Standards in Biomedical Informatics

7

W. Edward Hammond, Charles Jaffe,
James J. Cimino, and Stanley M. Huff

After reading this chapter, you should know the answers to these questions:

- Why are standards important in biomedical informatics?
- What data standards are necessary to be able to exchange data seamlessly among systems?
- What organizations are active in standards development?
- What aspects of biomedical information management are supported today by standards?
- What is the process for creating consensus standards?
- What factors and organizations influence the creation of standards?

W.E. Hammond (✉)
Duke Center for Health Informatics, Duke University Medical Center, Room 12053, Hock Plaza, 2424 Erwin Road, Durham, NC 27705, USA
e-mail: william.hammond@duke.edu

C. Jaffe, PhD
Health Level Seven International, 122 15th Street, #2250, Del Mar, CA 92014, USA

J.J. Cimino, MD
Laboratory for Informatics Development, NIH Clinical Center, 10 Center Drive, Room 6-2551, Bethesda, MD 20892, USA
e-mail: ciminoj@mail.nih.gov

S.M. Huff, MD
Medical Informatics, Intermountain Healthcare, South Office Building, Suite 600, 5171 Cottonwood Street, Murray, UT 84107, USA
e-mail: stan.huff@imail.org, coshuff@ihc.com

7.1 The Idea of Standards

Ever since Eli Whitney developed interchangeable parts for rifle assembly, standards have been created and used to make things or processes work more easily and economically—or, sometimes, to work at all. A standard can be defined in many physical forms, but essentially it comprises a set of rules and definitions that specify how to carry out a process or produce a product. Sometimes, a standard is useful because it provides a way to solve a problem that other people can use without having to start from scratch. Generally, though, a standard is useful because it permits two or more disassociated people to work in some cooperative way. Every time you screw in a light bulb or play a music CD, you are taking advantage of a standard. Some standards make things work more easily. Some standards evolve over time,[1] others are developed deliberately.

The first computers were built without standards, but hardware and software standards quickly became a necessity. Although computers work with values such as 1 or 0, and with "words" such as 10101100, humans need a more readable language (see Chap. 5). Thus, standard

[1] The current standard for railroad-track gauge originated with Roman chariot builders, who set the axle length based on the width of two horses. This axle length became a standard as road ruts developed, requiring that the wheels of chariots—and all subsequent carriages—be the right distance apart to drive in the ruts. When carriage makers were called on to develop railway rolling stock, they continued to use the same axle standard.

E.H. Shortliffe, J.J. Cimino (eds.), *Biomedical Informatics*,
DOI 10.1007/978-1-4471-4474-8_7,

character sets, such as ASCII and EBCDIC, were developed. The first standard computer language, COBOL, was written originally to simplify program development but was soon adopted as a way to allow sharing of code and development of software components that could be integrated. As a result, COBOL was given official standard status by the American National Standards Institute (ANSI).[2] In like manner, hardware components depend on standards for exchanging information to make them as interchangeable as were Whitney's gun barrels.

A 1987 technical report from the International Standards Organization (ISO) states that "any meaningful exchange of utterances depends upon the prior existence of an agreed upon set of semantic and syntactic rules" (International Standards Organization 1987). In biomedical informatics, where the emphasis is on collection, manipulation, and transmission of information, standards are greatly needed but have only recently begun to be available. At present, the standards scene is evolving so rapidly that any description is inevitably outdated within a few months. This chapter emphasizes the need for standards in general, standards development processes, current active areas of standards development, and key participating organizations that are making progress in the development of usable standards.

7.2 The Need for Health Informatics Standards

Standards are generally required when excessive diversity creates inefficiencies or impedes effectiveness. The health care environment has traditionally consisted of a set of loosely connected, organizationally independent units. Patients receive care across primary, secondary, and tertiary care settings, with little bidirectional communication and coordination among the services. Patients are cared for by one or more primary physicians, as well as by specialists. There is little coordination and sharing of data between inpatient care and outpatient care. Both the system and patients, by choice, create this diversity in care. Within the inpatient setting, the clinical environment is divided into clinical specialties that frequently treat the patient without regard to what other specialties have done. Ancillary departments function as detached units, performing their tasks as separate service units, reporting results without follow-up about how those results are used or whether they are even seen by the ordering physician. Reimbursement requires patient information that is often derived through a totally separate process, based on the fragmented data collected in the patient's medical record and abstracted specifically for billing purposes. The resulting set of diagnosis and procedure codes often correlates poorly with the patient's original information (Jollis et al. 1993).

7.2.1 Early Standards to Support the Use of IT in Health Care

Early interest in the development of standards was driven by the need to exchange data between clinical laboratories and clinical systems, and then between independent units within a hospital. Therefore, the first standards were for data exchange and were referred to as messaging standards. Early systems were developed within independent service units, functional applications such as ADT (admission-discharge-transfer) and billing, and within primary care and specialty units. The first uses of computers in hospitals were for billing and accounting purposes and were developed on large, monolithic mainframe computers (see Chap. 13). Initially the cost of computers restricted expansion into clinical areas. But in the 1960s, hospital information systems (HISs) were developed to support service operations within a hospital. These systems followed a pattern of diversity similar to that seen in the health care system itself. As new functions were added in the 1970s, they were implemented on mainframe computers and were managed by a data processing staff that usually was independent of the clinical and even of the administrative staff. The advent of the minicomputer supported the development of departmental systems, such as those for the

[2] Interestingly, medical informaticians were responsible for the second ANSI standard language: MUMPS (now known as M).

clinical laboratory, radiology department, or pharmacy. With the advent of minicomputers, departmental systems were introduced but connectivity to other parts of the hospital was either by paper or independent electronic systems. It was common to see two terminals sitting side-by-side with an operator typing data from one system to another. Clinical systems, as they developed, continued to focus on dedicated departmental operations and clinical-specialty systems and thus did not permit the practicing physician to see a unified view of the patient. Most HISs were either supported entirely by a single vendor or were still functionally independent and unconnected.

In the 1980s, the need to move laboratory data directly into developing electronic health records systems (although this term was not used then), early standards were created in ASTM (formerly, the American Society for Testing and Materials, see Sect. 7.3.2) for the transfer of laboratory data from local and commercial laboratories (American Society for Testing and Materials 1999). In the late 1980s, Simborg and others developed an HIS by interfacing independent systems using a "Best of Breed" approach to create an integrated HIS (Simborg et al. 1983) Unfortunately, the cost of developing and maintaining those interfaces was prohibitive, and the need for a broader set of standards was realized. This effort resulted in the creation of the **standards development organization** (SDO) Health Level Seven (**HL7**) in 1987. Other SDOs were created in this same time frame: EDIFACT by the United Nations and ASC X12N by ANSI to address standards for claims and billing, IEEE for device standards, ACR/NEMA (later DICOM) for imaging standards, and NCPDP for prescription standards. Internationally, the 1990s saw the creation of the European Normalization Committee (CEN) and ISO Technical Committee 215 for Health Informatics (TC251). These organizations are described in more detail in Sect. 7.3.2.

7.2.2 Transitioning Standards to Meet Present Needs

Early standards were usually applied within a single unit or department in which the standards addressed mainly local requirements. Even then, data acquired locally came from another source introducing the need for additional standards. These many pressures caused health care information systems to change the status quo such that data collected for a primary purpose could be reused in a multitude of ways. Newer models for health care delivery, such as integrated delivery networks, health maintenance organizations (HMOs), preferred provider organizations (PPOs), and now accountable care organizations (ACOs) have increased the need for coordinated, integrated, and consolidated information (see Chap. 11), even though the information comes from disparate departments and institutions. Various management techniques, such as continuous quality improvement and case management, require up-to-date, accurate abstracts of patient data. Post hoc analyses for clinical and outcomes research require comprehensive summaries across patient populations. Advanced tools, such as clinical workstations (Chap. 5) and decision-support systems (Chap. 3), require ways to translate raw patient data into generic forms for tasks as simple as summary reporting and as complex as automated medical diagnosis. All these needs must be met in the existing setting of diverse, interconnected information systems—an environment that cries out for implementation of standards.

One obvious need is for standardized identifiers for individuals, health care providers, health plans, and employers so that such participants can be recognized across systems. Choosing such an identifier is much more complicated than simply deciding how many digits the identifier should have. Ideal attributes for these sets of identifiers have been described in a publication from the ASTM (1999). The identifier must include a check digit to ensure accuracy when the identifier is entered by a human being into a system. A standardized solution must also determine mechanisms for issuing identifiers to individuals, facilities, and organizations; for maintaining databases of identifying information; and for authorizing access to such information (also see Chap. 5).

The Centers for Medicare and Medicaid Services (CMS), has defined a National Provider

Identifier (NPI) as a national standard. This number is a seven-character alphanumeric base identifier plus a one-character check digit. No meaning is built into the number, each number is unique, it is never reissued, and alpha characters that might be confused with numeric characters have been eliminated (e.g., 0, 1, 2, 4, and 5 can be confused with O, I or L, Z, Y, and S). CMS was tasked to define a Payer ID for identifying health care plans. The Internal Revenue Service's employer identification number has been adopted as the Employer Identifier.

The most controversial issue is identifying each individual or patient. Many people consider assignment and use of such a number to be an invasion of privacy and are concerned that it could be easily linked to other databases. Public Law 104–191, passed in August 1996 (see Sect. 7.3.2), required that Congress formally define suitable identifiers. Pushback by privacy advocates and negative publicity in the media resulted in Congress declaring that this issue would not be moved forward until privacy legislation was in place and implemented (see Chap. 10). The Department of Health and Human Services has recommended the identifiers discussed above, except for the person identifier. This problem has still not been resolved, although the momentum for creating such a unique personal identified seems to be increasing. The United States is one of the few developed countries without such an identifier.

7.2.3 Settings Where Standards Are Needed

A hospital admissions system records that a patient has the diagnosis of diabetes mellitus, a pharmacy system records that the patient has been given gentamicin, a laboratory system records that the patient had certain results on kidney function tests, and a radiology system records that a doctor has ordered an X-ray examination for the patient that requires intravenous iodine dye. Other systems need ways to store these data, to present the data to clinical users, to send warnings about possible drug-drug interactions, to recommend dosage changes, and to follow the patient's outcome. A standard for coding patient data is nontrivial when one considers the need for agreed-on definitions, use of qualifiers, differing (application-specific) levels of granularity in the data, and synonymy, not to mention the breadth and depth that such a standard would need to have.

The inclusion of medical knowledge in clinical systems is becoming increasingly important and commonplace. Sometimes, the knowledge is in the form of simple facts such as the maximum safe dose of a medication or the normal range of results for a laboratory test. Much medical knowledge is more complex, however. It is challenging to encode such knowledge in ways that computer systems can use (see Chap. 22), especially if one needs to avoid ambiguity and to express logical relations consistently. Thus the encoding of clinical knowledge using an accepted standard would allow many people and institutions to share the work done by others. One standard designed for this purpose is the Arden Syntax, discussed in Chap. 3.

Because the tasks we have described require coordination of systems, methods are needed for transferring information from one system to another. Such transfers were traditionally accomplished through custom-tailored point-to-point interfaces, but this technique has become unworkable as the number of systems and the resulting permutations of necessary connections have grown. A current approach to solving the multiple-interface problem is through the development of messaging standards. Such messages must depend on the preexistence of standards for patient identification and data encoding.

Although the technical challenges are daunting, methods for encoding patient data and shipping those data from system to system are not sufficient for developing practical systems. Security must also be addressed before such exchanges can be allowed to take place. Before a system can divulge patient information, it must ensure that requesters are who they say they are and that they are permitted access to the requested information (see Chap. 5). Although each clinical system can have its own security features, system

builders would rather draw on available standards and avoid reinventing the wheel. Besides, the secure exchange of information requires that interacting systems use standard technologies. Fortunately, many researchers are busy developing such standards.

7.3 Standards Undertakings and Organizations

It is helpful to separate our discussion of the general process by which standards are created from our discussion of the specific organizations and the standards that they produce. The process is relatively constant, whereas the organizations form, evolve, merge, and are disbanded. This section will discuss how standards are created then identify the many SDOs and an overview of the types of standards they create. This section will also identify other groups and organizations that contribute or relate to standards activities.

7.3.1 The Standards Development Process

The process of creating standards is biased and highly competitive. Most standards are created by volunteers who represent multiple, disparate stakeholders. They are influenced by direct or indirect self-interest rather than judgment about what is best or required. The process is generally slow and inefficient; multiple international groups create competitive standards; and new groups continue to be formed as they become aware of the need for standards and do not look to see what standards exist. Yet, the process of creating standards largely works, and effective standards are created.

There are four ways in which a standard can be produced:

1. Ad hoc method: A group of interested people and organizations (e.g., laboratory-system and hospital-system vendors) agree on a standard specification. These specifications are informal and are accepted as standards through mutual agreement of the participating groups. A standard example produced by this method is the DICOM standard for medical imaging.
2. De facto method: A single vendor controls a large enough portion of the market to make its product the market standard. An example is Microsoft's Windows.
3. Government-mandate method: A government agency, such as CMS or the National Institute for Standards and Technology (NIST) creates a standard and legislates its use. An example is the HIPPA standard.
4. Consensus method: A group of volunteers representing interested parties works in an open process to create a standard. Most health care standards are produced by this method. An example is the Health Level 7 (HL7) standard for clinical-data interchange (Fig. 7.1).

The process of creating a standard proceeds through several stages (Libicki 1995). It begins with an identification stage, during which someone becomes aware that there exists a need for a standard in some area and that technology has reached a level that can support such a standard. For example, suppose there are several laboratory systems sending data to several central hospital systems—a standard message format would allow each laboratory system to talk to all the hospital systems without specific point-to-point interface programs being developed for each possible laboratory-to-laboratory or laboratory-to-hospital combination. If the time for a standard is ripe, then several individuals can be identified and organized to help with the conceptualization stage, in which the characteristics of the standard are defined. What must the standard do? What is the scope of the standard? What will be its format? In the early years of standards development, this approach led the development of standards, and the process was supported by vendors and providers. As those early standards have become successful, the need for "gap-standards" has arisen. These gap-standards have no champion but are necessary for completeness of an interoperable data exchange network. The need for these standards is not as obvious as for the primary standards, people are less likely to volunteer to do work, putting stress on the voluntary approach.

Fig. 7.1 Standards development meetings. The development of effective standards often requires the efforts of dedicated volunteers, working over many years. Work is often done in small committee meetings and then presented to a large group to achieve consensus. Here we see meetings of the HL7. Vocabulary Technical Committee (*top*) and an HL7 plenary meeting (*bottom*). See Sect. 7.5.2 for a discussion on HL7

In such cases, the need of such standards must be sold to the volunteers or developed by paid professionals.

Let us consider, for purposes of illustration, how a standard might be developed for sending laboratory data in electronic form from one computer system to another in the form of a message. The volunteers for the development might include laboratory system vendors, clinical users, and consultants. One key discussion would be on the scope of the standard. Should the standard deal only with the exchange of laboratory data, or should the scope be expanded to include other types of data exchange? Should the data elements being exchanged be sent with a XML tag identifying the data element, or should the data be defined positionally? In the ensuing discussion stage, the participants will begin to create an outline that defines content, identifies critical issues, and produces a time line. In the discussion, the pros and cons of the various concepts are discussed. What will be the specific form for the standard? For example, will it be message based or document based? Will the data exchange be based on a query or on a trigger event? Will the standard define the message content, the message syntax, the terminology, the network protocol, or a subset of these issues? How will a data model or information model be incorporated?

The participants are generally well informed in the domain of the standard, so they appreciate the needs and problems that the standard must address. Basic concepts are usually topics for heated discussion; subsequent details may follow at an accelerated pace. Many of the participants will have experience in solving problems to be addressed by the standard and will protect their own approaches. The meanings of words are often debated. Compromises and loosely defined terms are often accepted to permit the process to move forward. For example, the likely participants would be vendors of competing laboratory systems and vendors of competing HISs. All participants would be familiar with the general problems but would have their own proprietary approach to solving them. Definitions of basic concepts normally taken for granted, such as what constitutes a test or a result, would need to be clearly stated and agreed on.

