with the majority of private industry. ICD-9-CM contains both diagnostic and procedure codes within both inpatient and outpatient settings. Volumes 1 (tabular listing) and 2 (index) contain diagnosis codes while Volume 3 contains only procedure codes.

The National Center for Health Statistics (NCHS) and the Centers for Medicare and Medicaid Services (CMS) are responsible for all changes and modifications to ICD-9-CM and the standard is updated every October 1st. Table 11.3 shows the structure of ICD-9-CM.

While the ninth revision has been dominant, the 10th revision for the ICD was endorsed by the World Health Assembly in 1990 and began being used by WHO Member States in 1994. The revision increased the number of codes to more than 155,000 up from the 17,000 in ICD-9-CM.

ICD-10 has been adopted throughout the world, however ICD-10-CM has been a challenge in the United States. In August of 2008 the United States Department of Health and Human Services (HHS) proposed that ICD-10-CM would replace ICD-9-CM on October 1, 2013, however HHS delayed this to October 1, 2014, and then again until October 1, 2015.

The majority of the dissent to the implementation of the ICD-10-CM system in the United States has focused on its use within electronic health record systems. In particular, detractors question the benefits and see limitations in increasing the number of codes from 17,000 to 155,000. While many would argue that having more granularity will allow for better representation of the diseases and services rendered, this level of granularity can have humorous (V97.33 "Sucked into jet engine") and likely unnecessary (Y92.146 "Swimming-pool of prison as the place of occurrence of the external cause") consequences.

Table 11.3 ICD-9-CM index of diseases

Diseases covered	ICD-9-CM codes
Infectious and parasitic diseases	001–139
Neoplasms	140–239
Endocrine, nutritional and metabolic diseases, and immunity disorders	240–279
Diseases of the blood and blood-forming organs	280–289
Mental Disorders	290–319
Diseases of the nervous system and sense organs	320–389
Diseases of the circulatory system	390–459
Diseases of the respiratory system	460–519
Diseases of the digestive system	520–579
Diseases of the genitourinary system	580–629
Complications of pregnancy, childbirth, and the puerperium	630–679
Diseases of the skin and subcutaneous tissue	680–709
Diseases of the musculoskeletal system and connective tissue	710–739
Congenital anomalies	740–759
Certain conditions originating in the perinatal period	760–779
Symptoms, signs, and ill-defined conditions	780–799
Injury and poisoning	800–999
Supplementary classification of factors influencing health status and contact with health services	V01–V89
Supplementary classification of external causes of injury and poisoning	E800–E999

The resolution of these additional codes has also been a challenge for electronic health records that will have to correctly "map" the codes to what is being performed clinically.

Current Procedural Terminology (CPT®) – www.ama-assn.org/go/cpt

The Current Procedural Terminology (CPT) is a standardized terminology that is owned and maintained by the American Medical Association (AMA) and first published in 1966. Its original focus was on codes for surgical procedures. While CPT now covers a broader domain, it surgical procedures codes have been used widely.

In 1996, the Health Insurance Portability and Accountability Act (HIPAA) enacted and the Department of Health and Human Services (HHS) published the "Final Rule" that selected CPT for reporting physician services for payments for CMS. CPT is broken into three categories:

Category I: Numeric codes for procedures that are within the scope of medical practice in the United States (e.g. 45385 "Colonoscopy with polypectomy"). Within this category there is an assignment of relative value units (RVUs) by the Relative Value Scale (RVS) Update Committee (RUC).

Category II: Alphanumeric codes for tracking performance measurement (e.g. 3725 F "Screening for depression"). These codes all contain the "F" designation. It is anticipated that these codes will be important for Pay-for-Performance (P4P) measures.

Category III: Temporary codes for new or emerging procedures and are removed after 5 years from the time of publication (e.g. 0346 T "Ultrasound, elastography"). These codes all contain the "T" designation.

