



# Public Health Information Standards

# 8

J. A. Magnuson, Riki Merrick, and James T. Case

## Learning Objectives

1. List and describe two common categories of standards: process standards and content standards.
2. Discuss the anticipated benefits of standards use and the obstacles to adoption of standards.
3. Identify and discuss examples of standards used in public health systems.
4. Describe the general process for standards development.
5. Review details about three of the commonly-used standards in public health, HL7<sup>®</sup>, LOINC<sup>®</sup>, and SNOMED CT<sup>®</sup>.

“speak” different languages, and coordinate time-sensitive materials and actions, and that are often critical to the health and safety of individuals or populations; examples might include systems for air traffic controllers, police, and hospitals. Public health (PH) systems may not seem as obvious a choice, but they also fit into this category. Air traffic controllers would be unable to ensure the safety of modern air travel if there were no standard terminology or communication procedures. Similarly, without the ability to share and exchange data with its many partners, both private and governmental, public health would be unable to fulfill its primary goal of population focused prevention of disease, injury, disability, or environmental impact.

There are many PH systems which utilize standards, including communicable disease reporting; surveillance of emergency department (ED), emergency medical services (EMS), pharmacy, over-the-counter (OTC), and poison control data; air and water quality testing; and data sharing with private and public entities/programs. This chapter will highlight laboratory reporting of communicable disease data as its primary example system.

## Introduction

With a little imagination, one can picture many systems that must communicate over distances,

---

J. A. Magnuson (✉)  
Department of Medical Informatics and Clinical  
Epidemiology, Oregon Health & Science University,  
Portland, OR, USA  
e-mail: [magnusoj@ohsu.edu](mailto:magnusoj@ohsu.edu)

R. Merrick  
Vernetzt, LLC, Wilton, CA, USA

APHL, Silver Spring, MD, USA

J. T. Case  
SNOMED International, London, UK  
e-mail: [jca@snomed.org](mailto:jca@snomed.org)

## Context for Public Health Standards

As public health continues its advance into the arena of *electronic data interchange (EDI)*, interoperability, or the capacity to exchange and

utilize data between systems, becomes increasingly critical. Examples of EDI in public health are many and varied, such as:

- *Communicable disease reporting* from laboratories (Electronic Laboratory Reporting or ELR)
- *Wide-ranging surveillance* of sources such as emergency department, emergency medical services (EMS), pharmacy, over-the-counter (OTC), poison control, and absenteeism data
- *Interoperability Initiatives*, including the Interoperability Standards Advisory (ISA) [1] process through which the Office of the National Coordinator for Health Information Technology (ONC) coordinates healthcare interoperability standards and implementation specifications. The Meaningful Use (MU) objectives (identified by the Centers for Medicare & Medicaid Services as part of Incentive Programs to promote the adoption of Electronic Health Record Systems), included public health choices for immunizations, syndromic surveillance, and ELR [2].
- *Data sharing and exchange* within-state and between state partners such as outside states, local or regional health departments, tribal governments, and federal agencies

Standards are important for all instances of public health EDI, including the examples above. But despite the criticality of interoperability (the ability to communicate or share data), not all PH systems have achieved this goal. Jernigan et al. [3] list three main causes of non-communicating or silo-ed public health systems:

- *Functional requirements*: design differences may be based on function, for example the function of case management vs. the function of population surveillance.
- *Policy requirements*: policy restrictions upon systems, such as those that could restrict choices of software.
- *External restrictions* imposed by federal funding. Some of the silo-ed systems still in use today are actually required. One such current example is the Enhanced HIV/AIDS Reporting

System (eHARS), a browser-based HIV surveillance system used by state and local health departments to submit de-identified data electronically to the Centers for Disease Control and Prevention (CDC) national database [4].