The writing of the draft standard is usually the work of a few dedicated individuals—typically people who represent the vendors in the field. Other people then review that draft; controversial points are discussed in detail and solutions are proposed and finally accepted. Writing and refining the standard is further complicated by the introduction of people new to the process who have not been privy to the original discussions

and who want to revisit points that have been resolved earlier. The balance between moving forward and being open is a delicate one. Most standards-writing groups have adopted an **open standards development policy**: Anyone can join the process and can be heard. Most standards development organizations—certainly those by accredited groups— support an open balloting process. A draft standard is made available to all interested parties, inviting comments and recommendations. All comments are considered. Negative ballots must be addressed specifically. If the negative comments are persuasive, the standard is modified. If they are not, the issues are discussed with the submitter in an attempt to convince the person to remove the negative ballot. If neither of these efforts is successful, the comments are sent to the entire balloting group to see whether the group is persuaded to change its vote. The resulting vote then determines the content of the standard. Issues might be general, such as deciding what types of laboratory data to include (pathology? blood bank?), or specific, such as deciding the specific meanings of specific fields (do we include the time the test was ordered? specimen drawn? test performed?).

A standard will generally go through several versions on its path to maturity. The first attempts at implementation are frequently met with frustration as participating vendors interpret the standard differently and as areas not addressed by the standard are encountered. These problems may be dealt with in subsequent versions of the standard. Backward compatibility is a major concern as the standard evolves. How can the standard evolve, over time, and still be economically responsible to both vendors and users? An implementation guide is usually produced to help new vendors profit from the experience of the early implementers.

A critical stage in the life of a standard is early implementation, when acceptance and rate of implementation are important to success. This process is influenced by accredited standards bodies, by the federal government, by major vendors, and the marketplace. The maintenance and promulgation of the standard are also important to ensure widespread availability and continued value of the standard. Some form of conformance testing is ultimately necessary to ensure that vendors adhere to the standard and to protect its integrity.

Producing a standard is an expensive process in terms of both time and money. Vendors and users must be willing to support the many hours of work, usually on company time; the travel expense; and the costs of documentation and distribution. In the United States, the production of a consensus standard is voluntary, in contrast to in Europe and elsewhere, where most standards development is funded by governments.

An important aspect of standards is conformance, a concept that covers compliance with the standard and also usually includes specific agreements among users of the standard who affirm that specific rules will be followed. A conformance document identifies specifically what data elements will be sent, when, and in what form. Even with a perfect standard, a conformance document is necessary to define business relationships between two or more partners.

The creation of the standard is only the first step. Ideally the first standard would be a Draft Standard for Trial Use (DSTU), and two or more vendors would implement and test the standard to identify problems and issues. Those items would be corrected, and in a short period of time (usually 1 year) the standards would be advanced to a normative stage.

Even then, the process is only beginning. Implementation that conforms to the standard is essential if the true value of the standard is to be realized. The use of most standards is enhanced by a certification process in which a neutral body certifies that a vendor's product, in fact, does comply and conform to the standard.

There is currently no body that certifies conformance of specific standards from a vendor. There is, however, the certification of an application that uses standards. In 2010, the Office of the National Coordinator (ONC)[3] engaged with the CCHIT to certify EHR products. That certification process evolved in 2011 to include

[3] http://www.healthit.gov/newsroom/about-onc (accessed 4/26/13)

eight groups that could certify EHR products, and to date over 500 EHR products have been certified. The certification process is still undergoing change.

7.3.2 Data Standards Organizations

Sometimes, standards are developed by organizations that need the standard to carry out their principal functions; in other cases, coalitions are formed for the express purpose of developing a particular standard. The latter organizations are discussed later, when we examine the particular standards developed in this way. There are also standards organizations that exist for the sole purpose of fostering and promulgating standards. In some cases, they include a membership with expertise in the area where the standard is needed. In other cases, the organization provides the rules and framework for standard development but does not offer the expertise needed to make specific decisions for specific standards, relying instead on participation by knowledgeable experts when a new standard is being studied.

This section describes in some detail several of the best known SDOs. Our goal has been to familiarize you with the names or organizational and historical aspects of the most influential health-related standards groups. Additional organizations are listed in Table 7.1. In Sect. 7.5 we describe many of the most important standards. For a detailed understanding of an organization or the standards it has developed, you will need to refer to current primary resources. Many of the organizations maintain Web sites with excellent current information on their status.

7.3.2.1 ISO Technical Committee 215—Health Informatics

In 1989, interests in the European Committee for Standardization (CEN) and the United States led to the creation of Technical Committee (TC) 215 for Health Information within ISO.

TC 215 meets once in a year as a TC and once as a Joint Working Group. TC 215 follows rather rigid procedures to create ISO standards. Thirty-five countries are active participants in the TC with another 23 countries acting as observers. While the actual work is done in the working groups, the balloting process is very formalized—one vote for each participating country. For most work there are a defined series of steps, beginning with a New Work Item Proposal and getting five countries to participate; a Working Document, a Committee Document; a Draft International Standard, a Final Draft International Standard (FDIS); and finally an International Standard. This process, if fully followed, takes several years to produce an International Standard. Under certain conditions, a fast track to FDIS is permitted. Technical Reports and Technical Specifications are also permitted.

The United States has been assigned the duties of Secretariat, and that function is carried out by ANSI. Currently, AHIMA acts for ANSI as Secretariat. AHIMA also serves as the U.S. Technical Advisory Group Administrator, which represents the U.S. position in ISO.

A recent change in ISO policy is permitting standards developed by other bodies to move directly to become ISO standards. Originally, CEN and ISO developed an agreement, called the Vienna Agreement, which permits CEN standards to move into ISO for parallel development and be balloted in each organization. In 2000, a new process was added with the Institute of Electrical and Electronics Engineers (IEEE) called the ISO/IEEE Standards partners in which IEEE standards could be moved directly to ISO for approval as ISO standards. HL7 also has an agreement with ISO which permits HL7 standards to be submitted to TC 215 for approval to become ISO standards.

7.3.2.2 European Committee for Standardization Technical Committee 251

The European Committee for Standardization (CEN) established, in 1991, Technical Committee 251 (TC 251—not to be confused with ISO TC 215 described below) for the development of standards for health care informatics. The major goal of TC 251 is to develop standards for communication among independent medical information systems so that clinical and management

Table 7.1 List of health-related standards groups

Organization	Description
Accredited Standards Organization X12 (ASC X12)	ASC X12 was charted by the American National Standards Institute (ANSI) to develop and maintain several data interchange standards.
American National Standards Institute	ANSI is a private, nonprofit membership organization responsible for approving official American National Standards. ANSI assists standards developers and users from the private sector and from government to reach consensus on the need for standards.
ASTM	ASTM (formerly known as the American Society for Testing and Materials) develops standard test methods for materials, products, systems, and services.
Clinical Data Interchange Standards Consortium (CDISC)	CDISC creates standards in support of the clinical research community. Its membership includes pharma, academic researchers, vendors and others.
European Committee for Standardization Technical Committee 251	The European Committee for Standardization (CEN) established Technical Committee 251 in 1991. The major goal of TC 251 is to develop standards for communication among independent medical information systems so that clinical and management data produced by one system could be transmitted to another system.
GS1	GS1 is a global standards organization with over one million members world-wide. It has a presence in over 100 countries. Its primary standards relate to the supply chain and for assigning object identifiers and standards for bar codes.
Health Level Seven International (HL7)	HL7 was founded in 1987 to create standards for the exchange of clinical data. HL7 is ANSI accredited, and many of the HL7 standards have been designed for required use by the U.S. government as part of Meaningful Use.
Institution of Electrical and Electronic Engineering (IEEE)	The IEEE is an international SDO that has developed standards in many areas. In the health care area, the applicable standards are for the interfacing of instruments and mobile devices.
Integrating the Health care Enterprise	The goal of the Integrating the Health care Enterprise (IHE) initiative is to stimulate integration of health care information resources. IHE enables vendors to direct product development resources toward building increased functionality rather than redundant interfaces.
International Health Terminology Standards Organization (IHTSDO)	The primary purpose of IHTSDO is the continued development and maintenance of SNOMED-CT. Member countries make SNOMED-CT freely available to its citizens. IHTSDO has a number of Special Interest Groups, including Anesthesia, Concept Model, Education, Implementation, International Pathology & Laboratory Medicine, Mapping, Nursing, Pharmacy, and Translation.
ISO Technical Committee 215—Health Informatics	Formed by the European Committee for Standardization (CEN) and the United States in 1989 to create ISO standards in health informatics.
Joint Initiative Council (JIC)	The Joint Initiative Council was formed to enable the creation of common, timely health information standards by addressing gaps, overlaps, and competitive standards efforts by jointly creating, as equal partners, health IT standards. Current members of JIC include ISO, CEN, HL7, CDISC, IHTSDO, and GS1. Ownership of standards created by this group is shared among the participating partners.
National Council for Prescription Drug Programs (NCPDP)	NCPDP creates data interchange standards for the pharmacy services, most specifically for the prescribing process (ePrescribing) and reimbursement for medications.
National Institute of Standards and Technology	Non-regulatory federal agency within the U.S. Department of Commerce whose mission is to develop and promote measurement, standards, and technology to enhance productivity, facilitate trade, and improve the quality of life. In health care, NIST is providing measurement tools, manufacturing assistance, research and development support and quality guidelines as an effort to contain health care costs without compromising quality.

(continued)

Table 7.1 (continued)

Organization	Description
National Quality Forum	Private, not for profit, whose purpose is to develop and implement a national strategy for health care quality measurement and reporting.
openEHR	International organization that develops standards loosely based on ISO 13606. Key work is the development of archetypes.
Standards Development Organizations Charter Committee (SCO)	The SCO was created through an initiative of NCPDP to harmonize the efforts of U.S. SDOs. Membership is limited to SDOs, but associate membership is award to other organizations with an interest in U.S. Health Data Standards.
The Digital Imaging and Communications in Medicine (DICOM)	DICOM was created through a joint effort by the American College of Radiology and NEMA (and was thus initially known as the ACR/NEMA) to develop standards for imaging and waveforms.
U.S. Technical Advisory Group	Represents the U.S. interests in ISO.
Workgroup for Electronic Data Interchange	Broad health care coalition to promote greater health care electronic commerce and connectivity; one of the four organizations named specifically in HIPAA to be consulted in the development of health care standards that would be selected to meet HIPAA requirements.

data produced by one system could be transmitted to another system. The organization of TC 251 parallels efforts in the United States through various working groups. These groups similarly deal with data interchange standard, medical record standards, code and terminology standards, imaging standards, and security, privacy and confidentiality. Both Europe and the United States are making much effort toward coordination in all areas of standardization. Draft standards are being shared. Common solutions are being accepted as desirable. Groups are working together at various levels toward a common goal.

CEN has made major contributions to data standards in health care. One important CEN prestandard ENV 13606 on the electronic health record (EHR) is being advanced by CEN as well as significant input from Australia and the OpenEHR Foundation. There is an increasing cooperation among the CEN participants and several of the U.S. standards bodies. CEN standards may be published through the ISO as part of the Vienna Agreement.

7.3.2.3 Health Level Seven International (HL7)

Health Level 7 was founded as an ad hoc standards group in March 1987 to create standards for the exchange of clinical data, adopting the name "HL7" to reflect the application (seventh) level of the OSI (see Sect. 7.5.1) reference model. The primary motivation was the creation of a Hospital Information System from "Best of Breed" components. The HL7 data interchange standard (version 2.n series) reduced the cost of interfacing between disparate systems to an affordable cost. Today HL7 is one of the premier SDOs in the world. It has become an international standards body with approximately 40 Affiliates, over 500 organizational members and over 2,200 individual members. HL7 is ANSI accredited, and many of the HL7 standards are required by the U.S. government as part of the certification requirements of Meaningful Use (see Chap. 13). The HL7 standards are described in Sect. 7.5.2.3, below.

7.3.2.4 Integrating the Health Care Enterprise

The goal of the Integrating the Health care Enterprise (IHE) initiative is to stimulate integration of health care information resources. While information systems are essential to the modern health care enterprise, they cannot deliver full benefits if they operate using proprietary protocols or incompatible standards. Decision makers need to encourage comprehensive integration among the full array of imaging and information systems.

IHE is sponsored jointly by the Radiological Society of North America (RSNA) and the

HIMSS. Using established standards and working with direction from medical and information technology professionals, industry leaders in health care information and imaging systems cooperate under IHE to agree upon implementation profiles for the transactions used to communicate images and patient data within the enterprise. Their incentive for participation is the opportunity to demonstrate that their systems can operate efficiently in standards-based, multi-vendor environments with the functionality of real HISs. Moreover, IHE enables vendors to direct product development resources toward building increased functionality rather than redundant interfaces.

7.3.2.5 International Health Terminology Standards Organization (IHTSDO)

IHTSDO was founded in 2007 with nine charter members. Currently 19 countries, including the United States, are members. The primary purpose of IHTSDO is the continued development and maintenance of the Systematized Nomenclature of Medicine – Clinical Terms (SNOMED-CT). Member countries make SNOMED-CT freely available to its citizens. SNOMED has a number of Special Interest Groups: Anesthesia, Concept Model, Education, Implementation, International Pathology & Laboratory Medicine, Mapping, Nursing, Pharmacy, and Translation. SNOMED-CT is described in Sect. 7.4.4.7, below.

7.3.2.6 American National Standards Institute

ANSI is a private, nonprofit membership organization founded in 1918. It originally served to coordinate the U.S. voluntary census standards systems. Today, it is responsible for approving official American National Standards. ANSI membership includes over 1,100 companies; 30 government agencies; and 250 professional, technical, trade, labor, and consumer organizations.

ANSI does not write standards; rather, it assists standards developers and users from the private sector and from government to reach consensus on the need for standards. It helps them to avoid duplication of work, and it provides a forum for resolution of differences. ANSI administers the only government-recognized system for establishing American National Standards. ANSI also represents U.S. interests in international standardization. ANSI is the U.S. voting representative in the ISO and the International Electrotechnical Commission (IEC). There are three routes for a standards development body to become ANSI approved so as to produce an American National Standard: Accredited Organization; Accredited Standards Committee (ASCs); and Accredited Canvass.

An organization that has existing organizational structure and procedures for standards development may be directly accredited by ANSI to publish American National Standards, provided that it can meet the requirements for due process, openness, and consensus. HL7 (discussed in Sect. 7.5.2) is an example of an ANSI Accredited Organization.

ANSI may also create internal ASCs to meet a need not filled by an existing Accredited Organization. ASC X12 is an example of such a committee.

The final route, Accredited Canvass, is available when an organization does not have the formal structure required by ANSI. Through a canvass method that meets the criterion of balanced representation of all interested parties, a standard may be approved as an American National Standard. ASTM (discussed below) creates its ANSI standards using this method.

7.3.2.7 ASTM

ASTM (formerly known as the American Society for Testing and Materials) was founded in 1898 and chartered in 1902 as a scientific and technical organization for the development of standards for characteristics and performance of materials. The original focus of ASTM was development of standard test methods. The charter was subsequently broadened in 1961 to include products, systems, and services, as well as materials. ASTM is the largest nongovernment source of standards in the United States. It has over 30,000 members who reside in over 90 different countries. ASTM is a charter member of ANSI. ASTM Committee E31 on

Table 7.2 ASTM E31 subcommittees

Subcommittee	Medical information standard
E31.01	Controlled Vocabularies for Health Care Informatics
E31.10	Pharmaco-Informatics Standards
E31.11	Electronic Health Record Portability
E31.13	Clinical Laboratory Information Management Systems
E31.14	Clinical Laboratory Instrument Interface
E31.16	Interchange of Electrophysiological Waveforms and Signals
E31.17	Access, Privacy, and Confidentiality of Medical Records
E31.19	Electronic Health Record Content and Structure
E31.20	Data and System Security for Health Information
E31.21	Health Information Networks
E31.22	Health Information Transcription and Documentation
E31.23	Modeling for Health Informatics
E31.24	Electronic Health Record System Functionality
E31.25	XML for Document Type Definitions in Health care
E31.26	Personal (Consumer) Health Records
E31.28	Electronic Health Records

Computerized Systems is responsible for the development of the medical information standards. Table 7.2 shows the domains of its various subcommittees.

7.4 Detailed Clinical Models, Coded Terminologies, Nomenclatures, and Ontologies

As discussed in Chap. 2, the capture, storage, and use of clinical data in computer systems is complicated by lack of agreement on terms and meanings. In recent years there has also been a growing recognition that just standardizing the terms and codes used in medicine is not sufficient to enable interoperability. The structure or form of medical data provides important context for computable understanding of the data. Terms and codes need to be interpreted in the context of clinical information models. The many terminologies and detailed clinical modeling activities discussed in this section have been developed to ease the communication of coded medical information.