The AMA retains a panel of 11 physicians that are nominated by multiple medical societies, health insurance plans, and by CMS. This CPT Editorial Panel meets three times a year to discuss new and emerging technologies and any difficulties with the CPT® codes.

Healthcare Common Procedure Coding System (HCPCS) – www.cms.gov/medhcpcsgeninfo/

The Healthcare Common Procedure Coding System (HCPCS), pronounced "hick picks", is a billing coding system based on the AMA CPT that was started in 1978 to provide a descriptive standard for billing services in health care.

Similar to CPT codes, the HCPCS was required for reporting services to CMS in 1996 due to HIPAA. This coding is necessary for billing Medicare, Medicaid, and other commercial health insurance programs. HCPCS is broken into two categories with a now retired category III:

Category I: Numeric codes from AMA CPT (See section on CPT).

Category II: Alphanumeric codes for primarily non-physician services and not represented in Category I (e.g. B4034 "Enteral feeding supply kit; syringe fed, per day").

Category II codes are broken down into: (1) Permanent National Codes, maintained by the CMS HCPCS Workgroup, (2) Dental Codes, maintained by the Current Dental Terminology (CDT) from the American Dental Association (ADA), (3) Miscellaneous Codes that represent services that are not currently available in a coded manner (e.g. a service that is provide while a new code is under the HCPCS review process), and (4) Temporary National Codes that allows for providing a code prior to the next January 1 annual update to HCPCS.

Category III: Local codes (now discontinued since the end of 2003).

Diagnosis-Related Group (DRG) – www.cms.gov/medicare/medicare-fee-for-service-payment/acuteinpatientpps/

The Diagnosis-related Group (DRG), also known as the MS-DRG, is a system that describes a bundle of services that a hospital might provide. The system was maintained by the U.S. Congress in 1982 for creation of a prospective payment system (PPS) to control costs. In this setting Medicare paid a flat rate per case for inpatient hospital care to reward hospitals for efficiency. The groupings allowed for a likely expenditure based on the underlying disease and co-morbidities that a hospital was likely to accrue.

Initially the system for "bundled" services was trialed in 1980 in New Jersey and became much more wide-spread in the early 1980s for Medicare services. These services can range from "Chest pain" (313) to "Liver transplant" (006) and consist of 470 unique codes with number 999 (previously 470) as "ungroupable". The DRG codes are maintained by the Department of Health and Human Services (HHS).

Clinical Standard Terminologies

SNOMED Clinical Terms (SNOMED CT®) – www.nlm.nih.gov/snomed/

SNOMED Clinical Terms (SNOMED CT) is a multi-lingual clinical terminology that is used in many countries. The terminology was originally created by the College of American Pathologists (CAP) and was the combination merger from SNOMED Reference Terminology (SNOMED RT) and the United Kingdom National Health Services Clinical Terms (Read Codes).

SNOMED CT is owned, maintained, and distributed by the International Health Terminology Standards Development Organization (IHTSDO) in Denmark. The IHTSDO itself is owned and governed by more than twenty-seven member organizations. In the United States, the National Library of Medicine (NLM) is the member organization of the IHTSDO and distributes SNOMED CT under the IHTSDO's uniform international license at no cost for use within the U.S. SNOMED CT is one of the suite of standard terminologies designated by the United States government for electronic health information exchange.

The NLM makes SNOMED CT available both through the UMLS and also in the native SNOMED CT file formats produced by the IHTSDO. The NLM also maintains a SNOMED CT browsing service at https://uts.nlm.nih.gov/snomedctBrowser.html that requires a free account for access.

SNOMED CT is a concept-oriented terminology that allows for machine readability numeric (e.g. 104817019 "left cusp of aortic valve"). The structure of the terminology is from larger concept (e.g. "body structure") towards more specific concepts (e.g. "anatomical or acquired body structure" ->"anatomical structure" ->"body organ structure" ->"organ part" ->"cardiovascular organ part" ->"heart part" ->"cardiac internal structure" ->"cardiac valve structure" ->"aortic valve structure" ->"structure of cusp of aortic valve" ->"structure of left cusp of aortic valve"). The system allows for relationships both to and from the specific concept (e.g. "is a" and "part of") as well as synonyms (e.g. "left coronary cusp").