In this chapter, two additional causes of public health silos are added, both closely tied to standards:

- *System architecture* or more specifically, lack of system architecture. If the importance of system integration and architecture is either unknown or discounted when building a new system, then the outcome will be an isolated, non-integrated system. Harmonization, or at least accommodation, of standards is a critical factor in system integration.
- *Exchange partner variations*. There is a wide variety of exchange partners inherent in public health matters, and the corresponding barrier of asking all these partners to agree upon and incorporate any chosen standard can be formidable. Public health exchange partners include local, regional, state, and federal public health agencies; other government agencies, such as social service agencies; the public and its personal health records; laboratories; hospitals; and other data generating entities. As public health becomes more adept at sharing its wealth of data and information these partners, either currently or in the future, will expand to include new data *receiving* entities, such as bi-directional exchange with laboratories or with clinical Electronic Health Record (EHR) systems.

---

## The Value of Standards

One of the most efficient ways to prevent data silos, achieve system interoperability, and promote the value of data is through the utilization of standards. Establishing and gaining consensus for standards is not an easy task, however, and historically public health has lagged industry (though not healthcare in general) in agreeing upon and utilizing standards. But to keep per-

spective on the difficulty of such an endeavor, an observer need only consider that although the metric system was introduced in France in 1799, the United States is today the only industrialized country that does not utilize it as its official standard of measurement (the metric system is certainly accepted in the US, but it is not yet the official standard). In a nutshell—implementation of standards is hard.

---

## Obstacles to Adoption of Standards

In order for a standard to be both useful and accepted by the community, there must be agreement among the stakeholders on the goals to be accomplished through adoption of standards. This agreement may be a challenging objective in itself, especially if the industries and groups are fragmented. Additionally, even the experts often disagree on details. The difficulties and costs inherent in the implementation of standards within any organization must be justified by stated objectives for the exchange of data or the utilization of aggregate data from multiple institutions.

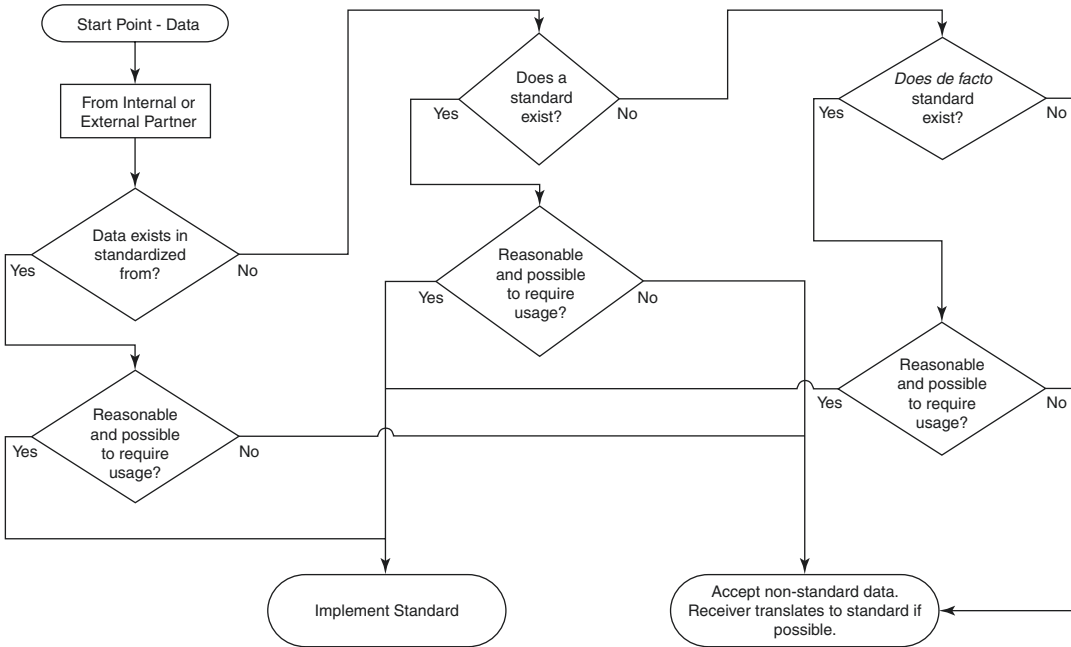
Ideally, standards are developed by a panel of experts and formally approved by a *standards development organization (SDO)* such as the International Organization for Standardization (ISO) [5] or the American National Standards Institute (ANSI) [6]. In practice, many “standards” are the product of legacy use within an industry or group. Such *de facto* standards can be extremely useful when no formal standards are available.