7.4.1 Motivation for Structured and Coded Data

The structuring and encoding of medical information is a basic function of most clinical systems. Standards for such structuring and encoding can serve two purposes. First, they can save system developers from reinventing the wheel. For example, if an application allows caregivers to compile problem lists about their patients, using a standard structure and terminology saves developers from having to create their own. Second, using commonly accepted standards can facilitate exchange of data, applications, and clinical decision support logic among systems. For example, if a central database is accepting clinical data from many sources, the task is greatly simplified if each source is using the same logical data structure and coding scheme to represent the data. System developers often ignore available standards and continue to develop their own solutions. It is easy to believe that the developers have resisted adoption of standards because it is too much work to understand and adapt to any system that was "not invented here." The reality, however, is that the available standards are often inadequate for the needs of the users (in this case, system developers). As a result, no standard terminology enjoys the wide acceptance sufficient to facilitate the second function: exchange of coded clinical information.

The need for detailed clinical models is directly related to the second goal discussed above, that of creating interoperability between systems. The subtle relationship between terminologies and models is best understood using a couple of examples. If a physician wants to record the idea that a patient had "chest pain that radiated to the back", the following coded terms could be used from SNOMED-CT (see Sect. 7.4.4, below):

51185008	Chest (Thoracic structure)
22253000	Pain
8754004	Radiating to
302552004	Back (Entire back)

However, by just reordering the codes, as shown below, one could use the same codes to represent the idea of "back pain radiating to the chest."

302552004	Back (Entire back)
22253000	Pain
8754004	Radiating to
51185008	Chest (Thoracic structure)

In this simple example, changing the order of the codes changes the implied meaning. Creating an ordering of the codes is one way of imposing structure. The need for structure is often overlooked because people can make sense of the set of codes because of medical context and knowledge. Chest pain radiating to the back is much more common than back pain that radiates to the chest, so most clinicians seeing the data would not be confused. However, computer systems do not have the same kind of intutition. To render the representation so that it is unambiguous even to a computer one can make an explicit structure that includes the elements of *finding*, *finding location*, *radiates to location*:
Finding: PainFinding
Location: Chest
Radiates to location: Back

Using the SNOMED-CT codes as values in a data structure removes the ambiguity that exists if just the list of codes is used as a representation of finding.

A second example of why detailed clinical models are needed is closely related to the idea of **pre-coordination** and **post-coordination**. In this example, we are trying to represent the measurement of a patient's dry weight. Two different sites could choose two different ways of representing the measurement:

Site#1:	Dry weight:	70	●kg	○lbs			
Site #2:	Weight:	70	●kg	○lbs	○Wet	●Dry	○Ideal

At the first site, they have chosen a single concept or label to represent the idea of *dry weight*. The second site decided to use one concept for the generic idea of a *weight measurement*, and to couple that measurement, with a second piece of information indicating the *kind of weight* (dry, wet, or ideal). The first site is using a precoordinated approach, while the second site is using a postcoordinated approach. If you put the measurements from these two different sites into spreadsheets or databases, the information might be structured as shown below:

Site #1:	Patient ID	Date	Observation	Value	Unit
	12345678	10/20/2012	Dry weight	70	kg
	12345678	10/26/2012	Current weight	72	kg

Site #2:	Patient ID	Date	Observation	Type	Value	Unit
	12345678	10/20/2012	Weight	Dry	70	kg
	12345678	10/26/2012	Weight	Current	72	kg

It is intuitively clear that the information content of both the pre- and post-coordination representations is identical. However, if data from both sites were being combined to support a clinical study, it is equally clear that the data from the two sites cannot be referenced or manipulated in exactly the same way. The information that is in the *Observation Type* column in the first site is being represented by information in two columns (*Observation Type* and *Weight type*) in the second site. The logic to query and then calculate the desired weight loss for a patient is different for data from the two sites. Even though the information content is equivalent in the two cases, a computer would need models to know how to transform data from one site into data that could be used interoperably with data from the second site.

The weight example is a very simple case. Problem list data, family history data, complex physical exam observations, and the use of negation all have much greater complexity and degrees of freedom in how the data could be represented, and the need for models to formally represent the explicit structure is even more

evident. Because of this interdependence of structure with terms and codes, we will discuss terminologies and detailed clinical models together in this section. We will first discuss detailed clinical models, and then discuss how terminologies relate to the models.

7.4.2 Detailed Clinical Models

The creation of unambiguous data representation is a combination of creating appropriate structures (models) for representing the form of the data and then linking or "binding" specific sets of codes to the coded elements in the structures. Several modeling languages or formalisms have been found to be useful in describing the structure of the data. They include:

- UML – the Unified Modeling Language, Object Management Group
- ADL – Archetype Definition Language, OpenEHR Foundation
- CDL – Constraint Definition Language, General Electric and Intermountain Health care
- MIF – Model Interchange Format, Health Level Seven International Inc.
- OWL – Web Ontology Language, World Wide Web Consortium

All languages used for clinical modeling need to accomplish at least two major things: they need to show the "logical" structure of the data, and they need to show how sets of codes from standard terminologies participate in the logical structure. Defining the logical structure is simply showing how the named parts of a model relate to one another. Model elements can be contained in other elements, creating hierarchies of elements. It is also import to specify which elements of the model can occur more than once (cardinality), which elements are required, and which are optional. Terminology binding is the act of creating connections between the elements in a model and concepts in a coded terminology. For each coded element in a model, the set of allowed values for the coded element are specified. The HL7 Vocabulary Working Group has created a comprehensive discussion of how value sets can be defined and used with information models [HL7 Core Vocabulary Foundation].

There are many clinical information modeling activities worldwide. Some of the most important activities are briefly listed below.

- HL7 Activities
  - HL7 Detailed Clinical Models – This group has developed a method for specifying clinical models based on the HL7 Reference Information Model (RIM) that guarantees that data that conforms to the model could be sent in HL7 Version 3 messages.
  - HL7 Clinical Document Architecture (CDA) Templates – This group has defined a standard way of specifying the structure of data to be sent in XML documents that conform to the CDA standard.
  - HL7 TermInfo – This Workgroup of HL7 has specified a set of guidelines for how SNOMED-CT codes and concepts should be used in conjunction with the HL7 RIM to represent data sent in HL7 Version 3 messages.
- The openEHR Foundation is developing models based on a core reference model and the Archetype Definition Language. This approach has been adopted by several national health information programs.
- EN 13606 is developing models based on the ISO/CEN 13606 standard and core reference model.
- Tolven is an open source initiative that creates models as part of an overall architecture to support open Electronic Health Record implementation.
- The US Veterans Administration (VA) is creating models for integrating data across all VA facilities and for integration with military hospitals that are part of the US Department of Defense. The modeling is done primarily using Unified Modeling Language.
- US Department of Defense is creating models for integrating data across all DoD facilities and for integration with VA facilities. The modeling is done primarily using Unified Modeling Language.
- The National Health Service in the United Kingdom (UK) is developing the Logical

- **Object:** Any part of the perceivable or conceivable world
- **Name:** Designation of an object by a linguistic expression
- **Concept:** A unit of thought constituted through abstraction on the basis of properties common to a set of objects
- **Term:** Designation of a defined concept in a special language by a linguistic expression
- **Terminology:** Set of terms representing the system of concepts of a particular subject field
- **Nomenclature:** System of terms that is elaborated according to preestablished naming rules
- **Dictionary:** Structured collection of lexical units, with linguistic information about each of them
- **Vocabulary:** Dictionary containing the terminology of a subject field

Fig. 7.2 Terminologic terms, adapted from ISO Standard 1087. Terms not defined here—such as definition, lexical unit, and linguistic expression—are assumed by the Standard to have common meanings

Record Architecture to provide models for interoperability across all health care facilities in the UK. The modeling is done primarily using UML.

- Clinical Element Models - Intermountain Health care and General Electric have created a set of detailed clinical models using a core reference model and Constraint Definition Language. The models are free-for-use and are available for download from the Internet.
- SHARE Models – CDISC is creating models to integrate data collected as part of clinical trials.
- SMArt Team – This group at Boston Children's Hospital is defining standard application programming interfaces (APIs) for services that store and retrieve medical data.
- Clinical Information Modeling Initiative (CIMI) – This is an international consortium that has the goal of establishing a free-for-use repository of detailed clinical models, where the models are expressed in a single common modeling language with explicit bindings to standard terminologies.

7.4.3 Vocabularies, Terminologies, and Nomenclatures

In discussing coding systems, the first step is to clarify the differences among a **terminology**, a **vocabulary**, and a **nomenclature**. These terms are often used interchangeably by creators of coding systems and by authors discussing the subject. Fortunately, although there are few accepted standard terminologies, there is a generally accepted standard about terminology: ISO Standard 1087 (Terminology—Vocabulary). Figure 7.2 lists the various definitions for these terms. For our purposes, we consider the currently available standards from the viewpoint of their being terminologies.

The next step in the discussion is to determine the basic use of the terminology. In general, there are two different levels relevant to medical data encoding: abstraction and representation. **Abstraction and classification** entail examination of the recorded data and then selection of items from a terminology with which to label the data. For example, a patient might be admitted to the hospital and have a long and complex course; for the purposes of billing, however, it might be relevant that the patient was diagnosed only as having had a myocardial infarction. Someone charged with abstracting the record to generate a bill might then reduce the entire set of information to a single code. **Representation**, on the other hand, is the process by which as much detail as possible is coded. Thus, for the medical record example, the representation might include codes for each physical finding noted, laboratory test performed and its result, and medication administered.

When we discuss a controlled terminology, we should consider the domain of discourse. Virtually any subject matter can be coded, but there must be a good match with any standard selected for the purpose. For example, a terminology that only included diseases might be a poor choice for coding entries on a problem list because it might lack items such as "abdominal pain," "cigarette smoker," or "health maintenance."

The next consideration is the content of the standard itself. There are many issues, including the degree to which the standard covers the terminology of the intended domain; the degree to which data are coded by assembly of terms into descriptive phrases (post-coordination) versus selection of a single, precoordinated term; and the overall structure of the terminology (list, strict hierarchy, multiple hierarchy, semantic network, and so on). There are also many qualitative issues to consider, including the availability of synonyms and the possibility of redundant terms (i.e., more than one way to encode the same information).

Finally, we should consider the methods by which the terminology is maintained. Every standard terminology must have an ongoing maintenance process, or it will rapidly become obsolete. The process must be timely and must not be too disruptive to people using an older version of the terminology. For example, if the creators of the terminology choose to rename a code, what happens to the data previously recorded with that code?

7.4.4 Specific Terminologies

With these considerations in mind, let us survey some of the available controlled terminologies. People often say, tongue in cheek, that the best thing about standards is that there are so many from which to choose. We give introductory descriptions of a few current and common terminologies. New terminologies appear annually, and existing proprietary terminologies often become publicly available. When reviewing the following descriptions, try to keep in mind the background motivation for a development effort. All these standards are evolving rapidly, and one should consult the Web sites or other primary sources for the most recent information.

7.4.4.1 International Classification of Diseases and Its Clinical Modifications

One of the best known terminologies is the International Classification of Diseases (ICD). First published in 1893, it has been revised at roughly 10-year intervals, first by the Statistical International Institute and later by the World Health Organization (WHO). The Ninth Edition (ICD-9) was published in 1977 (World Health Organization, 1977) and the Tenth Edition (ICD-10) in 1992 (World Health Organization 1992). The ICD-9 coding system consists of a core classification of three-digit codes that are the minimum required for reporting mortality statistics to WHO. A fourth digit (in the first decimal place) provides an additional level of detail; usually .0 to .7 are used for more specific forms of the core term, .8 is usually "other," and .9 is "unspecified." Codes in the ICD-10 coding system start with an alpha character and consist of three to seven characters. In both systems, terms are arranged in a strict hierarchy, based on the digits in the code. For example, bacterial pneumonias are classified as shown in Figs. 7.3 and 7.4. In addition to diseases, ICD includes several "families" of terms for medical-specialty diagnoses, health status, disease-related events, procedures, and reasons for contact with health care providers.

ICD-9 has generally been perceived as inadequate for the level of detail desired for statistical reporting in the United States (Kurtzke 1979). In response, the U.S. National Center for Health Statistics published a set of **clinical modifications** (CM) (Commission on Professional and Hospital Activities 1978). **ICD-9-CM**, as it is known, is compatible with ICD-9 and provides extra levels of detail in many places by adding fourth-digit and fifth-digit codes. Figure 7.3 shows a sample of additional material. Most of the diagnoses assigned in the United States are coded in ICD-9-CM, allowing compliance with international treaty (by conversion to ICD-9) and supporting billing requirements (by conversion to diagnosis-related groups or DRGs). A clinical modification for ICD-10 has also been created; examples are shown in Fig. 7.4.

7.4.4.2 Diagnosis-Related Groups

Another U.S. creation for the purpose of abstracting medical records is the DRGs, developed initially at Yale University for use in

003 Other salmonella infections
- 003.2 Localized salmonella infections
  - 003.22 Salmonella pneumonia *

020 Plague
- 020.3 Primary pneumonic plague
- 020.4 Secondary pneumonic plague
- 020.5 Pneumonic plague, unspecified

021 Tularemia
- 021.2 Pulmonary tularemia *

022 Anthrax
- 022.1 Pulmonary anthrax

481 Pneumococcal pneumonia

482 Other bacterial pneumonia
- 482.0 Pneumonia due to *Klebsiella pneumoniae*
- 482.1 Pneumonia due to Pseudomonas
- 482.2 Pneumonia due to *Hemophilus influenzae*
- 482.3 Pneumonia due to Streptococcus
  - 482.30 Pneumonia due to Streptococcus, unspecified *
  - 482.31 Pneumonia due to Group A Streptococcus *
  - 482.32 Pneumonia due to Group B Streptococcus *
  - 482.39 Other streptococcal pneumonia *
- 482.4 Pneumonia due to Staphylococcus
  - 482.40 Pneumonia due to Staphylococcus, unspecified *
  - 482.41 Pneumonia due to *Staphylococcus aureus* *
  - 482.49 Other Staphylococcus pneumonia *
- 482.8 Pneumonia due to other specified bacteria
  - 482.81 Pneumonia due to anaerobes *
  - 482.82 Pneumonia due to *Escherichia coli* *
  - 482.83 Pneumonia due to other Gram-negative bacteria *
  - 482.84 Legionnaires' disease *
  - 482.89 Pneumonia due to other specified bacteria *
- 482.9 Bacterial pneumonia, unspecified

483 Pneumonia due to other specified organism
- 483.0 *Mycoplasma pneumoniae* *

484 Pneumonia in infectious diseases classified elsewhere *
- 484.3 Pneumonia in whooping cough *
- 484.5 Pneumonia in anthrax *

Fig. 7.3 Examples of codes in ICD-9 and ICD-9-CM (*) showing how bacterial pneumonia terms are coded. Tuberculosis terms, pneumonias for which the etiologic agent is not specified, and other intervening terms are not shown. Note that some terms, such as "Salmonella Pneumonia" were introduced in ICD-9-CM as a children of organism-specific terms, rather than under 482 (other bacterial pneumonia)

prospective payment in the Medicare program (3M Health Information System, updated annually). In this case, the coding system is an abstraction of an abstraction; it is applied to lists of ICD-9-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 severity of illness. The principal bases for the groupings are factors that affect cost and length of stay. Thus, a medical record containing the ICD-9-CM primary diagnosis of pneumococcal pneumonia (481) might be coded with one of 18 codes (Fig. 7.5), depending on associated conditions and procedures; additional codes are possible if the pneumonia is a secondary diagnosis.