SNOMED CT provides both pre-coordination and post-coordinated expressions. A pre-coordination expression refers to a single concept defining one clinical idea (e.g. "burn of skin"). Post-coordination expressions describes two or more terms that can be combined by an expression to represent a new meaning (e.g. "burn of skin" by "hot water" on "index finger") [14].

SNOMED CT contrasts from ICD-9/10-CM in that it is designed as a relationship of concepts that goes above and beyond the simple listing that is present in

ICD. This has both advantages and disadvantages. While the relationship status of SNOMED CT allows for complex associations including synonyms, these do not easily aggregate towards a specific concept that a clinician would make use of for either billing or for reporting a disease state.

As described by Cimino in his *Desiderata for controlled medical vocabularies in the twenty-first century* the optimal vocabulary must have "vocabulary content, concept orientation, concept permanence, non-semantic concept identifiers, polyhierarchy, formal definitions, rejection of "not elsewhere classified" terms, multiple granularities, multiple consistent views, context representation, graceful evolution, and recognized redundancy"[11, 15]. This listing of attributes is considered to be the "holy grail" of medical terminologies and contains many ideas that themselves could constitute a chapter in a book. For illustration purposes, we highlight that according to the Desiderata, an optimal vocabulary possesses broad and comprehensive concepts that have only one meaning with terms neither changed nor deleted, and uses unique identifiers that have no semantic (logical) meaning. Astute readers will note that ICD violates many of the Desiderata criteria.

Logical Observation Identifiers Names and Codes (LOINC®) – loinc.Org

Logical Observation Identifiers Names and Codes (LOINC) is a universal code system for identifying laboratory tests and clinical observations in electronic messaging. LOINC was developed by the Regenstrief Institute and the LOINC Committee in 1994 to support exchange and aggregation for care delivery, outcomes management, and research.

The current release of LOINC, version 2.50 (December 2014), contains more than 73,000 concepts covering the full scope of laboratory testing (chemistry, microbiology, etc.) and a broad range of clinical measurements (e.g. vital signs, EKG, patient reported outcomes, etc.). LOINC has a sophisticated data model for representing answer lists, panels of individual observations, and other details like help text, units of measure, and more. Based on formal naming conventions, LOINC also carries names for document titles (discharge summary, radiation oncology consult note, etc.), radiology reports and section headings (social history, objective, etc.).

LOINC has become widely adopted as the standard for laboratory and clinical observations in the USA and internationally. Today, there are more than 37,000 registered users from 166 countries and it has been translated into 18 variants of 12 languages. Many countries have adopted LOINC as a national standard. Within the USA, the Meaningful Use program requires LOINC in messages reporting laboratory test results, exchanging medical summaries, and sending data to cancer registries and public health agencies.

Each LOINC term is assigned a unique identifier (the LOINC code) and a fully specified name containing six main axes:

1. component (e.g. what is measured, evaluated or observed),
2. kind of property (e.g. mass, substance, catalytic activity),
3. time aspect (e.g. 24 h collection),

4. system type (e.g. context or specimen type within which the observation was made),
5. type of scale (e.g. ordinal, nominal, narrative)
6. type of method (e.g. procedure used to make the measurement or observation).

The combination of axis values produce names that are detailed enough to distinguish among similar observations. Of the six axes, only the method is optional and used only when necessary to distinguish among clinical important differences.

The Regenstrief Institute continues to develop and maintain LOINC. New concepts are added to LOINC based on submissions from end users, with new releases being published twice yearly. In addition to distributing the terminology, Regenstrief makes available at no cost a variety of supporting tools and resources, including the Regenstrief LOINC® Mapping Assistant (RELMA®) and online search application at http://search.loinc.org.