The process of developing a standard differs somewhat between SDOs, but there remains a basic similarity, illustrated here by the ISO process. The International Organization for Standardization follows a six-step process [7] when developing a standard. The process begins with (step 1) a proposal to the appropriate technical committee (TC), and then (step 2) a working draft is developed by a group of experts and (step 3) shared with the TC. Next, the draft is (step 4) released for comment by all ISO national members and (step 5) the final draft, after reconcilia-

tion of comments, is sent to all ISO members for a vote. If approved by the vote, the draft finally becomes (step 6) an official ISO International Standard.

Agreeing upon a standard is only the beginning. Despite the critical advantages and benefits of standards, it is important to remember that not all partners in a data exchange may share equally in those benefits. The implementation and utilization of standards is often resource-intensive, and many times the essential costs of standards implementations are borne by partners who may not share in the benefits. A good illustration of this situation involves *electronic laboratory reporting* or *ELR*. Around 2000, public health began asking the laboratories legally required to send reportable condition data to public health (including hospital, private, and public health laboratories) to report electronically through the new institution of ELR. Implementation of ELR systems created a potential for faster disease reporting [8]. This new reporting path, however, required that the data submitted be standardized in both format and content. These new requirements for standardization often created significant expense for laboratories, especially those with multi-jurisdictional clients requiring multi-jurisdictional reporting. The laboratory must be able to retrieve, format, and transmit data from the *Laboratory Information System (LIS)* or *Laboratory Information Management System (LIMS)*, originally signaling industrial settings, though that distinction is fading and the terms are often considered to be interchangeable), as well as apply standard codes for laboratory tests and results. This resulting standardization is highly valuable to public health, but confers little practical return on investment for the submitting laboratory. Some funding has been made available to assist laboratories, either directly by public health entities or tangentially by measures such as the Meaningful Use incentives. Nevertheless, in most cases the efforts have been funded largely by the laboratories themselves.

As demonstrated in the previous example, decisions to develop, select, implement, or require standards should not be reached without careful consideration. Figure 8.1 illustrates



**Fig. 8.1** Sample decision process flow for standards utilization

a decision process flow that represents effective contemplation of such standards issues.

### Standards Categories

Categorization of public health standards can help to simplify the subject, but there are a number of different categorical schemas from which to choose. For example, in February 2006, the Health Information Technology Standards Panel (HITSP) [9] separated health information technology standards into the categories and corresponding examples shown in Table 8.1.

In this chapter, a slightly different categorization will be used, and public health informatics standards will be considered as divided into two fundamental categories, *process* standards and *data* or *content* standards. *Process standards* include procedure and policy standards. Examples of process standards include security policies, data use agreements, workflow, architectural, and metadata standards (creating some overlap with data standards). *Data (or content) standards* address common terms and methods,

**Table 8.1** One example of public health standards categorization, based on work done by Public Health Data Standards Consortium [10]

Standards categories	Examples
Data standards	Vocabularies and terminologies, such as LOINC® or SNOMED CT®
Information content standards	Reference information models, such as HL7® Reference Information model (RIM) [11]
Information exchange standards	Message-based and structured document-based, such as HL7®
Identifier standards	Identifiers, such as the National Provider Identifier (NPI) [12]
Privacy and security standards	Access control, audit, electronic consent
Functional standards	Work processes, workflow and dataflow models
Other standards	Internet standards, transport mechanisms

and increase the ability to share data between systems, i.e., interoperability and integration. The theoretical components of data standards are (a) vocabulary, (b) format, and (c) transmission. Transmission standards include privacy and con-

fidentiality components, and so tend to overlap somewhat with process standards.

## Process Standards

As mentioned earlier, transmission standards and process standards have some degree of overlap in the areas of privacy and confidentiality. *Privacy* generally refers to a ‘people’ context, a state of being free from unauthorized intrusion or invasion. This concept is as applicable to medical records as it is to a person’s own residence. *Confidentiality* is viewed more in the context of information, usually dealing with accessing and sharing information or data.