A01 Typhoid and paratyphoid fevers
- A01.0 Typhoid Fever
  - A01.03 Typhoid Pneumonia *

A02 Other salmonella infection
- A02.2 Localized salmonella infections
  - A02.22 Salmonella pneumonia *

A20 Plague
- A20.2 Pneumonic plague

A22 Anthrax
- A22.1 Pulmonary anthrax

A37 Whooping cough
- A37.0 Whooping cough due to *Bordetella pertussis*
  - A37.01 Whooping cough due to *Bordetella pertussis* with pneumonia *
- A37.1 Whooping cough due to *Bordetella parapertussis*
  - A37.11 Whooping cough due to *Bordetella parapertussis* with pneumonia *
- A37.8 Whooping cough due to other Bordetella species
  - A37.81 Whooping cough due to other Bordetella species with pneumonia *
- A37.9 Whooping cough, unspecified
  - A37.91 Whooping cough, unspecified species with pneumonia *

A50 Congenital syphilis
- A50.0 Early congenital syphilis, symptomatic
  - A50.04 Early congenital syphilitic pneumonia *

A54 Gonococcal infection
- A54.8 Other gonococcal infection
  - A54.84 Gonococcal pneumonia *

J13 Pneumonia due to *Streptococcus pneumoniae*

J14 Pneumonia due to *Hemophilus influenzae*

J15 Bacterial pneumonia, not elsewhere classified
- J15.0 Pneumonia due to *Klebsiella pneumoniae*
- J15.1 Pneumonia due to Pseudomonas
- J15.2 Pneumonia due to staphylococcus
  - J15.20 Pneumonia due to staphylococcus, unspecified *
  - J15.21 Pneumonia due to *Staphylococcus aureus* *
  - J15.29 Pneumonia due to other staphylococcus *
- J15.3 Pneumonia due to streptococcus, group B
- J15.4 Pneumonia due to other streptococci
- J15.5 Pneumonia due to *Escherichia coli*
- J15.6 Pneumonia due to other aerobic Gram-negative bacteria
- J15.7 Pneumonia due to *Mycoplasma pneumoniae*
- J15.8 Other bacterial pneumonia
- J15.9 Bacterial pneumonia, unspecified

P23 Congenital pneumonia
- P23.2 Congenital pneumonia due to staphylococcus
- P23.3 Congenital pneumonia due to streptococcus, group B
- P23.4 Congenital pneumonia due to *Escherichia coli*
- P23.5 Congenital pneumonia due to Pseudomonas
- P23.6 Congenital pneumonia due to other bacterial agents

Fig. 7.4 Examples of codes in ICD-10 and ICD-10-CM (*) showing how bacterial pneumonia terms are coded. Tuberculosis terms, pneumonias for which the etiologic agent is not specified, and other intervening terms are not shown. Note that ICD-10 classifies Mycoplasma pneumoniae as a bacterium, while ICD-9 does not. Also, neither ICD-10 nor ICD-10-CM have a code for "Melioidosis Pneumonia," but ICD-10-CM specifies that the code A24.1 Acute and fulminating melioidosis (not shown) should be used

7.4.4.3 International Classification of Primary Care

The World Organization of National Colleges, Academies and Academic Associations of General Practitioners/Family Physicians (WONCA) publishes the International Classification of Primary Care (ICPC) with the WHO, the latest version of which is ICPC-2, published in 1998. ICPC-2 is a classification of some 1,400 diagnostic concepts that are partially mapped into ICD-9 and ICD-10. ICPC-2 contains all 380 concepts of the International Classification of Health Problems in Primary Care (ICHPPC), Third Edition, including reasons for an encounter. ICPC provides seven axes of terms and a structure to combine them to represent clinical encounters. Although the granularity of the terms is generally larger than that of other classification schemes (e.g.,

Respiratory disease w/ major chest operating room procedure, no major complication or comorbidity	75
Respiratory disease w/ major chest operating room procedure, minor complication or comorbidity	76
Respiratory disease w/ other respiratory system operating procedure, no complication or comorbidity	77
Respiratory infection w/ minor complication, age greater than 17	79
Respiratory infection w/ no minor complication, age greater than 17	80
Simple Pneumonia w/ minor complication, age greater than 17	89
Simple Pneumonia w/ no minor complication, age greater than 17	90
Respiratory disease w/ ventilator support	475
Respiratory disease w/ major chest operating room procedure and major complication or comorbidity	538
Respiratory disease, other respiratory system operating procedure and major complication	539
Respiratory infection w/ major complication or comorbidity	540
Respiratory infection w/ secondary diagnosis of bronchopulmonary dysplasia	631
Respiratory infection w/ secondary diagnosis of cystic fibrosis	740
Respiratory infection w/ minor complication, age not greater than 17	770
Respiratory infection w/ no minor complication, age not greater than 17	771
Simple Pneumonia w/ minor complication, age not greater than 17	772
Simple Pneumonia w/ no minor complication, age not greater than 17	773
Respiratory infection w/ primary diagnosis of tuberculosis	798

Fig. 7.5 Diagnosis-related group codes assigned to cases of bacterial pneumonia depending on co-occurring conditions or procedures (mycobacterial disease is not shown here except as a cooccurring condition). "Simple Pneumonia" codes are used when the primary bacterial pneumonia corresponds to ICD-9 code 481, 482.2, 482.3, or 482.9, and when there are only minor or no complications. The remaining ICD-9 bacterial pneumonias (482.0, 482.1, 482.2, 482.4, 482.8, 484, and various other codes such as 003.22) are coded as "Respiratory Disease" or "Respiratory Infection." Cases in which pneumonia is a secondary diagnosis may also be assigned other codes (such as 798), depending on the primary condition

all pneumonias are coded as R81), the ability to represent the interactions of the concepts found in a medical record is much greater through the **post-coordination** of atomic terms. In post-coordination, the coding is accomplished through the use of multiple codes as needed to describe the data. Thus, for example, a case of bacterial pneumonia would be coded in ICPC as a combination of the code R81 and the code for the particular test result that identifies the causative agent. This method is in contrast to the **pre-coordination** approach in which every type of pneumonia is assigned its own code.

7.4.4.4 Current Procedural Terminology

The American Medical Association developed the Current Procedural Terminology (CPT) in 1966 (American Medical Association, updated annually) to provide a precoordinated coding scheme for diagnostic and therapeutic procedures that has since been adopted in the United States for billing and reimbursement. Like the DRG codes, CPT codes specify information that differentiates the codes based on cost. For example, there are different codes for pacemaker insertions, depending on whether the leads are "epicardial, by thoracotomy" (33200), "epicardial, by xiphoid approach" (33201), "transvenous, atrial" (33206), "transvenous, ventricular" (33207), or "transvenous, atrioventricular (AV) sequential" (33208). CPT also provides information about the reasons for a procedure. For example, there are codes for arterial punctures for "withdrawal of blood for diagnosis" (36600), "monitoring" (36620), "infusion therapy" (36640), and "occlusion therapy" (75894). Although limited in scope and depth (despite containing over 8,000 terms), CPT-4 is the most widely accepted nomenclature in the United States for reporting physician procedures and services for federal and private insurance third-party reimbursement.

7.4.4.5 Diagnostic and Statistical Manual of Mental Disorders

The American Psychiatric Association published the Fifth Edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) in May 2013. DSM-5 is the standard classification of mental disorders used by mental health professionals in the U.S. and contains a listing of diagnostic criteria for every psychiatric disorder recognized by the U.S. healthcare system. The previous edition, DSM-IV, was originally published in 1994 and revised in 2000 as DSM-IV-TR. DSM-5 is coordinated with ICD-10.

7.4.4.6 Read Clinical Codes

The Read Clinical Codes comprise a set of terms designed specifically for use in coding electronic medical records. Developed by James Read in the 1980s (Read and Benson 1986; Read 1990), the first version was adopted by the British National Health Service (NHS) in 1990. Version 2.0 was developed to meet the needs of hospitals for cross-mapping their data to ICD-9. Version 3.0 (NHS Centre for Coding and Classification 1994) was developed to support not only medical record summarization but also patient-care applications directly. Whereas previous versions of the Read Codes were organized in a strict hierarchy, Version 3.0 took an important step by allowing terms to have multiple parents in the hierarchy, i.e., the hierarchy became that of a directed acyclic graph. Version 3.1 added the ability to make use of term modifiers through a set of templates for combining terms in specific, controlled ways so that both pre-coordination and post-coordination are used. Finally, the NHS undertook a series of "clinical terms" projects that expanded the content of the Read Codes to ensure that the terms needed by practitioners are represented in the Codes (NHS Centre for Coding and Classification 1994). Clinical Terms version 3 was merged with SNOMED in 2001 to create the first version of SNOMED-CT (see next section).

7.4.4.7 SNOMED Clinical Terms and Its Predecessors

Drawing from the New York Academy of Medicine's Standard Nomenclature of Diseases and Operations (SNDO) (Plunkett 1952; Thompson and Hayden 1961) (New York Academy of Medicine 1961), the College of American Pathologists (CAP) developed the Standard Nomenclature of Pathology (SNOP) as a multiaxial system for describing pathologic findings (College of American Pathologists, 1971) through post-coordination of topographic (anatomic), morphologic, etiologic, and functional terms. SNOP has been used widely in pathology systems in the United States; its successor, the Systematized Nomenclature of Medicine (SNOMED) has evolved beyond an abstracting scheme to become a comprehensive coding system.

Largely the work of Roger Côté and David Rothwell, SNOMED was first published in 1975, was revised as SNOMED II in 1979, and then greatly expanded in 1993 as the Systematized Nomenclature of Human and Veterinary Medicine—SNOMED.

International (Côté and Rothwell 1993). Each of these versions was **multi-axial**; coding of patient information was accomplished through the post-coordination of terms from multiple axes to represent complex terms that did not exist as single codes in SNOMED. In 1996, SNOMED changed from a multi-axial structure to a more logic-based structure called a Reference Terminology (Spackman et al. 1997a, b; Campbell et al. 1998), intended to support more sophisticated data encoding processes and resolve some of the problems with earlier versions of SNOMED (see Fig. 7.6). In 1999, CAP and the NHS announced an agreement to merge their products into a single terminology called SNOMED Clinical Terms (SNOMED-CT) (Spackman 2000), containing terms for over 344,000 concepts (see Fig. 7.7). SNOMED-CT is currently maintained by a not-for-profit association called the International Health Terminology Standards Development Organization (IHTSDO).

Despite the broad coverage of SNOMED-CT, it continues to allow users to create new, ad hoc terms through post-coordination of existing terms. While this increases the expressivity, users must be careful not to be too expressive because there are few rules about how the post-coordination coding should be done, the same expression might end up being represented differently by different coders. For example,

Concept: Bacterial pneumonia
- Concept Status Current
- Fully defined by …
  - Is a
    - Infectious disease of lung
    - Inflammatory disorder of lower respiratory tract
    - Infective pneumonia
    - Inflammation of specific body organs
    - Inflammation of specific body systems
    - Bacterial infectious disease
  - Causative agent:
    - Bacterium
  - Pathological process:
    - Infectious disease
  - Associated morphology:
    - Inflammation
  - Finding site:
    - Lung structure
  - Onset:
    - Subacute onset
    - Acute onset
    - Insidious onset
    - Sudden onset
  - Severity:
    - Severities
  - Episodicity:
    - Episodicities
  - Course:
    - Courses
- Descriptions:
  - Bacterial pneumonia (disorder)
  - Bacterial pneumonia
- Legacy codes:
  - SNOMED: DE-10100
  - CTV3ID: X100H

Fig. 7.6 Description-logic representation of the SNOMED-CT term "Bacterial Pneumonia." The "Is a" attributes define bacterial pneumonia's position in SNOMED-CT's multiple hierarchy, while attributes such as "Causative Agent" and "Finding Site" provide definitional information. Other attributes such as "Onset" and "Severities" indicate ways in which bacterial pneumonia can be postcoordinated with others terms, such as "Acute Onset" or any of the descendants of the term "Severities." "Descriptions" refers to various text strings that serve as names for the term, while "Legacy Codes" provide backward compatibility to SNOMED and Read Clinical Terms

"acute appendicitis" can be coded as a single disease term, as a combination of a modifier ("acute") and a disease term ("appendicitis"), or as a combination of a modifier ("acute"), a morphology term ("inflammation") and a topography term ("vermiform appendix"). Users must therefore be careful when post-coordinating terms, not to recreate a meaning that is satisfied by an

- Pneumonia
  - Bacterial pneumonia
    - Proteus pneumonia
    - Legionella pneumonia
    - Anthrax pneumonia
    - Actinomycotic pneumonia
    - Nocardial pneumonia
    - Meningococcal pneumonia
    - Chlamydial pneumonia
      - Neonatal chlamydial pneumonia
      - Ornithosis
        - Ornithosis with complication
        - Ornithosis with pneumonia
    - Congenital bacterial pneumonia
      - Congenital staphylococcal pneumonia
      - Congenital group A hemolytic streptococcal pneumonia
      - Congenital group B hemolytic streptococcal pneumonia
      - Congenital *Escherichia coli* pneumonia
      - Congenital pseudomonal pneumonia
    - Chlamydial pneumonitis in all species except pig
      - Feline pneumonitis
    - Staphylococcal pneumonia
    - Pulmonary actinobacillosis
    - Pneumonia in Q fever
    - Pneumonia due to Streptococcus
      - Group B streptococcal pneumonia
      - Congenital group A hemolytic streptococcal pneumonia
      - Congenital group B hemolytic streptococcal pneumonia
      - Pneumococcal pneumonia
        - Pneumococcal lobar pneumonia
        - AIDS with pneumococcal pneumonia
    - Pneumonia due to Pseudomonas
      - Congenital pseudomonal pneumonia
    - Pulmonary tularemia
    - Enzootic pneumonia of calves
    - Pneumonia in pertussis
    - AIDS with bacterial pneumonia
    - Enzootic pneumonia of sheep
    - Pneumonia due to *Klebstella pneumoniae*
    - *Hemophilus influenzae* pneumonia
    - Porcine contagious pleuropneumonia
    - Pneumonia due to pleuropneumonia-like organism
    - Secondary bacterial pneumonia
  - Pneumonic plague
    - Primary pneumonic plague
    - Secondary pneumonic plague
  - Salmonella pneumonia
    - Pneumonia in typhoid fever
  - Infective pneumonia
    - Mycoplasma pneumonia
      - Enzootic mycoplasmal pneumonia of swine
  - Achromobacter pneumonia
  - Bovine pneumonic pasteurellosis
  - Corynebacterial pneumonia of foals
  - Pneumonia due to *Escherichia coli*
  - Pneumonia due to *Proteus mirabilis*

Fig. 7.7 Examples of codes in SNOMED-CT, showing some of the hierarchical relationships among bacterial pneumonia terms. Tuberculosis terms and certain terms that are included in SNOMED-CT for compatability with other terminologies are not shown. Note that some terms such as "Congenital group A hemolytic streptococcal pneumonia" appear under multiple parent terms, while other terms, such as "Congenital staphylococcal pneumonia" are not listed under all possible parent terms (e.g., it is under "Congenital pneumonia" but not under "Staphylococcal pneumonia"). Some terms, such as "Pneumonic plague" and "Mycoplasma pneumonia" are not classified under Bacterial Pneumonia, althogh the causative agents in their descriptions ("Yersinia pestis" and "Myocplasma pneumioniae", respectively) are classified under "Bacterium", the causative agent of Bacterial pneumonia

already existing single code. SNOMED-CT's description logic, such as the example in Fig. 7.6, can help guide users when selecting modifiers.

7.4.4.8 Galen

In Europe, a consortium of universities, agencies, and vendors, with funding from the Advanced Informatics in Medicine initiative (AIM), has formed the GALEN project to develop standards for representing coded patient information (Rector et al. 1995). GALEN developed a reference model for medical concepts using a formalism called Structured Meta Knowledge (SMK) and a formal representation language called GALEN Representation and Integration Language (GRAIL). Using GRAIL, terms are defined through relationships to other terms, and grammars are provided to allow combinations of terms into sensible phrases. The reference model is intended to allow representation of patient information in a way that is independent of the language being recorded and of the data model used by an electronic medical record system. The GALEN developers are working closely with CEN TC 251 (see Sect. 7.3.2) to develop the content that will populate the reference model with actual terms. In 2000, an open source foundation called OpenGALEN[4] was developed. Active development of OpenGALEN has since ceased. However, the website is still active and all of the GALEN content and educational materials are available for download and use free of charge.

7.4.4.9 Logical Observations, Identifiers, Names, and Codes

An independent consortium, led by Clement J. McDonald and Stanley M. Huff, has created a naming system for tests and observations. The system is called Logical Observation Identifiers Names and Codes (LOINC).[5] The coding system contains names and codes for laboratory tests, patient measurements, assessment instruments, document and section names, and radiology exams. Figure 7.8 shows some typical fully specified names for common laboratory tests. The standard specifies structured coded semantic information about each test, such as the substance measured and the analytical method used.

7.4.4.10 Nursing Terminologies

Nursing organizations and research teams have been extremely active in the development of standard coding systems for documenting and evaluating nursing care. One review counted a total of 12 separate projects active worldwide (Coenen et al. 2001), including coordination with SNOMED and LOINC. These projects have arisen because general medical terminologies fail to represent the kind of clinical concepts needed in nursing care. For example, the kinds of problems that appear in a physician's problem list (such as "myocardial infarction" and "diabetes mellitus") are relatively well represented in many of the terminologies that we have described, but the kinds of problems that appear in a nurse's assessment (such as "activity intolerance" and "knowledge deficit related to myocardial infarction") are not. Preeminent nursing terminologies include the North American Nursing Diagnosis Association (NANDA) codes, the Nursing Interventions Classification (NIC), the Nursing Outcomes Classification (NOC), the Georgetown Home Health Care Classification (HHCC), and the Omaha System (which covers problems, interventions, and outcomes).