Drug Standard Terminologies

RxNorm – http://www.nlm.nih.gov/research/umls/rxnorm/

RxNorm is a standardized terminology set that provides normalized names for clinical drugs and links to synonyms within First Databank, Micromedex, MediSpan, Gold Standard Drug Database, NDF-RT from the Veterans Health Administration, and Multum. RxNorm is maintained by the National Library of Medicine (NLM) and gives both generic and branded names of prescription and over-the-counter drugs in the United States.

Since RxNorm is an aggregation of multiple sources of drug information each concept is sourced for a common meaning (e.g. 198013 "Naproxen Tab 250 MG=Naproxen 250 mg In 1 TABLET ORAL TABLET). This process allows for (1) ingredient (e.g. naproxen, (2) strength (e.g. 250 mg), and (3) dose form (e.g. "tab").

RxNorm is available through the NLM for free in the United States with a UMLS® Terminology Services (UTS) account. RxNorm is released as an update the first Monday of each month.

National Drug File Reference Terminology (NDF-RT) - http://www.nlm.nih.gov/research/umls/sourcereleasedocs/current/NDFRT/

The National Drug File Reference Terminology (NDF-RT) is a standardized terminology system for modeling drug characteristics including ingredients, chemical structure, dose form, physiologic effect, mechanism of action, pharmakinetics, and related diseases. The NDF-RT was created and is maintained by the United States Department of Veterans Affairs, Veterans Health Administration (VHA) and is part of RxNorm.

National Drug Code (NDC) – http://www.fda.gov/Drugs/InformationOnDrugs/ucm142438.htm

The Drug Listing Act of 1972 established through the Food and Drug Administration (FDA) that all drugs that were manufactured must be registered. This created a National Drug Code (NDC) that utilizes a three-segment number that serves as a universal identifier of the product.

The NDC number (e.g. 21695-255-90 "Lipitor 40 mg Tablet Film Coated") uniquely identifies the product NDC (e.g. 21695), the proprietary name (e.g. Lipitor), the non-proprietary name (e.g. atorvastatin calcium), the route (e.g. oral), the substance name (e.g. atorvastatin calcium), the package description (e.g. 90 tablet film coated in one bottle), and the labeler name (e.g. Rebel Distributors Corp). This highly detailed system allows for direct ability to identify a specific marketed product with the substance as well as the method for distribution.

The NDC is maintained by the FDA and is listed in the NDC Directory. This directly is updated daily to maintain an up-to-date listing of manufactured drugs. The NDC is a challenging system to utilize in comparison to RxNorm as it does not allow easy grouping of similar products. However, the detailed codings allow for excellent pharmacy integration and they can be mapped through the UMLS® to matching codes.

Other Healthcare Related Terminologies

There are many standardized nursing terminologies in use today, such as the Clinical Care Classification (CCC), Omaha System, Nursing Intervention Classification (NIC), Nursing Outcomes Classification, International Classification for Nursing Practice (ICNP) and the Perioperative Nursing Data Set (PNDS). Programs such as Meaningful Use have promoted a parsimonious set of core standards for EHRs, including standard terminologies such as SNOMED CT, LOINC, and RxNorm. There are active programs of work to enhance the nursing-related content in the dominant global healthcare terminologies (e.g. SNOMED CT and LOINC), including harmonization and mapping of elements from several nursing-specific terminologies. Given the trajectory towards using fewer distinct standard terminologies we will not describe each of the nursing specific terminologies in detail here. Table 11.4 shows the standard nursing terminology and a summary of what is contained.

Current Dental Terminology (CDT) and SNODENT – http://www.ada.org/en/publications/cdt and http://www.ada.org/en/member-center/member-benefits/practice-resources/dental-informatics/snodent

The Current Dental Terminology (CDT) is the HIPAA standard (since 2000) for dental procedures and for electronic dental claims. The CDT is maintained by the American Dental Association (ADA) and has previously been included in HCPCS Category II, however it is now maintained exclusively by the ADA and updated annually. An example code is D3347, "retreatment of previous root canal therapy – bicuspid". CDT is distributed within the UMLS®. SNODENT is an internationally recognized subset of SNOMED CT that is curated by the ADA. SNODENT supplements the CDT. SNODENT focuses on diagnostic and patient features, while the CDT focuses on procedures and treatments.