## Security Policies

Data *integrity* (freedom from errors or flaws) and confidentiality are often the prime focus of security concerns. Data integrity must be maintained during any transaction; for example, when reporting data, the data received must be exactly the same as the data that were sent. Data confidentiality is a critical concern in public health, which frequently requires the exchange of clinical or laboratory data containing patient identifiable information. Almost all public health agencies are concerned with confidentiality, since they routinely deal with sensitive data that are their legal responsibility to safeguard. A breach in security that allowed patient identifiable data to be made public would jeopardize the ability of a public health agency to perform its data gathering duties, as well as damaging its public reputation as a trustworthy government agency.

The confidentiality and privacy of an institution’s data depends in large part upon enterprise security—the administrative, physical, and technical security measures enacted by the institution to safeguard its systems. Physical security measures (e.g., locked doors and security patrols), administrative measures (e.g., limiting access rights of employees, providing management and financial support for security policies, prohibiting downloading/playing of music on computers), and technical measures (e.g., firewalls, encryption, digital certificates)—all must be part of an

effective enterprise security solution. Effective security policies will address these issues, and may be authored locally or involve collaboration between entities or jurisdictions. HIPAA, the Health Insurance Portability and Accountability Act of 1996 [13], included efforts to improve health data security nationally. Both HIPAA and security are discussed in detail in other parts of this book, including Chaps. 4 and 10.

## Data Use Agreements (DUA)

Data use agreements (DUAs) are legal agreements between entities that are intended to ensure appropriate safeguarding and use of shared information or data. DUAs will include details of the agreed-upon security measures and confidentiality requirements, such as the conditions under which data may be accessed and disclosed. An effective DUA will also include measures to ensure tracking of data and data use, to enforce compliance with the DUA and provide evidence in the case of a security breach or unauthorized use.

## Metadata

Metadata is often described as “Data about Data,” and entails structured information that facilitates usage and management of an information resource [14]. Metadata not only makes it easier to generate value from a resource, it enables continued usage of the resource by providing vital descriptive and identifying information for future users. For this discussion, we will review three important divisions of metadata:

- *Descriptive metadata*—generally provides the context for the data, which can be used for discovery and identification, e.g., title, abstract, author, and keywords
- *Structural metadata*—describes the parts of compound objects, e.g., sections of a document
- *Administrative metadata*—information for resource management, e.g., a database creation date and development platform. May be considered to contain the concepts of rights management metadata (intellectual property rights) and preservation metadata (archival information).

## Data or Content Standards

As discussed earlier in this chapter, data or content standards are divided into three categories, (a) format, (b) vocabulary, and (c) transmission. All three categories are involved in public health (or other health system) EDI, as illustrated in the foodborne outbreak investigation example profiled in Table 8.2.

The following sections discuss the vocabulary and format standards in greater detail. In order for systems to successfully communicate or interface, there must be both functional and semantic interoperability. *Functional interoperability* occurs when systems are able to physically communicate or share data, whereas *semantic interoperability* involves interpretation of data via a common language or *vocabulary*. Interoperability is facilitated by standards of all categories.

## Data Format Standards

Information exchange standards define *structure* (parts) and *syntax* (arrangement), including to some extent the vocabulary, of the electronic communication and are referred to as the standard ways of sending and receiving information [15]. These standards can be compared to the grammar requirements in a language.

Health Level Seven (HL7<sup>®</sup>) [16] is an international standard that is the most widely used formatting standard for health data. Created by developers in the 1980s, it is present in most hospital systems and has been adopted by public health as a data format standard.

HL7<sup>®</sup> is a complex and flexible set of format protocols that can encompass a staggering array of data requirements. The flexibility of HL7<sup>®</sup> can be a ‘good news/bad news’ attribute—while it can accommodate an enormous variety of data situations (definitely good), users can also create an