Despite the proliferation of standards for nursing terminologies, gaps remain in the coverage of this domain (Henry and Mead 1997). The International Council of Nurses and the International Medical Informatics Association Nursing Informatics Special Interest Group have worked together to produce the International Classification for Nursing Practice (ICNP®). This system uses a post-coordinated approach for describing nursing diagnoses, actions, and outcomes.

7.4.4.11 Drug Codes

A variety of public and commercial terminologies have been developed to represent terms used for prescribing, dispensing and administering drugs. The WHO Drug Dictionary is an international classification of drugs that provides proprietary drug names used in different countries, as

[4] www.opengalen.org (accessed 4/26/13)

[5] loinc.org (accessed 4/26/13)

Blood glucose	GLUCOSE:MCNC:PT:BLD:QN:
Plasma glucose	GLUCOSE:MCNC:PT:PLAS:QN:
Serum glucose	GLUCOSE:MCNC:PT:SER:QN:
Urine glucose concentration	GLUCOSE:MCNC:PT:UR:QN:
Urine glucose by dip stick	GLUCOSE:MCNC:PT:UR:SQ:TEST STRIP
Glucose tolerance test at 2 hours	GLUCOSE^2H POST 100 G GLUCOSE PO: MCNC:PT:PLAS:QN:
Ionized whole blood calcium	CALCIUM.FREE:SCNC:PT:BLD:QN:
Serum or plasma ionized calcium	CALCIUM.FREE:SCNC:PT:SER/PLAS:QN:
24-hour calcium excretion	CALCIUM.TOTAL:MRAT:24H:UR:QN:
Whole blood total calcium	CALCIUM.TOTAL:SCNC:PT:BLD:QN:
Serum or plasma total calcium	CALCIUM.TOTAL:SCNC:PT:SER/PLAS:QN:
Automated hematocrit	HEMATOCRIT:NFR:PT:BLD:QN: AUTOMATED COUNT
Manual spun hematocrit	HEMATOCRIT:NFR:PT:BLD:QN:SPUN
Urine erythrocyte casts	ERYTHROCYTE CASTS:ACNC:PT:URNS:SQ: MICROSCOPY.LIGHT
Erythrocyte MCHC	ERYTHROCYTE MEAN CORPUSCULAR HEMOGLOBIN CONCENTRATION:MCNC:PT:RBC:QN:AUTOMATED COUNT
Erythrocyte MCH	ERYTHROCYTE MEAN CORPUSCULAR HEMOGLOBIN:MCNC:PT:RBC:QN: AUTOMATED COUNT
Erythrocyte MCV	ERYTHROCYTE MEAN CORPUSCULAR VOLUME:ENTVOL:PT:RBC:QN:AUTOMATED COUNT
Automated Blood RBC	ERYTHROCYTES:NCNC:PT:BLD:QN: AUTOMATED COUNT
Manual blood RBC	ERYTHROCYTES:NCNC:PT:BLD:QN: MANUAL COUNT
ESR by Westergren method	ERYTHROCYTE SEDIMENTATION RATE:VEL:PT:BLD:QN:WESTERGREN
ESR by Wintrobe method	ERYTHROCYTE SEDIMENTATION RATE:VEL:PT:BLD:QN:WINTROBE

Fig. 7.8 Examples of common laboratory test terms as they are encoded in LOINC. The major components of the fully specified name are in separate columns and consist of the analyte, the property (e.g., *Mcnc* mass concentration, *Scnc* substance concentration, *Acnc* arbitrary concentration, *Vfr* volume fraction, *EntMass* entitic mass, *EntVol* entitic volume, *Vel* velocity, and *Ncnc* number concentration), the timing (*Pt* point in time), the system (specimen), and the method (*Ord* ordinal, *Qn* quantitative)

well as all active ingredients and the chemical substances, with Chemical Abstract numbers. Drugs are classified according to the Anatomical-Therapeutic-Chemical (ATC) classification, with cross-references to manufacturers and reference sources. The current dictionary contains 25,000 proprietary drug names, 15,000 single ingredient drugs, 10,000 multiple ingredient drugs, and 7,000 chemical substances. The dictionary now covers drugs from 34 countries and grows at a rate of about 2,000 new entries per year.

The National Drug Codes (NDC), produced by the U.S. Food and Drug Administration (FDA), is applied to all drug packages. It is widely used in the United States, but it is not as comprehensive as the WHO codes. The FDA designates part of the code based on drug manufacturer, and each manufacturer defines the specific codes for their own products. As a result, there is no uniform class hierarchy for the codes, and codes may be reused at the manufacturer's discretion. Due in part to the inadequacies of the NDC codes, pharmacy information systems typically purchase proprietary terminologies from knowledge base vendors. These terminologies map to NDC, but provide additional information about therapeutic classes, allergies, ingredients, and forms.

The need for standards for drug terminologies led to a collaboration between the FDA, the U.S. National Library of Medicine (NLM), the Veterans Administration (VA), and the pharmacy knowledge base vendors that has produced a representational model for drug terms called RxNorm. The NLM provides RxNorm to the

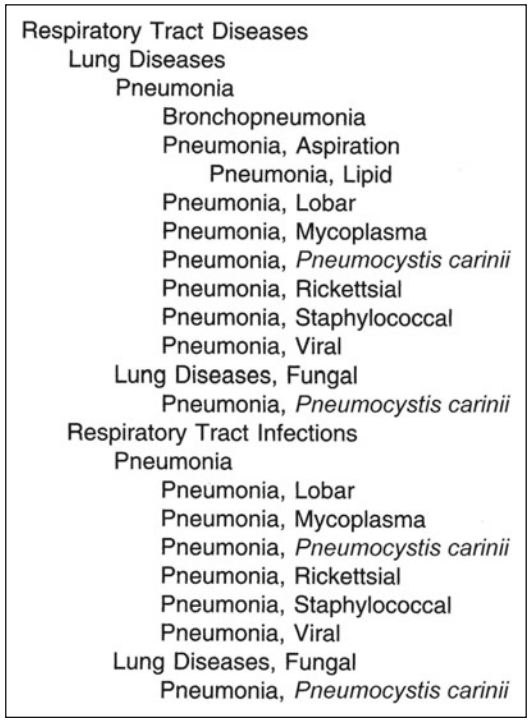

Fig. 7.9 Partial tree structure for the Medical Subject Headings showing pneumonia terms. Note that terms can appear in multiple locations, although they may not always have the same children, implying that they have somewhat different meanings in different contexts. For example, Pneumonia means "lung inflammation" in one context (line 3) and "lung infection" in another (line 16)

public as part of the Unified Medical Language System (UMLS) (see below) to support mapping between NDC codes, the VA's National Drug File (VANDF) and various proprietary drug terminologies (Nelson et al. 2002). RxNorm currently contains 14,000 terms.

7.4.4.12 Medical Subject Headings

The Medical Subject Headings (MeSH), maintained by the NLM (updated annually), is the terminology by which the world medical literature is indexed. MeSH arranges terms in a structure that breaks from the strict hierarchy used by most other coding schemes. Terms are organized into hierarchies and may appear in multiple places in the hierarchy (Fig. 7.9). Although it is not generally used as a direct coding scheme for patient information, it plays a central role in the UMLS.

7.4.4.13 RadLex

RadLex is a terminology produced by the Radiology Society of North America (RSNA). With more than 30,000 terms, RadLex is intended to be a unified language of radiology terms for standardized indexing and retrieval of radiology information resources. RadLex includes the names of anatomic parts, radiology devices, imaging exams and procedure steps performed in radiology. Given the scope of the radiology domain, many RadLex terms overlap with SNOMED-CT, and LOINC.

7.4.4.14 Bioinformatics Terminologies

For the most part, the terminologies discussed above fail to represent the levels of detail needed by biomolecular researchers. This has become a more acute problem with the advent of bioinformatics and the sequencing of organism genomes (see Chap. 24). As in other domains, researchers have been forced to develop their own terminologies. As these researchers have begun to exchange information, they have recognized the need for standard naming conventions as well as standard ways of representing their data with terminologies. Prominent efforts to unify naming systems include the Gene Ontology (GO) (Harris et al. 2004) from the Gene Ontology Consortium and the gene naming database of the HUGO Gene Nomenclature Committee (HGNC). A related resource is the RefSeq database of the National Center for Biotechnology Information (NCBI) which contains identifiers for reference sequences.

7.4.4.15 Unified Medical Language System

In 1986, Donald Lindberg and Betsy Humphreys, at the NLM, began working with several academic centers to identify ways to construct a resource that would bring together and disseminate controlled medical terminologies. An experimental version of the UMLS was first published in 1989 (Humphreys 1990); the UMLS has been updated annually since then. Its principal component is the Metathesaurus, which contains over 8.9 million terms collected from over 160 different sources (including many of those that we

Fig. 7.10 Growth of the UMLS. The UMLS Metathesaurus contains over 8.9 million terms collected from over 160 different sources and attempts to relate synonymous and similar terms from across the different sources into over 2.6 million concepts. The content continues to grow dynamically in response to user needs (Source: U.S. National Library of Medicine)

have discussed), and attempts to relate synonymous and similar terms from across the different sources into over 2.6 million concepts (Fig. 7.10). Figure 7.11 lists the preferred names for many of the pneumonia concepts in the Metathesaurus; Fig. 7.12 shows how like terms are grouped into concepts and are tied to other concepts through semantic relationships.

7.5 Data Interchange Standards

The recognition of the need to interconnect health care applications led to the development and enforcement of **data interchange standards**. The conceptualization stage began in 1980 with discussions among individuals in an organization called the American Association for Medical Systems and Informatics (AAMSI). In 1983, an AAMSI task force was established to pursue those interests in developing standards. The discussions were far ranging in topics and focus. Some members wanted to write standards for everything, including a standard medical terminology, standards for HISs, standards for the computer-based patient record, and standards for data interchange. Citing the need for data interchange between commercial laboratories and health care providers, the task force agreed to focus on data-interchange standards for clinical laboratory data. Early activities were directed mainly toward increasing interest of AAMSI members in working to create health care standards.

The development phase was multifaceted. The AAMSI task force became subcommittee E31.11 of the ASTM and developed and published ASTM standard 1238 for the exchange of clinical-laboratory data. Two other groups—many members of which had participated in the earlier AAMSI task force—were formed to develop standards, each with a slightly different emphasis: HL7 and Institute of Electrical and Electronics Engineering (IEEE) Medical Data Interchange Standard. The American College of Radiology (ACR) joined with the National Electronic Manufacturers Association (NEMA) to develop a standard for the transfer of image data. Two other groups developed related standards independent of the biomedical informatics community: (1) ANSI X12 for the transmission of commonly used business transactions, including health care claims and benefit data, and (2) National Council for Prescription Drug Programs (NCPDP) for the transmission of third-party drug claims. Development was further complicated by the independent creation of standards by several groups in Europe, including EDIFACT United Nations/Electronic Data Interchange For Administration, Commerce and Transport which is the EDI standard developed under the United Nations. The adoption of EDIFACT has been hampered by competing standards that are based upon XML.

7.5.1 General Concepts and Requirements

The purpose of a data-interchange standard is to permit one system, the **sender**, to transmit to another system, the **receiver**, all the data required to accomplish a specific communication, or **transaction set**, in a precise, unambiguous fashion. To complete this task successfully, both systems must know what format and content is being

C0004626: Pneumonia, Bacterial
C0023241: Legionnaires' Disease
C0032286: Pneumonia due to other specified bacteria
C0032308: Pneumonia, Staphylococcal
C0152489: Salmonella pneumonia
C0155858: Other bacterial pneumonia
C0155859: Pneumonia due to *Klebsiella pneumoniae*
C0155860: Pneumonia due to Pseudomonas
C0155862: Pneumonia due to Streptococcus
C0155865: Pneumonia in pertussis
C0155866: Pneumonia in anthrax
C0238380: PNEUMONIA, KLEBSIELLA AND OTHER GRAM NEGATIVE BACILLI
C0238381: PNEUMONIA, TULAREMIC
C0242056: PNEUMONIA, CLASSIC PNEUMOCOCCAL LOBAR
C0242057: PNEUMONIA, FRIEDLAENDER BACILLUS
C0275977: Pneumonia in typhoid fever
C0276026: Hemophilus influenzae pneumonia
C0276039: Pittsburgh pneumonia
C0276071: Achromobacter pneumonia
C0276080: Pneumonia due to *Proteus mirabilis*
C0276089: Pneumonia due to *Escherichia coli*
C0276523: AIDS with bacterial pneumonia
C0276524: AIDS with pneumococcal pneumonia
C0339946: Pneumonia with tularemia
C0339947: Pneumonia with anthrax
C0339952: Secondary bacterial pneumonia
C0339953: Pneumonia due to *Escherichia coli*
C0339954: Pneumonia due to proteus
C0339956: Typhoid pneumonia
C0339957: Meningococcal pneumonia
C0343320: Congenital pneumonia due to staphylococcus
C0343321: Congenital pneumonia due to group A hemolytic streptococcus
C0343322: Congenital pneumonia due to group B hemolytic streptococcus
C0343323: Congenital pneumonia due to *Escherichia coli*
C0343324: Congenital pneumonia due to pseudomonas
C0348678: Pneumonia due to other aerobic Gram-negative bacteria
C0348680: Pneumonia in bacterial diseases classified elsewhere
C0348801: Pneumonia due to streptococcus, group B
C0349495: Congenital bacterial pneumonia
C0349692: Lobar (pneumococcal) pneumonia
C0375322: Pneumococcal pneumonia {*Streptococcus pneumoniae* pneumonia}
C0375323: Pneumonia due to Streptococcus, unspecified
C0375324: Pneumonia due to Streptococcus Group A
C0375326: Pneumonia due to other Streptococcus
C0375327: Pneumonia due to anaerobes
C0375328: Pneumonia due to *Escherichia coli*
C0375329: Pneumonia due to other Gram-negative bacteria
C0375330: Bacterial pneumonia, unspecified

Fig. 7.11 Some of the bacterial pneumonia concepts in the Unified Medical Language System Metathesaurus

sent and must understand the words or terminology, as well as the delivery mode. When you order merchandise, you fill out a form that includes your name and address, desired items, quantities, colors, sizes, and so on. You might put the order form in an envelope and mail it to the supplier at a specified address. There are standard requirements, such as where and how to write the receiver's (supplier's) address, your (the sender's) address, and the payment for delivery (the postage stamp). The receiver must have a mailroom, a post office box, or a mailbox to receive the mail.