Unified Medical Language System (UMLS) – https://uts.nlm.nih.gov/

Due to the plethora of terminologies that have developed from multiple mechanisms there was a pressing need for a standard to link the different standards.

Table 11.4 Listing of standardized nursing terminologies and the content they cover

Nursing standardized terminology	Content/scope	Reference
Clinical Care Classification (CCC)	Provides a standardized coded framework for assessing, documenting, and evaluating nursing care holistically, over time, across settings, population groups, and geographic locations.	www.sabacare.com/
Omaha System	Consists of three relational components: (1) Problem Classification Scheme (client assessment), (2) Intervention Scheme (care plans and services), (3) Problem Rating Scale for Outcomes (client change/evaluation)	www.omahasystem.org/
Nursing Intervention Classification (NIC)	Comprehensive, research-based, standardized classification of interventions that nurses perform including clinical documentation, communication of care across settings, integration of data across systems and settings, effectiveness research, productivity measurement, competency evaluation, reimbursement, and curricular design.	www.nursing.uiowa.edu/cncce/nursing-interventions-classification-overview
Nursing Outcomes Classification (NOC)	Comprehensive, standardized classification of patient/client outcomes developed to evaluate the effects of interventions provided by nurses or other health care professionals	www.nursing.uiowa.edu/cncce/nursing-outcomes-classification-overview
International Classification for Nursing Practice (ICNP®)	A classification of nursing phenomena, nursing actions, and nursing outcomes that describes nursing practice.	www.icn.ch/what-we-do/international-classification-for-nursing-practice-icnpr/
Perioperative Nursing Data Set (PNDS)	Perioperative-specific, standardized nursing vocabulary recognized by the American Nursing Association (ANA) and is mapped to SNOMED CT®.	www.aorn.org/PNDS/

The Unified Medical Language System (UMLS) was established in 1986 to provide this service. This "Rosetta Stone" of the healthcare terminologies has integrated 58 separate sources (2014AB Release) that include the main terminologies for clinical data exchange and system integration.

The UMLS is maintained by the United States' National Institute of Health (NIH) National Library of Medicine (NLM). The UMLS provides a "Metathesaurus" that allows for terms and codes from multiple standards (e.g. ICD-10-CM, LOINC, CPT) with a unique identifier. There is also a "Semantic Network" that shows the broad categories (semantic types) and their relationships (semantic relations). Lastly it provides a lexical tools to provide natural language processing.

Data Exchange Standards

Health Level Seven International (HL7) – www.hl7.org

Health Level Seven International (HL7) was founded in 1987 and is an American National Standards Institute (ANSI) accredited standards developing organization for providing a framework for the exchange, integration, sharing, and retrieval of electronic health data. Of all the standards to know as a clinical informatician, the HL7 standards are probably the most critical because they provide the means for transmitting data across healthcare information systems (both local and external).

The "seven" of HL7 comes from the seventh level (application level) of the International Organization for Standardization (ISO) seven-layers of communications model for Open System Interconnection (OSI). HL7 develops its standards as a collaborative effort of many volunteers. HL7 members include individuals and organizations of many types, such as commercial entities, governmental, and non-government agencies.

HL7 has published several primary versions of its standards, but currently there are two widely used and both have some controversy. HL7 version 2 was released in 1987 with the most recent update to 2.7 in 2011. This version is used in 35 countries and 95 % of US healthcare organizations. HL7 version 2.x is organized into several different message types (e.g. Admission Discharge Transfer or Observation Result). Each message type has a set of segments that contain fields delimited by the "|" character. For example an observation result (ORU) message would contain segments such as the MSH (Message Header), the PID (Patient Identification), the OBR (Observation Request), and the Observation (OBX). And inside the Observation segment, there are separate fields to identify the data type, observation identifier, observation result value, units of measure, etc. Figure 11.1 shows an example HL7 version 2.x message. While it is important for informaticians to be aware of this general structure a detailed understanding of the specifics is not required (e.g. you can always Google it).