**Table 8.2** Examples of standards utilized during a public health foodborne outbreak investigation

Trigger	Actions	Standards involved
Public health—foodborne outbreak investigation		
Attendees at a celebration dinner become ill	Local PH begins an investigation	<i>Process</i> —privacy of individuals
Laboratory test identifies <i>Salmonella</i> in food sample	Laboratory prepares report for PH	<i>Data (format, vocabulary)</i> —report must be formatted according to agreed-upon standards; likely HL7 <sup>®</sup> format containing LOINC <sup>®</sup> (test) and SNOMED CT <sup>®</sup> (result) codes; data integrity preserved
Report ready for transmission	Laboratory transmits report to PH	<i>Data (transmission)</i> —privacy and confidentiality <i>Process</i> —security policies, metadata. DUA may be involved if lab has requested one adhere to transmission standards
State PH receives data (depending upon configuration of PH in that state/locality)	State PH system parses report and forwards data to local PH	<i>Data (format, vocabulary, transmission)</i> —State system must understand and translate report; determine which partners should receive, and format appropriately <i>Process</i> —secure storage and management of data at state; data integrity preserved; secure transmission of data to local PH; adhere to transmission standards
Local PH receives data	Local PH utilizes data in their investigation	<i>Data (format, vocabulary)</i> —local PH may capture and store test and result codes or other standards <i>Process</i> —maintain security to ensure confidentiality and data integrity
Investigation concludes	Local and state PH share data and findings	<i>Data (format, vocabulary, transmission)</i> —appropriate standards utilized and maintained <i>Process</i> —maintain security to ensure confidentiality and data integrity; adhere to transmission standards

Note the overlap between process standards and data (transmission) standards

astounding number of variations upon the standard, which may lead to confusion and extra effort. HL7® supports several product families, of which the first three will be discussed later in this chapter (section “A Closer Look: Format Standards Paradigm—HL7®”)—Version 2.x; Version 3, with its Reference Implementation Model (RIM) and the Clinical Document Architecture (CDA); and Fast Healthcare Interoperability Resources (FHIR), functional models describing requirements of Electronic Health Records, Domain Analysis Models (DAM), and Clinical Decision Support rules language.

### Vocabulary Standards

Vocabulary standards are often explained using the metaphor of language. If people are speaking different languages, it will be difficult for them to communicate effectively. Similarly, if systems are using different vocabularies to refer to data content, it will be difficult for them to interoperate. Vocabulary standards can be considered to be either local or ‘universal’, depending upon the partners involved and how widely accepted the standard in question may be. It should be remembered that there are both advantages and disadvantages for local or universal standards. For example, local code sets may be more easily updated or changed by the source institution, but may make sharing data with other institutions much more difficult. ‘Universal’ codes enhance data sharing between systems and across regions, but may require specialized training to use, and may not be as flexible as local codes in adapting to local circumstances. As with different languages, one can also translate between the local and the universal codes.

One of the most important components of data standards is the consistent representation of clinical concepts or terms through the use of unique codes or identifiers. These are commonly referred to as *code systems*. Some of the areas where code systems are used in public health EDI include laboratory data, syndromic surveillance, and vital statistics. Laboratory-related EDI will be used as the primary example in this chapter, specifically the areas of:

- Laboratory tests
- Laboratory results
- Other subjects, such as specimen information, diagnoses and clinical findings, administration, or demographics

### Laboratory Test and Result Code Standards

#### Logical Observation Identifiers, Names and Codes (LOINC®)

The most widely adopted code system for laboratory observation coding in the United States is the Logical Observation Identifier Names and Codes (LOINC®) system [17]; LOINC® can be used to represent the name of both ordered and related performed tests. This code system is owned and maintained by the Regenstrief Institute [18]. Current LOINC® codes are 3–7 characters in length, and will expand as the code set continues to increase in content. LOINC® codes are constructed as the combination of a simple integer sequence number (beginning with “1”), a “dash” delimiter, followed by a Mod-10 check digit. LOINC® terms are composed of six major parts:

- Component/analyte—the substance or entity that is being measured or observed.
- Kind of property—the kinds of quantities or qualities relating to the same substance.
- Time aspect—measurement relates to either a point in time or a specified time interval. The vast majority of laboratory measurements are “point in time”.
- System type—for laboratory observations, this is equivalent to the sample type being analyzed.
- Scale—specifies the scale of measurement. The most common scales used in laboratory analyses are quantitative (QN), qualitative (QL), nominal (NOM), ordinal (ORD), and narrative (NAR).
- Method—this reflects the technique or procedure used to obtain the result.