A communications model, called the Open Systems Interconnection (OSI) reference model (ISO 7498–1), has been defined by the ISO (see Chap. 5 and the discussion of software for network communications). It describes seven levels

Bacterial pneumonia
Source: CSP93/PT/2596-5280; DOR27/DT/U000523; ICD91/PT/482.9; ICD91/IT/482.9
Parent: Bacterial Infections; Pneumonia; Influenza with Pneumonia
Child: Pneumonia, Mycoplasma
Narrower: Pneumonia, Lobar; Pneumonia, Rickettsial; Pneumonia, Staphylococcal; Pneumonia due to *Klebsiella pneumoniae;* Pneumonia due to Pseudomonas; Pneumonia due to *Hemophilus influenzae*
Other: *Klebsiella pneumoniae, Streptococcus pneumoniae*

Pneumonia, Lobar
Source: ICD91/IT/481; MSH94/PM/D011018; MSH94/MH/D011018; SNM2/RT/M-40000; ICD91/PT/481; SNM2/PT/D-0164; DXP92/PT/U000473; MSH94/EP/D011018; INS94/MH/D011018;INS94/SY/D011018
Synonym: Pneumonia, diplococcal
Parent: Bacterial Infections; Influenza with Pneumonia
Broader: Bacterial Pneumonia; Inflammation
Other: *Streptococcus pneumoniae*
Semantic: inverse-is-a: *Pneumonia*
has-result: *Pneumococcal Infections*

Pneumonia, Staphylococcal
Source: ICD91/PT/482.4; ICD91/IT/482.4; MSH94/MH/D011023; MSH94/PM/D011023; MSH94/EP/D011023; SNM2/PT/D-017X; INS94/MH/D011023; INS94/SY/D011023
Parent: Bacterial Infections; Influenza with Pneumonia
Broader: Bacterial Pneumonia
Semantic inverse-is-a: Pneumonia; Staphylococcal Infections

Pneumonia, Streptococcal
Source: ICD91/IT/482.3
Other: *Streptococcus pneumoniae*

Pneumonia due to Streptococcus
Source: ICD91/PT/482.3
ATX: Pneumonia AND Streptococcal Infections AND NOT Pneumonia, Lobar
Parent: Influenza with Pneumonia

Pneumonia in Anthrax
Source: ICD91/PT/484.5; ICD91/IT/022.1; ICD91/IT/484.5
Parent: Influenza with Pneumonia
Broader: Pneumonia in other infectious diseases classified elsewhere
Other: Pneumonia, Anthrax

Pneumonia, Anthrax
Source: ICD91/IT/022.1; ICD91/IT/484.5
Other: Pneumonia in Anthrax

Fig. 7.12 Some of the information available in the Unified Medical Language System about selected pneumonia concepts. Concept's preferred names are shown in italics. Sources are identifiers for the concept in other terminologies. Synonyms are names other than the preferred name. ATX is an associated Medical Subject Heading expression that can be used for Medline searches. The remaining fields (Parent, Child, Broader, Narrower, Other, and Semantic) show relationships among concepts in the Metathesaurus. Note that concepts may or may not have hierarchical relations to each other through Parent–Child, Broader–Narrower, and Semantic (is-a and inverseis-a) relations. Note also that Pneumonia, Streptococcal and Pneumonia due to Streptococcus are treated as separate concepts, as are Pneumonia in Anthrax and Pneumonia, Anthrax

of requirements or specifications for a communications exchange: physical, data link, network, transport, session, presentation, and application (Stallings 1987a; Tanenbaum 1987; Rose 1989). Level 7, the application level, deals primarily with the semantics or data-content specification of the transaction set or message. For the data-interchange standard, HL7 requires the definition of all the data elements to be sent in response to a specific task, such as the admission of a patient to a hospital. In many cases, the data content requires a specific terminology that can be understood by both sender and receiver. For example, if a physician orders a laboratory test that is to be processed by a commercial laboratory, the ordering system must ensure that the name of the test on the order is the same as the name that the laboratory uses. When a panel of tests is ordered, both systems must share a common understanding of the panel composition. This terminology understanding is best ensured through use of a terminology table that contains both the test name and a unique code. Unfortunately, several code sets exist for each data group, and none are complete. An immediate challenge to the medical-informatics community is to generate

one complete set. In other cases, the terminology requires a definition of the domain of the set, such as what are the possible answers to the data parameter "ethnic origin."

The sixth level, presentation, deals with what the syntax of the message is, or how the data are formatted. There are both similarities and differences at this level across the various standards bodies. Two philosophies are used for defining syntax: one proposes a position-dependent format; the other uses a tagged-field format. In the position-dependent format, the data content is specified and defined by position. For example, the sixth field, delimited by "|," is the gender of the patient and contains an M, F, or U or is empty. A tagged-field representation is "SEX = M."

The remaining OSI levels—session, transport, network, data link, and physical—govern the communications and networking protocols and the physical connections made to the system. Obviously, some understanding at these lower levels is necessary before a linkage between two systems can be successful. Increasingly, standards groups are defining scenarios and rules for using various protocols at these levels, such as TCP/IP (see Chap. 5). Much of the labor in making existing standards work lies in these lower levels.

Typically, a transaction set or message is defined for a particular event, called a **trigger event**. This trigger event, such as a hospital admission, then initiates an exchange of messages. The message is composed of several data segments; each data segment consists of one or more data fields. Data fields, in turn, consist of data elements that may be one of several data types. The message must identify the sender and the receiver, the message number for subsequent referral, the type of message, special rules or flags, and any security requirements. If a patient is involved, a data segment must identify the patient, the circumstances of the encounter, and additional information as required. A reply from the receiving system to the sending system is mandatory in most circumstances and completes the communications set.

It is important to understand that the sole purpose of the data-interchange standard is to allow data to be sent from the sending system to the receiving system; the standard is not intended to constrain the application system that uses those data. Application independence permits the data-interchange standard to be used for a wide variety of applications. However, the standard must ensure that it accommodates all data elements required by the complete application set.

7.5.2 Specific Data Interchange Standards

As health care increasingly depends on the connectivity within an institution, an enterprise, an integrated delivery system, a geographic system, or even a national integrated system, the ability to interchange data in a seamless manner becomes critically important. The economic benefits of data-interchange standards are immediate and obvious. Consequently, it is in this area of health care standards that most effort has been expended. All of the SDOs in health care have some development activity in data-interchange standards.

In the following sections we summarize many of the current standards for data-interchange. Examples are provided to give you a sense of the technical issues that arise in defining a data-exchange standard, but details are beyond the scope of this book. For more information, consult the primary resources or the Web sites for the relevant organizations.

7.5.2.1 Digital Imaging and Communications in Medicine (DICOM) Standards

With the introduction of computed tomography and other digital diagnostic imaging modalities, people needed a standard method for transferring images and associated information between devices, manufactured by different vendors, that display a variety of digital image formats. ACR formed a relationship with the NEMA in 1983 to develop such a standard for exchanging radiographic images, creating a unique professional/vendor group. The purposes of the ACR/NEMA standard were to promote a generic digital-image communication format, to facilitate the

development and expansion of picture-archiving and communication systems (PACSs; see Chap. 18), to allow the creation of diagnostic databases for remote access and to enhance the ability to integrate new equipment with existing systems. Later the group became an international organization with ACR becoming just a member organization. NEMA still manages the organization in the United States.

Version 1.0 of the DICOM standard, published in 1985, specified a hardware interface, a data dictionary, and a set of commands. This standard supported only point-to-point communications. Version 2.0, published in 1988, introduced a message structure that consisted of a command segment for display devices, a new hierarchy scheme to identify an image, and a data segment for increased specificity in the description of an image (e.g., the details of how the image was made and of the settings).

In the DICOM standard, individual units of information, called data elements, are organized within the data dictionary into related groups. Groups and elements are numbered. Each individual data element, as contained within a message, consists of its group-element tag, its length, and its value. Groups include command, identifying, patient, acquisition, relationship, image presentation, text, overlay, and pixel data.

The latest version of DICOM is Version 3.0, which incorporates an object-oriented data model and adds support for ISO standard communications. DICOM provides full networking capability and specifies levels of conformance. The standard itself is structured as a nine-part document to accommodate evolution of the standard. In addition, DICOM introduces explicit information objects for images, graphics, and text reports; introduces service classes to specify well-defined operations across the network, and specifies an established technique for identifying uniquely any information object. DICOM also specifies image-related management information exchange, with the potential to interface to HISs and radiology information systems. An updated Version 3.0 is published annually.

The general syntax used by DICOM in representing data elements includes a data tag, a data length specification, and the data value. That syntax is preserved over a hierarchical nested data structure of items, elements, and groups. Data elements are defined in a data dictionary and are organized into groups. A data set consists of the structured set of attributes or data elements and the values related to an information object. Data-set types include images, graphics, and text

The protocol architecture for DICOM Version 3.0 is shown in Fig. 7.13, which illustrates the communication services for a point-to-point environment and for a networked environment, identifies the communication services and the upper-level protocols necessary to support communication between DICOM Application Entities. The upper-layer service supports the use of a fully conformant stack of OSI protocols to achieve effective communication. It supports a wide variety of international standards-based network technologies using a choice of physical networks such as Ethernet, FDDI, ISDN, X.25, dedicated digital circuits, and other local area network (LAN) and wide area network (WAN) technologies. In addition, the same upper-layer service can be used in conjunction with TCP/IP transport protocols. DICOM is now producing a number of standards including structured reports and Web access to, and presentation of, DICOM persistent objects.

7.5.2.2 ASTM International Standards

In 1984, the first ASTM health care data-interchange standard was published: E1238, Standard Specification for Transferring Clinical Observations Between Independent Systems. This standard is used in large commercial and reference clinical laboratories in the United States and has been adopted by a consortium of French laboratory system vendors who serve 95 % of the laboratory volume in France. The ASTM E1238 standard is message based; it uses position-defined syntax and is similar to the HL7 standard (see next section). An example of the ASTM 1238 standard describing a message transmitted between a clinic and a commercial clinical laboratory is shown in Fig. 7.14. Related data-interchange standards include E1467 (from Subcommittee E31.16), Specification for Transferring Digital

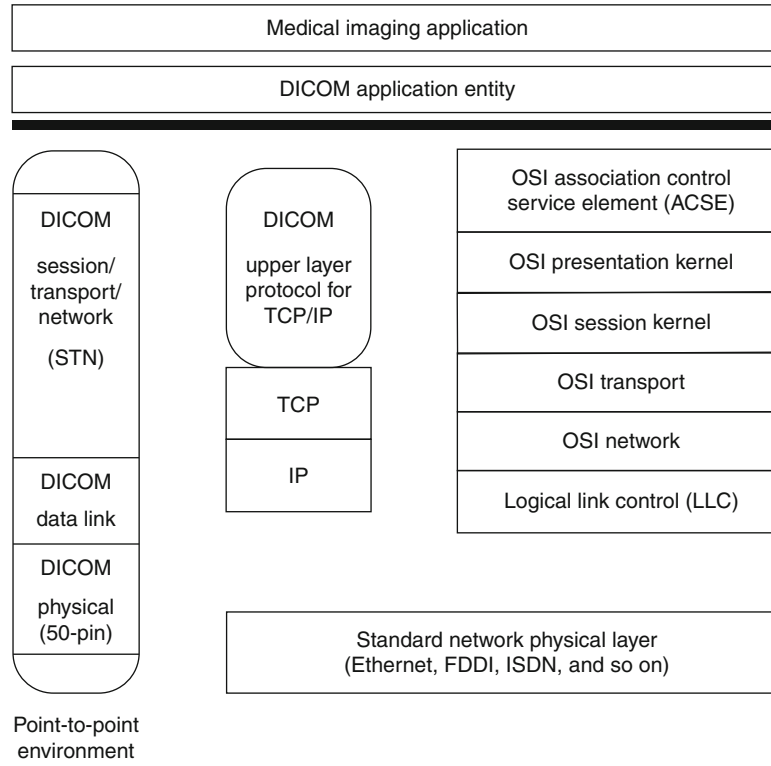

Fig. 7.13 DICOM communications-protocol architecture illustrating the different approaches to dealing with the OSI reference model communication levels

```
H|~^\&|95243|HAMMO001|COMMUNITY AND FAMILY MEDICINE|BOX 2914^DUKE
UNIVERSITY MEDICAL CENTER^DURHAM^NC|919-684-6721||SMITHKLINE
CLINICAL LABS|TEST MESSAGE|D|2|199401170932<cr>

P|1|999-99-9999|||GUNCH^MODINE^SUE||19430704|F|
RT 1, BOX 97^ZIRCONIA^NC^27401||704-982-
1234||DOCTOR^PRIMARY^A^^DR.<cr>

OBR|1|101||80018^CHEM 18|R||||N|||||M D&PRIMARY&A&DR.<cr>

OBR|2|102||85025^AUTO CBC|R||||N||||| MD&PRIMARY&A&DR.
```

Fig. 7.14 An example of a message in the ASTM 1238 format. The message consists of the header segment, H, the patient segment, P, and general order segments, OBR. Primary delimiters are the vertical bars (|); secondary delimiters are the carets (^). Note the similarities of this message to the HL7 message in Fig. 7.3

Neurophysiological Data Between Independent Computer Systems. Another important ASTM standard is E1460, Defining and Sharing Modular Health Knowledge Bases (Arden Syntax for Medical Logic Modules; see Chap. 20). In 1998, ownership of the Arden Syntax was transferred to HL7, where it will be developed by the Arden Syntax and Clinical Decision Support Technical Committee.

In 2005, ASTM developed an electronic format for a paper-based record, initially created by the Massachusetts Medical Society to improve data exchange when transferring patients. This standard became known as the Continuity of Care Record (CCR), and embodies a core data set, represented in XML. The published format of the CCR was jointly developed by ASTM International, the Massachusetts Medical Society[1] (MMS), the Health Care Information and Management Systems Society (HIMSS), the American Academy of Family Physicians (AAFP), the American Academy of Pediatrics,

and other health informatics vendors. The CCR represents a summary of the patient record, rather than clinical encounter. The CCR is both machine- and human-readable, which accelerated its adoption. In the delineation of the Meaningful Use requirements from the US Department of Health and Human Services, the CCR data set provided a constraint on the HL7 Clinical Document Architecture (CDA) specification (see below), now referred to as the Continuity of Care Document (CCD).

7.5.2.3 Health Level 7 Standards

The initial HL7 (see Sect. 7.3.2.3) standard was built on existing production protocols—particularly ASTM 1238. The HL7 standard is message based and uses an event trigger model that causes the sending system to transmit a specified message to the receiving unit, with a subsequent response by the receiving unit. Messages are defined for various trigger events. Version 1.0 was published in September 1987 and served mainly to define the scope and format of standards. Version 2.0, September 1988, was the basis for several data-interchange demonstrations involving more than ten vendors. Version 2.1, June 1990, was widely implemented in the United States and abroad. In 1991, HL7 became a charter member of ANSI; on June 12, 1994, it became an ANSI-accredited Standard Developers Organization (SDO). Version 2.2 was published in December 1994 and on February 8, 1996, it was approved by ANSI as the first health care data-interchange American National Standard. Version 2.3, March 1997, considerably expanded the scope by providing standards for the interchange of data relating to patient administration (admission, discharge, transfer, and outpatient registration), patient accounting (billing), order entry, clinical-observation data, medical information management, patient and resource scheduling, patient-referral messages, patient-care messages that support communication for problem-oriented records, adverse-event reporting, immunization reporting, and clinical trials, as well as a generalized interface for synchronizing common reference files.

Version 2.4, which became an ANSI standard in October 2000 introduced conformance query profiles and added messages for laboratory automation, application management, and personnel management. ANSI recently approved the HL7 Version 2.0 Extensible Markup Language (XML) Encoding Syntax. The XML capability of HL7 v2.xml makes messages Web-enabled. Version 2.5, which is more consistent and supports more functionality than any other previous version, became an ANSI standard in 2003. Figure 7.15 illustrates the exchange that occurs when a patient is transferred from the operating room (which uses a system called DHIS) to the surgical intensive-care unit (which uses a system called TMR). Note the similarity between these messages and the ASTM example.

Version 3.0 of the standard (currently in the ballot process) is object oriented and based on a **Reference Information Model** (**RIM**) being developed by HL7. The RIM has evolved from a number of commercial and academic health care data models, and it accommodates the data elements defined in the current Version 2.x HL7 standard.

The RIM is a collection of subject areas, scenarios, classes, attributes, use cases, actors, trigger events, interactions, and so on that depict the information needed to specify HL7 messages. In this sense it is more than a data-interchange standard, seeking to merge standards notions that include terminology and representation as well as data exchange. The stated purpose of the RIM is to provide a model for the creation of message specifications and messages for HL7. The RIM was approved as an ANSI standard in 2003, and has been introduced as an ISO standard. HL7 has also introduced a V3 suite of standards including V3 Abstract Data Types; Clinical Data Architecture, Release 2; and Context Management Standard (CCOW). In some health information system architectures, notably those of UK and Canada, there has been a successful blending of messaging specifications, incorporating those from both V2 and V3.

Since its initial development in 2001, the Clinical Document Architecture (CDA) standard has become globally adopted for a broad range of use. Now an ISO standard and advanced to Release 2, CDA is a document mark-up standard

```
MSH|^~&\DHIS|OR|TMR|SICU|199212071425|password|ADT|16603529|P|2.1<cr>

EVN|A02|199212071425||<cr>

PID|||Z99999^5^M11||GUNCH^MODINE^SUE|RILEY|19430704 |F||C|RT. 1, BOX
97^ZIRCONIA^NC^27401 |HEND|(704)982-1234|(704)983-1822||S|C||245-33-
9999<cr>

PV1|1|I|N22^2204|||OR^03|0940^DOCTOR^HOSPITAL^A||| SUR|||||A3<cr>

OBR|7|||93000^EKG REPORT|R|199401111000|199401111330|||RMT||||19940111
11330|?|P030|||||199401120930|||||88-126666|A111|VIRANYI^ANDREW<cr>

OBX|1|ST|93000.1^VENTRICULAR RATE(EKG)||91|/MIN|60-100<cr>

OBX|2|ST|93000.2^ATRIAL RATE(EKG)||150|/MIN|60-100<cr>

. . .

OBX|8|ST|93000&IMP^EKG DIAGNOSIS|1|^ATRIAL FIBRILATION<cr>
```

Fig. 7.15 An example of an HL7 ADT transaction message. This message includes the Message Heading segment, the EVN trigger definition segment, the PID patient-identification segment, the PV1 patient-visit segment, the OBR general-order segment, and several OBX results segments

for the structure and semantics of an exchanged "clinical document." CDA is built upon the RIM and relies upon reusable templates for its ease of implementation. A CDA document is a defined and complete information object that can exist outside of a message and can include text, images, sounds, and other multimedia content. CDA supports the following features: persistence, stewardship, potential for authentication, context, wholeness, and human-readability. In the US, CDA is one of the core components of data exchange for Meaningful Use. The competing implementation processes for CCD profile development were successfully harmonized into a broadly adopted *Consolidated* Continuity of Care Document (*C*CCD).