HL7 version 3 was begun in 1995 and initially published in 2005. The version 3 of HL7 was a large departure from the previous version, and thus has been met with significant resistance. HL7 version 3 utilizes a human readable Reference Information Model (RIM) in the Extensible Markup Language (XML). Figure 11.2 shows a HL7

```
PID|||12322^^^ Assigning authority ^MR ^Savage^Robert^^^^^L^|
```

Fig. 11.1 Healthcare Level Seven (HL7) Version 2 example.(Reprinted with permission from Health Level Seven® International)

Fig. 11.2 Healthcare Level Seven (HL7) Version 3 example.(Reprinted with permission from Health Level Seven® International)

```
- <ORU_R01.ORDER_OBSERVATION>
  - <OBR>
      <OBR.1>1</OBR.1>
    - <OBR.2>
        <EI.1>845439</EI.1>
        <EI.2>GHH OE</EI.2>
      </OBR.2>
    - <OBR.3>
        <EI.1>1045813</EI.1>
        <EI.2>GHH LAB</EI.2>
      </OBR.3>
    - <OBR.4>
        <CE.1>1554-5</CE.1>
        <CE.2>GLUCOSE</CE.2>
        <CE.3>LN</CE.3>
      </OBR.4>
    - <OBR.7>
        <TS.1>200202150730</TS.1>
      </OBR.7>
    - <OBR.16>
        <XCN.1>555-55-5555</XCN.1>
      - <XCN.2>
          <FN.1>PRIMARY</FN.1>
        </XCN.2>
        <XCN.3>PATRICIA P</XCN.3>
        <XCN.7>MD</XCN.7>
      - <XCN.9>
          <HD.1>LEVEL SEVEN HEALTHCARE, INC.</HD.1>
        </XCN.9>
      </OBR.16>
      <OBR.25>F</OBR.25>
    - <OBR.32>
      - <NDL.1>
          <CNN.1>444-44-4444</CNN.1>
        - <CNN.2>
            <FN.1>HIPPOCRATES</FN.1>
          </CNN.2>
          <CNN.3>HOWARD H</CNN.3>
          <CNN.7>MD</CNN.7>
        </NDL.1>
      </OBR.32>
    </OBR>
  - <ORU_R01.OBSERVATION>
    - <OBX>
        <OBX.1>1</OBX.1>
        <OBX.2>SN</OBX.2>
      - <OBX.3>
          <CE.1>1554-5</CE.1>
          <CE.2>GLUCOSE POST 12H CFST</CE.2>
          <CE.3>LN</CE.3>
        </OBX.3>
      - <OBX.5>
          <SN.2>182</SN.2>
        </OBX.5>
      - <OBX.6>
          <CE.1>mg/dl</CE.1>
        </OBX.6>
        <OBX.7>70-105</OBX.7>
        <OBX.8>H</OBX.8>
        <OBX.11>F</OBX.11>
      </OBX>
    </ORU_R01.OBSERVATION>
</ORU_R01.ORDER_OBSERVATION>
```