Of these LOINC® parts, the code, analyte, property, timing, and scale are required. Both system

and method are able to be specified in other parts of an HL7<sup>®</sup> message, which is the primary vehicle for using LOINC<sup>®</sup> coded terms.

An example of a fully specified LOINC<sup>®</sup> term and its component parts is shown below:

**13203-5: *Borrelia burgdorferi***

**AB.IGM:PrThr:PT:CSF:ORD:IB**

- Component/analyte—*Borrelia burgdorferi* AB.IGM. The specific immunoglobulin subclass IgM stimulated in response to the presence of *Borrelia burgdorferi* (the organism that causes Lyme disease) antigen.
- Property—PrThr. Presence/Threshold. Used for LOINC<sup>®</sup> terms whose results are reported using an ordered categorical scale, regardless of whether or not an internal threshold was used to make that determination.
- Time aspect—PT. Point or moment in time, i.e., the time the sample was collected.
- System—CSF. Sample type, cerebrospinal fluid.
- Scale—ORD. Ordinal, a qualitative ordered list of values such as “Detected,” “Not detected,” “Positive” or “Negative.”
- Method—IB. Measurement method, “Immune blot.”

**SNOMED CT<sup>®</sup>**

Standardized laboratory result coding for non-numeric values is increasingly being represented using the Systematized Nomenclature of Medicine—Clinical Terms (SNOMED CT<sup>®</sup>). SNOMED CT<sup>®</sup> was initially produced by the College of American Pathologists (CAP), which entered into an agreement with the US National Library of Medicine (NLM), funded by the Department of Health and Human Services, to offer open access to the US for the International release content of SNOMED CT<sup>®</sup> and the US edition of SNOMED CT<sup>®</sup>. Since 2007, the code system has been owned and managed by SNOMED International<sup>®</sup> (previously known as the International Health Terminology Standards Development Organisation (IHTSDO<sup>®</sup>) in the United Kingdom [19]. Section “A Closer Look: Vocabulary Standards—SNOMED CT<sup>®</sup>” of this chapter discusses SNOMED CT<sup>®</sup> in detail.

SNOMED CT<sup>®</sup> is a comprehensive, multi-lingual clinical terminology, with over 350,000 terms covering domains including clinical findings, diseases, procedures, organisms, observable entities, drugs, vaccines, and medical devices. It contains scientifically validated clinical content with the goal of providing a consistent representation of clinical content in electronic health record systems. The content is constructed and maintained using description logics, which allows for computable definitions that facilitate data analytics.

**Other Data Content Coding Standards**

Laboratory test and result standards are of course not the only coding standards of importance to public health informatics. A sampling of other important code systems includes:

- *Procedural codes*: The Current Procedural Terminology (CPT<sup>®</sup>) code system contains content developed (and copyrighted) by the American Medical Association. CPT<sup>®</sup> codes are 5-digit alphanumeric that classify medical service and are used for insurance billing [20]. As an example, the 2019 CPT<sup>®</sup> codes and Medicare payment information show that, for the region of New Mexico, US, “Application of short leg cast (below knee to toes)” has the assigned CPT<sup>®</sup> code of 29405, and a cost (facility: in hospital) of US \$60.11. It is of note that fee-based code systems, especially those that are generally accepted for reimbursement, are often far more advanced in their acceptance, adoption, and implementations. This reflects the popular adage, “Money talks and people listen.”
- *Geographic codes*: In 2006, the Geographic Names Information System (GNIS) Feature ID became the official federal reference to named geographic entities [21]. Using this system in 2019, the White House in Washington DC has an ID of 531723. However, public health often uses the legacy standard, the Federal Information Processing Standards (FIPS), to identify geographic areas such as states and counties [22]. Using FIPS, the Washington, DC code is 11001. Another



option for geographic coding is derived from census tracts. The 2010 US Census identified 74,134 tracts [23]. Census tracts are assigned at a county level, and are comprised of a 4-digit code (between 0001–9999), and a possible 2-digit additional code (between 0.01–0.98). While census tracts are unique within a county, they may be re-used for other counties within a state or in other states. For this reason, census tracts are also assigned an additional state (2-digit) and county (3-digit) FIPS code.