In order to ease the path to implementation of CDA, HL7 has developed a more narrowly defined specification called greenCDA, which limits the requirements of the RIM, provides greater ease of template composition, and consumes much less bandwidth for transmission. An additional effort to promote CDA adoption was achieved with the release of the CDA Trifolia repository, which, in addition to offering a library of templates, includes tooling for template modification as well as a template-authoring language. This has enabled the adoption of native CDA for exchange of laboratory data, clinical summaries, and electronic prescriptions and well as for clinical decision support.

Fast Health care Interoperability Resources (FHIR, pronounced "Fire") is a new highly innovative approach to standards development, first introduced by HL7in 2011. FHIR was created in order to overcome the complexity of development based upon the HL7 Reference Information Model (RIM), without losing the successful interoperability that model-driven data interchange demands. At the same time, FHIR delivers greater ease of implementation than other high-level development processes. It is designed to be compatible with legacy systems that conform to V2 and/or V3 messaging, and it supports system-development utilizing broadly deployed Clinical Document Architecture (CDA) platforms and ubiquitous templated CDA implementations, such as *Consolidated CDA*.

Although FHIR is built upon more than a decade of the development and refinement of the RIM, FHIR utilizes unique methodologies, artifacts, tooling, and publishing approach. While FHIR is based upon the RIM, it does not require implementers to know the RIM or know the modeling language upon which it was built. FHIR defines a limited set of data elements (or *resources*) as XML objects, but provides extension mechanisms for creating any elements which are incomplete or missing. The resulting structures are native XLM objects which do not require knowledge of the

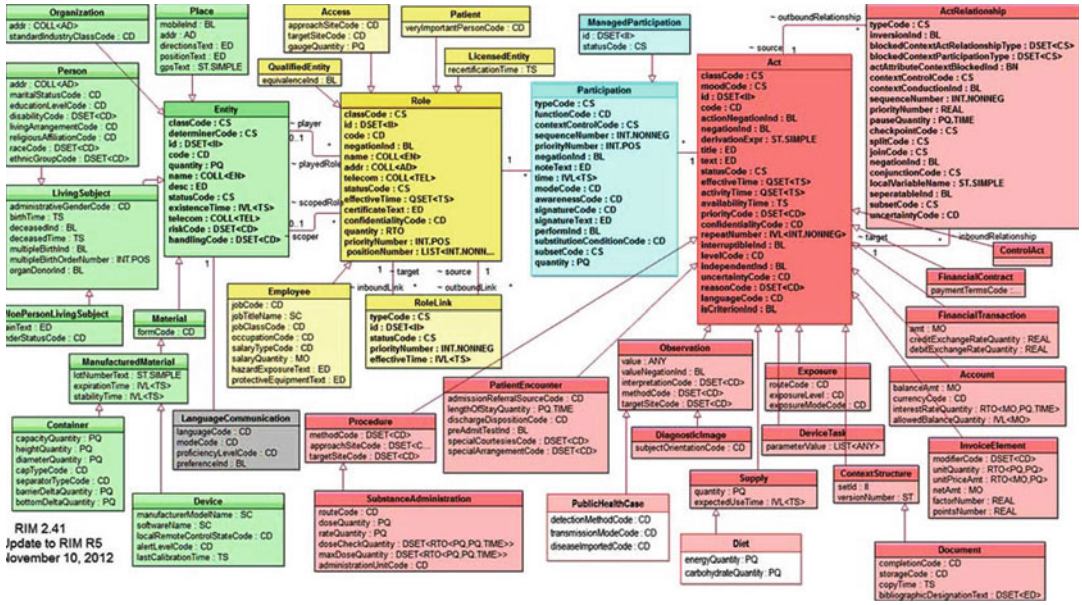

Fig. 7.16 Current version of the HL7 Reference Information Model. Most newer standards from HL7 are based on this model. The model is under constant revision

RIM abstraction in order to be implemented. Fundamentally, each clinical concept is created as a single resource, which need not change over time. The resources remain as the smallest unit of abstraction, and the creation of each resource is based upon *RESTful* design principles.

Inherently, development can precede around a *services* (SOA) model, which will support cloud-based applications. While a RESTful framework is enabled, it is not required. In addition, a well-defined ontology persists in the background, but knowledge of the vocabulary is not necessary for implementation. Fundamental to FHIR, all resources, as well as all resource attributes have a free-text expression, an encoded expression or both. Thus, FHIR supports a human-readable format, which is so valuable to the implementations supported by CDA.

Finally, FHIR is built with new *datatypes*, conformant with the familiar ISO 21090 format. As such, these datatypes are far simpler to use, with much of the complexity captured in the *extensions*. This allows mapping to other models, including those developed using *archetypes*, upon which the CEN format for electronic medical records is predicated. This allows an inherently much smaller library of resources, all mapped to the HL7 RIM, and which can be maintained in perpetuity. FHIR developers have estimated that fewer than 150 such resources will define all of health care. Other concepts can be described as extensions.

As it is now envisioned, legacy systems will not map their interfaces to FHIR. Instead, FHIR will most likely serve in those environments or applications in which classical V2 messaging and/or CDA do not currently exist. This provides a unique opportunity for creation of both new applications in mature computing environments and for low and medium resource countries without legacy implementations. Nonetheless, migrations from V2 or V3 environments to FHIR implementations are achievable through native tooling.

Most often, HL7 is recognized for its messaging standards, but there is a large contribution to technical specifications that support the development and implementation of these messaging standards. These standards are also based on the RIM. A copy of the current version of the RIM is shown in Fig. 7.16.

Other HL7 standards that may be of interest the reader include:

- Clinical Decision Support (CDS) Standards
  - Arden Syntax – Begun at Columbia University, migrated to ASTM, and currently resides in HL7; syntax for CDS
  - GELLO – common expression language for clinical guidelines
- Functional Requirements for Electronic Health Records
  - Electronic Health Record – System Functional Model (EHR-S FM)
  - Personal Health Record (PHR FM)
- Functional Profiles – defines function and conformances
  - Behavioral Health
  - Child Health
  - Clinical Research
  - Records Management and Evidentiary Support
  - Long Term Care
  - Vital Records
  - Pharmacist/Pharmacy Provider
  - Public Health
- Domain Analysis Models – an informative document that defines work and data flow; actors and entities, and data elements
  - Immunization
  - Cardiology; Acute Coronary Syndrome
  - Clinical Trials Registration and Results
  - Tuberculosis Surveillance, Diagnosis, Treatment and Research
  - Virtual Medical Record for Clinical Decision Support
  - Vital Records
- Implementation Guides
  - Orders and Observations Ambulatory Lab Result (v2.5.1)
  - CDA R2: Quality Reporting Document Architecture
  - S&I Framework Lab Results Interface (v2.5.1)
  - Clinical Genomics; Fully LOINC-Qualified Genetic Variation Model
  - CDA R2: IHE Health Story Consolidation
  - URL-Based Implementations of Context-Aware Information Retrieval (Infobutton)
  - Virtual Medical Record for Clinical Decision Support for GELLO
  - Electronic Laboratory Reporting to Public Health
  - Interoperable Laboratory Result Reporting to EHR
  - Infobutton Service Oriented Architecture
  - Continuity of Care Document
  - Laboratory Test Compendium Framework
  - Blood Bank Donation Services
  - Emergency Medical Services Run Report
  - CDA R2: Consent Directives
  - Unique Object Identifiers
  - CDA R2: Level 1 and 2 Care Record Summary
  - CDA R2: Personal Health care Monitoring Report
  - CDA R2: History and Physical Notes
  - CDA R2: Patient Assessments
  - CDA R2: Procedure Note
  - CDA R2: Unstructured Documents
  - Imaging Integration; Basic Imaging Reports in CDA and DICOM
  - CDA R2: Neonatal Care Reports
  - CDA R2: Care Record Summary Discharge Summary
  - CDA R2: Consult Notes
  - CDA R2: Operative Notes
  - CDA R@: Plan-to-plan Personal Health Record Data Transfer
  - CDA R2: greenCDA Modules for CCD
  - HL7-NCPDP Prescribing Coordination Mapping Document
  - CDA R2: Health care Associated Infection
  - CDA R2: Public Health Case Reports
  - Vital Records Death Reporting
  - Structured Product Labeling
  - Annotated ECG
  - Drug Stability Reporting
  - Regulated Product Submission
- Genomics
  - Family History/Pedigree
- Regulatory Standards
  - Structured Product Labeling
  - Individual Case Safety Report (ICSR)
- Common Terminology Services
- Claims Attachment
- Data Types
- HL7 Vocabulary Tables

7.5.2.4 Institute of Electrical and Electronics Engineers Standards

IEEE is an international organization that is a member of both ANSI and ISO. Through IEEE, many of the world's standards in telecommunications, electronics, electrical applications, and computers have been developed. There were two major IEEE standards projects in health care.

IEEE 1073, Standard for Medical Device Communications, has produced a family of documents that defines the entire seven-layer communications requirements for the **Medical Information Bus** (MIB). The MIB is a robust, reliable communication service designed for bedside devices in the intensive care unit, operating room, and emergency room (see Chap. 19 for further discussion of the MIB in patient-monitoring settings). These standards have been harmonized with work in CEN, and the results are being released as ISO standards. IEEE and HL7 have collaborated on several key standards, including those for mobile medical devices.

7.5.2.5 National Council for Prescription Drug Programs (NCPDP) Standards

NCPDP is an ANSI-accredited SDO and is a trade organization. Its mission is to create and promote data-interchange standards for the pharmacy services sector of the health care industry and to provide information and resources that educate the industry. Currently, NCPDP has developed three ANSI-approved standards: a telecommunication standard (Version 3.2 and Version 7.0), a SCRIPT standard (Version 5.0), and a manufacturer rebate standard (Version 3.01). The telecommunication standard provides a standard format for the electronic submission of third-party drug claims. The standard was developed to accommodate the eligibility verification process at the point of sale and to provide a consistent format for electronic claims processing. Primarily pharmacy providers, insurance carriers, third-party administrators, and other responsible parties use the standard. This standard addresses the data format and content, transmission protocol, and other appropriate telecommunication requirements. Version 5.1 (September, 1999) of this standard is one of the transactions standards required for use by HIPAA.

Version 1.0, released in 1988, used formats that were limited to only fixed fields.. Version 2.0 added only typographic corrections to the Version 1.0 standard. The major thrust of the changes in Versions 3.0 and 3.1, in 1989, was the change from fixed-field transactions to a hybrid or variable format in which the fields can be tailored to the required content of the message. The current release is Version 3.2 (February, 1992). It introduces the fixed-length Recommended Transaction Data Sets (RTDS), which define three different message types, and a separate Data Dictionary format. The Data Dictionary defines permissible values and default values for fields contained in the specification. An online, real-time version was developed in 1996.

The standard uses defined separator characters at a group and a field level. The telecommunications specifications for sending two or more prescriptions include three required sections (Transaction Header; Group Separator, First-Claim Information; and Group Separator, Second-Claim Information [R]) and three optional sections (Header Information, First-Claim Information, and Second-Claim Information [O]). The NCPDP communication standard is used in more than 60 % of the nation's total prescription volume.

The SCRIPT Standard and Implementation Guide was developed for transmitting prescription information electronically between prescribers and providers. The standard, which adheres to EDIFACT syntax requirements and utilizes ASC X12 data types where possible, addresses the electronic transmission of new prescriptions, prescription refill requests, prescription fill status notifications, and cancellation notifications.

7.5.2.6 Accredited Standards Committee X12 (ASC X12) Standards

ASC X12, an independent organization accredited by ANSI, has developed message standards for purchase-order data, invoice data, and other commonly used business documents. The

subcommittee X12N has developed a group of standards related to providing claim, benefits, and claim payment or advice. The specific standards that strongly relate to the health care industry are shown in Table 7.3.

The X12 standards define commonly used business transactions in a formal, structured manner called transaction sets. A transaction set is composed of a transaction-set header control segment, one or more data segments, and a transaction-set trailer control segment. Each segment is composed of a unique segment ID; one or more logically related simple data elements or composite data structures, each preceded by a data element separator; and a segment terminator. Data segments are defined in a data segment directory; data elements are defined in a data element directory; composite data structures are defined in a composite data structure directory; control segments and the binary segment are defined in a data segment directory.

A sample 835 Interchange Document is shown in Fig. 7.17. This standard is similar to ASTM and HL7 in that it uses labeled segments with positionally defined components.

There are several additional organizations that either create standards related to health care or have influence on the creation of standards.

7.5.2.7 American Dental Association Standards

In 1983, the American Dental Association (ADA) committee MD 156 became an ANSI-accredited committee responsible for all specifications for dental materials, instruments, and equipment. In 1992, a Task Group of the ASC MD 156 was established to initiate the development of technical reports, guidelines, and standards on electronic technologies used in dental practice. These include digital radiography, digital intraoral video cameras, digital voice-text-image transfer, periodontal probing devices, and CAD/CAM. Proposed standards include Digital Image Capture in Dentistry,

Table 7.3 ANSI X12N standards

Code	Title	Purpose
148	First Report of Injury, Illness or Incident	Facilitates the first report of an injury, incident, or illness
270	Health-Care Eligibility/Benefit Inquiry	Provide for the exchange of eligibility information and for response to individuals in a health care plan
271	Health-Care Eligibility/Benefit Information	
275	Patient Information	Supports the exchange of demographic, clinical, and other patient information to support administrative reimbursement processing as it relates to the submission of health-care claims for both health-care products and services reports the status of a submitted claim
276	Health-Care Claim Status Request	Queries the status of a submitted claim and reports the status of a submitted claim
277	Health-Care Claim Status Notification	
278	Health-Care Service Review Information	Provides referral certification and authorization
811	Consolidated Service Invoice/Statement	Facilitate health-plan premium billing and payment
820	Payment Order/Remittance Advice	
IHCLME Interactive Health-Care Claim/Encounter		Supports administrative reimbursement processing as it relates to the submission of health-care claims for both health-care products and services in an interactive environment
IHCE/BI Interactive Health-Care Eligibility/ Benefit Inquiry		Provide for the exchange of eligibility information and for response to individuals within a health plan
IHCE/BR Interactive Health-Care Eligibility/ Benefit Response		

```
ST*835*0001<n/l>
BPR*X*3685*C*ACH*CTX*01*122000065*DA*296006596*IDNUMBER*
SUPPLECODE*01*134999883*DA*867869899*940116<n/l>
TRN*1*45166*IDNUMBER<n/l>
DTM*009*940104<n/l>
N1*PR*HEALTHY INSURANCE COMPANY<n/l>
N3*1002 WEST MAIN STREET<n/l>
N4*DURHAM*NC*27001<n/l>
N1*PE*DUKE MEDICAL CENTER<n/l>
N3*2001 ERWIN ROAD<n/l>
N4*DURHAM*NC*27710<n/l>
CLP*078189203*1*6530*4895*CIN<n/l>
CAS*PR*1*150<n/l>
CAS*PR*2*550<n/l>
NM1*15*IAM*A*PATIENT<n/l>
REF*1K*942238493<n/l>
DTM*232*940101<n/l>
DTM*233*940131<n/l>
SE*22*0001<n/l>
```

Fig. 7.17 An example of ANSI X12 Interchange Document (Standard 835). This message is derived from a batch process, business-document orientation to a data-interchange model. The example does not include the control header or the functional-group header. The first line identifies the segment as a transaction-set header (ST). The last line is the transaction-set trailer (SE).The leading alphanumeric characters are tags that identify data content. For example, DTM is a date/time reference; N3 is address information; and BPR is the beginning segment for payment order/remittance advice

Infection Control in Dental Informatics, Digital Data Formats for Dentistry, Construction and Safety for Dental Informatics, Periodontal Probe Standard Interface, Computer Oral Health Record, and Specification for the Structure and Content of electronic medical record integration.

7.5.2.8 Uniform Code Council Standards

The Uniform Code Council (UCC) is an ANSI-approved organization that defines the universal product code. Standards include specifications for the printing of machine-readable representations (bar codes).