version 3 message. Although available for several years, version 3 of HL7 has not been widely adopted. This may be partially due to the reasoning for each version's creation. Version 2 was created by a small group of interface specialists and software vendors as an ad hoc standard, while version 3 was created by informaticians with a focus on modelling health care information and data for a wide range of use cases. Regardless of the controversy (or the reluctance to accept one or the other standard),

```
<gender>
  <coding>
    <system value="http://hl7.org/fhir/v3/AdministrativeGender"/>
    <code value="M"/>
    <display value="Male"/>
  </coding>
</gender>
<birthDate value="1974-12-25"/>
<deceasedBoolean value="false"/>

<address>
  <use value="home"/>
  <line value="534 Erewhon St"/>
  <city value="PleasantVille"/>
  <state value="Vic"/>
  <zip value="3999"/>
</address>

<contact>
  <relationship>
    <coding>
      <system value="http://hl7.org/fhir/patient-contact-relationship"/>
      <code value="partner"/>
    </coding>
  </relationship>

  <name>
    <family value="du">
      <!--  the "du" part is a family name prefix (VV in iso 21090)  -->
      <extension url="http://hl7.org/fhir/Profile/iso-21090#qualifier">
        <valueCode value="VV"/>
      </extension>
    </family>
    <family value="Marché"/>
    <given value="Bénédicte"/>
  </name>

  <telecom>
    <system value="phone"/>
    <value value="+33 (237) 998327"/>
  </telecom>
</contact>
```

Fig. 11.3 FHIR® version example.(Reprinted with permission from Health Level Seven® International)

they are both in active use and as an informatician you must be aware of both. Probably the most widely recognized component of the HL7 version 3 suite of standards is the HL7 Clinical Document Architecture (CDA). The CDA is a standard for specifying the structure and semantics of clinical documents. As part of the HL7 version 3 suite, the CDA derives its semantic content from the shared HL7 Reference Information Model and is implemented in Extensible Markup Language.

Fast Healthcare Interoperability Resources FHIR® (pronounced "fire") – (www.hl7.org/fhir/) is an emerging and likely impactful HL7 standard that looks to take to good portions of versions 2 and 3 and merge them with a focus on implementation. At the time of this writing FHIR has been published as a Draft Standard for Trial Use (DSTU), but there is also significant momentum for both FHIR development and implementation. Many major EHR system vendors have already started opening up new pathways to integration with their products using FHIR and are actively participating in FHIR's continued development. Figure 11.3 shows an example of FHIR.

Digital Imaging and Communications in Medicine (DICOM) – dicom.nema.org

The Digital Imaging and Communications in Medicine (DICOM) is an international medical imaging standard for handling, storing, printing, and transmitting across a specified network communications protocol with a specific file format definition.

DICOM was first published in 1993 and is implemented on nearly every radiology, cardiology, and radiotherapy device (e.g. CT, MRI, ultrasound). DICOM is maintained by the National Electrical Manufacturers Association (NEMA) and holds copyright to the standard. The DICOM image format (DICOM data object) includes name and medical record number so that these can never be separated from the image. DICOM

standards also leverage standard terminologies such as LOINC and SNOMED CT. Radiology is the primary domain of DICOM development and usage, but the standard has applicability in other domains such as Obstetrics and Gynecology and Cardiology.

Emerging Trends in Healthcare Data Standards

As this chapter goes to press, we note several trends that we anticipate will continue to drive evolution in healthcare data standards. The vision of creating a "learning health system" as articulated by the Institute of Medicine (IOM) [16] continues to drive many national initiatives. A key characteristic of such a learning health system is the ability to "capture the care experience on digital platforms for real-time generation and application of knowledge for care improvement". Further, the White House's recently announced focus "precision medicine" initiative [17] continues a movement towards disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person. The only way to accomplish these goals is with a broad scope of interoperable health IT products and services that use common standards for the content and structure of health data.

Against this backdrop, we anticipate that the development and adoption of data standards will spread in both breadth and depth. Data from providers and care settings outside of the current emphasis on hospitals and primary care need to communicate on the same platform. This extends to clinical research settings, healthcare devices, and data directly from patients through patient-reported outcomes measures, wearable devices, and more.

The current digital infrastructure operates largely on a messaging and document exchange paradigm. But with the rapidly growing number of different data sources and applications that operate on them coming online in the "internet of health things", we anticipate a rapid growth towards the use of open APIs such as HL7's FHIR standard. Likewise, as the scope of available electronic health data extends deeper into the richness of genetic, behavioral, social, and other environmental factors that influence health, terminology standards such as LOINC and SNOMED CT will need to expand their content coverage.