7.5.2.9 Health Industry Business Communications Council Standards

The Health Industry Business Communications Council (HIBCC) has developed the Health Industry Bar Code (HIBC) Standard, composed of two parts. The HIBC Supplier Labeling Standard describes the data structures and bar code symbols for bar coding of health care products. The HIBCC Provider Applications Standard describes data structures and bar code symbols for bar coding of identification data in a health care provider setting. HIBCC also issues and maintains Labeler Identification Codes that identify individual manufacturers. The HIBCC administers the Health Industry Number System, which provides a unique identifier number and location information for every health care facility and provider in the United States The HIBCC also administers the Universal Product Number Repository, which identifies specific products and is recognized internationally.

7.5.2.10 The Electronic Data Interchange for Administration, Commerce, and Transport Standard

The EDI for Administration, Commerce, and Transport (EDIFACT) is a set of international standards, projects, and guidelines for the electronic interchange of structured data

related to trade in goods and services between independent computer-based information systems (National Council for Prescription Drug Programs Data Dictionary, 1994). The standard includes application-level syntax rules, message design guidelines, syntax implementation guidelines, data element dictionary, code list, composite data-elements dictionary, standard message dictionary, uniform rules of conduct for the interchange of trade data by transmission, and explanatory material.

The basic EDIFACT (ISO 9735) syntax standard was formally adopted in September 1987 and has undergone several updates. In addition to the common syntax, EDIFACT specifies standard messages (identified and structured sets of statements covering the requirements of specific transactions), segments (the groupings of functionally related data elements), data elements (the smallest items in a message that can convey data), and code sets (lists of codes for data elements). The ANSI ASC X12 standard is similar in purpose to EDIFACT, and work is underway to coordinate and merge the two standards.

EDIFACT is concerned not with the actual communications protocol but rather with the structuring of the data that are sent. EDIFACT is independent of the machine, media, system, and application and can be used with any communications protocol or with physical magnetic tape.

7.5.2.11 Clinical Data Interchange Standards Consortium

The Clinical Data Interchange Standards Consortium (CDISC) was founded in 1997 in order to facilitate electronic regulatory submission of clinical trial data. The current standards include a study data model, a data analysis model, a lab data model and an operational data model that supports audit trails and metadata. In 2007, CDISC began a collaborative project with HL7, the National Institutes of Health and the USFDA to link research data with data derived from clinical care. This modeling effort, BRIDG (Biomedical Research Integrated Domain Group), has created a domain analysis model of clinical research based upon the HL7 Reference Information Model (RIM).

7.6 Today's Reality and Tomorrow's Directions

7.6.1 The Interface: Standards and Systems

Historically, interchange standards evolved to support sharing of information over complex networks of distributed systems. This served a simple business model in which data was pushed from disparate repositories with inconsistent architectures and data structures. This permitted the exchange of data for both business needs and patient care.

In today's medical environment, there are several competing forces that place a burden on standards requirements. The traditional scope of data sources included business level information, principally for payment needs. These were developed utilizing coding methodologies and business architecture that did not rely upon inclusion of primary clinical data into the reimbursement decision. With the advent of statutory requirements that demand justification of insurance claims and reimbursement, additional data forms and formats became essential. This lead to the development of claims attachment standards (see X12, above) that enabled more complex adjudication, comparative effectiveness, and accountable care. These standards will most certainly require structured, coded data rather than free-text and unstructured narrative.

Complexity of data requirements is constantly growing to better support evidence-based medicine, clinical decision support, personalized medicine, and accountable care. Each of these has overlapping, but fundamentally unique data streams. Moreover, the data provided at the point of care, if unfiltered, is likely to overwhelm the clinical decision making process. Elements of clinical data, such as events in pediatric years, must not compete for the attention of the caregiver. To an extent, this was solved with specifications, such as CCOW (see Sect. 7.5.2.3), which were developed to provide context aware data to that process. There are growing demands for increasing the depth and breadth of data delivered to that clinical environment. In addition,

these standards must support the implicit policy decisions about the nature of this data.

To date, clinical and pre-clinical information populates many of the alerts that clinicians receive at the point of care. Typically, these range from information supporting complex decision trees to the selection of testing and interventions. This has been abetted by increasing knowledge of genomic data and implication for therapeutic decisions. Although this has had its greatest impact on the chemotherapy of cancer, the importance in many other clinical domains, for more common conditions (including the treatment of diabetes, hypertension, and arthritis) is now recognized. Current architectural systems are ill-prepared to manage this process. Moreover, data formats for genomic and genetic information are disparate and often incompatible.

Data privacy requirements, and the variability of these requirements among legal entities, currently pose a different set of demands for information access technologies. For example, some states permit line-item exclusion of clinical data that is transferred between providers, based on the primacy of the information and the role of the caregiver. Other jurisdictions allow participation of health information exchanges to those individuals who agree only to dissemination of data from complete sources.

Existing data architectures enable a constant stream of data to be passed in an untended and unmonitored fashion. In evolving models, data request and acknowledgement require a more complex query and response logic. In fact, most inquiries demand the validation of the provider system and privileges that are afforded to both the caregiver and the primary data repository. This places another component of interface design between the respective systems, and necessitated the development of analogous provider indices and provider repositories. Concerns of both privacy and security must be met by these specifications. The system effectively asks not only who you are but why you want the information.

Much of this process overhead has been addressed by the design and architecture of health information exchanges. Often the business case supersedes the demand for clinical knowledge. At the same time, these exchanges are designed to behave in an entirely agnostic fashion, placing no demand on either the sender or recipient for data quality, but only for source identification. In fact, the metadata, so responsible for the value of the information, is often capable of specifying only its origin and value sets.

In today's clinical environment, there has been very little attention paid to the capture and validation of patient-initiated data. While so very critical to diagnosis and ongoing management, only scant standards exist for embedding patient derived information into the clinical record without intermediate human interaction and adjudication. When allowed by current systems, data provided by patients often lies within the audit trail, as a comment, rather than in the record as source data. Steps are sorely needed to define and attribute such data since it is so critical to many aspects of accountable care. Data obtained directly from patient sources is often attributed to "subjective" status, but it is no less objective that many clinician observations. Perhaps justification for that lies in the fact that this patient derived data is neither quantifiable nor codeable. This is supported by valid concerns about the patient's health care literacy, or lack thereof, but is no less required than validated decision support for caregivers.

Data obtained from clinical research and clinical data provided to inform clinical studies suffer from other concerns of failed interoperability. This is attributed, and rightfully so, to disparities of vocabularies utilized for patient care and those used in clinical research. This is most dramatically highlighted in the terminology deployed by regulation for adverse event reporting (MedDRA; Medical Dictionary of Regulatory Affairs). Mapping between the MedDRA dictionary and other clinical terminologies (SNOMED-CT, LOINC, ICD, and CPT) has not proven successful. Moreover, many aspects pertaining to study subject inquiries in clinical research are often designed to elicit yes-no responses (Have you smoked in the last 5 years), rather than data that many caregivers deem relevant. Yet, today, it is more critical than ever to enable clinical research to inform patient care and care derived data to

enable clinical trials. The business model of developing drugs for billon dollar markets ("blockbuster drugs") has proven itself to be unsustainable, as the cost of developing a new drug entity has now exceeded a billion dollars. From the clinical perspective, current estimates suggest that information from basic science research experiences delays of nearly 17 years before that knowledge can be incorporated into clinical care.

Semantic interoperability of clinical data inherently requires *data reuse*. It is not sufficient for systems to unambiguously exchange machine readable data. Data, once required only for third party payment, must be shared by other partners in the wellness and health care delivery ecosystems. Certainly, these data must be presented to research systems, as noted above. The data must also be available for public health reporting and analysis, for comparative effectiveness research, for accountable care measurement and for enhancement of decision support systems (including those for patients and their families). The immediate beneficiaries have been systems developed to support biosurveillance and pharmacovigilance. In practical terms, the business practices that govern our delivery systems (and the government policies and regulations that enable them) must enable these data streams to both enhance care and control costs.

7.6.2 Future Directions

The new models for health care require a very different approach. The concept of a patient-centric EHR (Chap. 15) requires the aggregation of any and all data created from, for, and about a patient into a single real or virtual record that provides access to the required data for effective care at the place and time of care. Health information exchange (HIE; see Chap. 12) at regional, state, national, and potentially at a global level is now the goal. This goal can only be reached through the effective use of information technology, and that use can only be accomplished through the use of common global standards that are ubiquitously implemented across all sites of care.

Three other future trends influence the need for new and different standards. The first is secondary use of data by multiple stakeholders. This requirement can only be met through semantic interoperability – a universal ontology that covers all aspects of health, health care, clinical research, management, and evaluation. Standards for expressing what is to be exchanged and under what circumstances are important as well as standards for the exchange of data. Included in multiple uses of data is reporting to other organizations such as immunization and infectious disease reports to the **Centers for Disease Control and Prevention** (CDC), performance reports to the **National Quality Forum** (NQF) and audit reports to **The Joint Commission** (TJC). Such systems as described also enable population health studies and health surveillance for natural and bioterrorism outbreaks.

The second trend area is the expansion of the types of data that are to be included in the EHR. The new emphasis on translational informatics will require new standards for the transport, inclusion into the EHR, and use of genetic information including genes, biomarkers, and phenotypic data. Imaging, videos, waveforms, audio, and consumer-generated data will require new types of standards. Effective use of these new types of data as well as exponential increases in the volume of data will require standards for decision support, standards for creating effective filters for presentation and data exchange, and new forms of presentation including visualization. New sources of data will include geospatial coding, health environmental data, social and community data, financial data, and cultural data. Queries and navigation of very large databases will require new standards. Establishing quality measures and trust will require new standards. Ensuring integrity and trust as data is shared and used by other than the source of data will require new standards addressing provenance and responsibility.

The third trend area is the use of mobile devices, smart devices, and personal health devices. How, when and where such devices should be used is still being explored. Standards will be required

for safe design, presentation, interface, integrity, and protection from interference.

True global interoperability will require a suite of standards starting with the planning of systems, the definition and packaging of the data, collection of the data including usability standards, the exchange of data, the storage and use of data, and a wealth of applications that enable the EHR for better care. IT systems must turn data collected into information for use – a process that will require the use of knowledge in real time with data to produce information for patient care. Selecting the correct knowledge from literature, clinical trials, and other forms of documentation will require standards. Knowledge representation, indexing, and linkages will require standards.

A major, and challenging, requirement to address these new types and use of data will be effective standards for privacy and security. These standards must protect, but not restrict the use of data and access to that data for determining and giving the best care possible. The aggregation of data requires an error-free way for patient identification that will permit merger of data across disparate sources. Sharing data also requires standards for the de-identification of patient data.

The effective management of all of these resources will require additional output from the standards communities. Standards for defining the required functionality of systems and ways for certifying adherence to required functionality is essential for connecting a seamless network of heterogeneous EHRs from multiple vendors. Testing of standards, including IHE Connectathons before wide-spread dissemination and perhaps mandated use of standards is critical to use and acceptance. Standards for registries, standards for the rules that govern the sharing of research data, standards for patient consent, and standards for identification of people, clinical trials, collaboration, and other similar areas are necessary. Profiles for use and application from the suite of standards are a necessity. Detailed implementation guides are key to use and implementation of standards. Tools that enable content population and use of standards are mandatory for easy use of standards.

Standards for these new and evolving business and social needs must be supported by changes in the standards development methodologies and harmonization. Legacy systems are not easily discarded. Recommendations for complete replacement of existing standards are neither politically expedient nor fiscally supportable. Currently, there is increasing attention to new approaches to standards development that speeds the creation process and improves the quality of standards that are developed. These evolving development platforms pay appropriate homage to existing standards and leverage previously developed models of development and analysis.

The use of the FHIR, which leverages the HL7 Reference Information Model, may provide a much needed solution While relying upon historically developed and refined interoperability specifications, it hides the complexity of authoring messages within the FHIR development process. This leads to more usable specifications, created in a dramatically abbreviated time frame. Other approaches to standards development, such as those focused upon services, are rapidly evolving. These services-aware architectures are governed by strict development principles that help ensure both interoperability and the ability of components to be reused.

Increasingly large data stores ("big data") have demanded some of these changes. These data have emanated from a highly diverse universe of scientific development. In fact, some of the new bio-analytic platforms for *in vitro* cellular research are generating data at a rate, which by some estimates, is faster than the data can be analyzed. Medical images, for which storage requirements are growing, must now be principally evaluated by human inspection. Newly evolving algorithms and the technologies to support them, initially developed for star wars-type image analysis, are replacing radiologist and pathologists for the establishment of diagnoses. These machines have proven to be faster and more accurate than their human counterparts. In the very near future, such instrumentation will supplant medical scientists the same way that comparable technologies replaced human inspection in the estimation of cell differentials

for blood counts. These new technologies are demanding the development of specifications and the terminologies to support them.

Tomorrow's technologies will transition from early vision through prototyping to commercial products in a more compressed life cycle. A model for this process in biomedical science was established with the emergence of the Human Genome Project (see Chap. 24). Within the next decade, routine genome determination and archiving, as well as their application to disease management, will require greatly enhanced solutions for data management and analysis. Innovative strategies for recognizing and validating biomarkers will grow exponentially from the current stable of imaging and cell surface determinants. These data streams will require adaption of existing decision support systems and comparative effectiveness paradigms. Lastly, scientific evidence supporting the diagnosis and management with the field of behavioral medicine will change the entire clinical spectrum and approach to evaluation and care. As we emerge from the dark ages of behavioral medicine, we will certainly require new systems for recognizing, diagnosing, naming and intervening on behalf of our patients.

In some sense, the development of standards is just beginning. The immediate future years will be important to create effective organizations that include the right experts in the right setting to produce standards that are in themselves interoperable. That goal still remains in the future.

Suggested Readings

Abbey, L. M., & Zimmerman, J. (Eds.). (1991). Dental informatics: integrating technology into the dental environment. New York: Springer. This text demonstrates that the issues of standards extend throughout the areas of application of biomedical informatics. The standards issues discussed in this chapter for clinical medicine are shown to be equally pertinent for dentistry.

American Psychiatric Association Committee on Nomenclature and Statistics (1994). *Diagnostic and Statistical Manual of Mental Disorders*. (4th ed.). Washington, D.C.: The American Psychiatric Association.

Benson, T. (2012). *Principles of health interoperability HL7 and SNOMED* (2nd ed.). London: Springer. This book presents a detail discussion of the HL7 version 2 messaging standard and a detail presentation of SNOMED-CT.

Boone, K. W. (2012). *The CDA™ book*. London: Springer. This book provides an excellent presentation of the HL7 Clinical Document Architecture and related topics.

Chute, C. G. (2000). Clinical classification and terminology: some history and current observations. Journal of the American Medical Informatics Association, 7(3), 298–303. This article reviews the history and current status of controlled terminologies in health care.

Cimino, J. J. (1998). Desiderata for controlled medical vocabularies in the twenty-first century. Methods of Information in Medicine, 37(4–5), 394–403. This article enumerates a set of desirable characteristics for controlled terminologies in health care.

Executive Office of the President; President's Council of Advisors on Science and Technology (2010). *Report to the President realizing the full potential of health information technology to Improve healthcare for Americans: the path forward*.

Henderson, M. (2003). *HL7 messaging*. Silver Spring: OTech Inc.. Description of HL7 V2 with examples. Available from HL7.

Institute of Medicine. (2003). *Patient safety: achieving a new standard for care*. Washington, D.C.: National Academy Press. Discusses approaches to the standardization of collection and reporting of patient data.

New York Academy of Medicine (1961). *Standard Nomenclature of Diseases and Operations*. (5th ed.). New York: McGraw-Hill.

Richesson, R. L., & Andrews, J. E. (2012). *Clinical research informatics*. London: Springer. This book includes a discussion of standards and applications from the clinical research perspectives.

Stallings, W. (1987). *Handbook of computer-communications standards*. New York: Macmillan. This text provides excellent details on the Open Systems Interconnection model of the International Standards Organization.

Stallings, W. (1997). *Data and computer communications*. Englewood Cliffs: Prentice-Hall. This text provides details on communications architecture and protocols and on local and wide area networks.

Questions for Discussion

1. Who should be interested in interoperability and health data standards?
2. What are the five possible approaches to accelerating the creation of standards?
3. Define five health care standards, not mentioned in the chapter, which might also be needed?
4. What role should the government play in the creation of standards?
5. At what level might a standard interfere with a vendor's ability to produce a unique product?
6. Define a hypothetical standard for one of the areas mentioned in the text for which no current standard exists. Include the conceptualization and discussion points. Specifically state the scope of the standard.