The deep complexity of health knowledge and systems means that these transitions will not be easy. For example, a recent paper describing LOINC's approach to representing genetic testing results [18] noted that although specifications for reporting fully structured genetic variation and cytogenetic results have existed for several years, most genetic test results are sent today as narrative text.

Conclusion

Standards are critical for developers of systems, interoperability, and the exchange of information across health information networks. Planned and future development and enhancements to standards can both assure us of comprehensive information on our patients as well as providing quality and efficient care.

References

1. McDonald CJ, Schadow G, Suico J, Overhage JM. Data standards in health care. Ann Emerg Med. 2001;38(3):303–11. Epub 2001/08/29.
2. Hammond WE. The making and adoption of health data standards. Health Affairs (Millerwood). 2005;24(5):1205–13.
3. Shortliffe EH, Cimino JJ. Biomedical informatics: computer applications in health care and biomedicine, vol. xxvi. 4th ed. London: Springer; 2013. p. 965.
4. Dolin RH, Alschuler L. Approaching semantic interoperability in health level seven. J Am Med Inform Assoc. 2011;18(1):99–103. Epub 2010/11/26.
5. Office of the national coordinator for Health Information Technology. Report to congress: update on the adoption of health information technology and related efforts to facilitate the electronic use and exchange of health information. Washington: Department of Health and Human Services; 2013.
6. Furukawa MF, King J, Patel V, Hsiao CJ, Adler-Milstein J, Jha AK. Despite substantial progress in EHR adoption, health information exchange and patient engagement remain low in office settings. Health Affairs (Millwood). 2014;33(9):1672–9. Epub 2014/08/12.
7. Baker D. HIT standards committee, NwHIN power team: final recommendations. [Powerpoint presentation] online: office of the national coordinator for health information technology; 2011 [cited 28 Apr 2015]; Available from: http://www.healthit.gov/sites/default/files/archive/HIT%20Standards%20Committee/2011/2011-09-28/3%20-%20Baker_NwHIN-PT_HITSC%28final%29.ppt.
8. Zhu VJ, Overhage MJ, Egg J, Downs SM, Grannis SJ. An empiric modification to the probabilistic record linkage algorithm using frequency-based weight scaling. J Am Med Inform Assoc. 2009;16(5):738–45.
9. Dixon BE, Colvard C, Tierney WM. Identifying health facilities outside the enterprise: challenges and strategies for supporting health reform and meaningful use. Inform Health Soc Care. 2014;1–15.
10. Giannangelo K. Healthcare code sets, clinical terminologies, and classification systems, vol. xiv. 2nd ed. Chicago: American Health Information Management Association; 2010. p. 336.
11. Cimino JJ. In defense of the desiderata. J Biomed Inform. 2006;39(3):299–306.
12. Ceusters W, editor. Terminology and ontology in semantic interoperability of electronic health records. Joint EU-WHO workshop. Brussels. 2005.
13. Cimino JJ, Zhu X. The practical impact of ontologies on biomedical informatics. Yearb Med Inform. 2006;124–35.
14. Benson T. Principles of health interoperability HL7 and SNOMED, vol. xxv. 2nd ed. Dordrecht: Springer; 2012. p. 316.
15. Cimino JJ. Desiderata for controlled medical vocabularies in the twenty-first century. Methods Inf Med. 1998;37(4–5):394–403.
16. Institute of Medicine. Best care at lower cost: the path to continuously learning health care in america. Washington, DC: The National Academies Press; 2013.
17. National institutes of health. precision medicine initiative. [Internet] Available at http://www.nih.gov/precisionmedicine/. Accessed 17 Apr 2015.
18. Deckard J, McDonald CJ, Vreeman DJ. Supporting interoperability of genetic data with LOINC. J Am Med Inform Assoc. 2015 pii: ocu012. doi:10.1093/jamia/ocu012. [Epub ahead of print] PubMed.