- *Industry and Occupation codes:* These code systems may be used by public health programs, such as programs tracking environmental issues like lead exposure. The Standard Occupational Classification (SOC) system, Bureau of Labor Statistics, classifies workers into occupational categories [24]. The 2018 dataset includes 867 detailed occupations, which are also grouped into broader categories. For example, Carpenters Assistant is classified as 47-3012 Helpers—Carpenters; Broad Occupation is 47-3010 Helpers, Construction Trades; Minor Group is 47-3000 Helpers, Construction Trades; and Major Group is 47-0000 Construction and Extraction Occupations. The North American Industry Classification System (NAICS) is the standard used by Federal statistical agencies in classifying business establishments for the purpose of collecting, analyzing, and publishing statistical data [25]. As an example, the 2017 NAICS Definition assigns code 238350 to “Finish Carpentry Contractors.” Unlike many countries, the US does not have a health worker registry with standardized roles such as physician, nurse, etc. The US also does not standardize health facility locations. These factors have resulted in silo-ed lists involving data that are not interoperable between agencies.
- *Demographic codes:* Information on demographic or population variables, such as gender, race, ethnicity, and age, are crucial to public health. Demographic codes, while not commonly used for public health data, include

the Race Value Set developed by HITSP [26]. As an example, this value set assigns a code of 2076-8 to the concept “Native Hawaiian or Other Pacific Islander.”

---

## A Closer Look: Format Standards Paradigm—HL7<sup>®</sup>

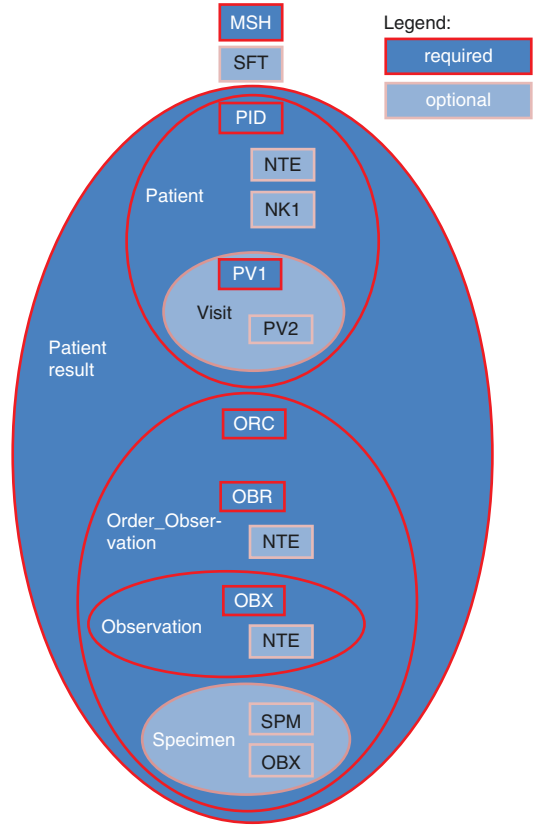
As previously mentioned, HL7<sup>®</sup> is the most widely used syntactic standard in Healthcare. The term ‘Health Level Seven’ refers to the Open Systems Interconnection (OSI) standard developed by the International Organization for Standardization (ISO) [5] in 1984. The OSI Reference Model defines the different stages that data must go through to travel over a network, and the seventh level (level 7) is the Application Level, which includes definition and structure of data in order to enable communication between computer systems. Since its inception in the 1980s it has undergone multiple transformations to adapt to the healthcare environment internationally. This section will look at three of the HL7 products that are most commonly used to date:

- *Version 2 (v2)* was developed to support message exchange between systems based on triggers; it was developed over time and had no formal underlying information model.
- *Version 3 (v3)* was developed as a response to the shortcomings of v2 and started with the development of the Reference Information Model (RIM) as well as clear definitions of user and application roles by formalizing transactions of clinical use cases.
- With the success of application programming interface (API) based technology in recent years HL7<sup>®</sup> has developed a new product called *Fast Healthcare Interoperability Resources (FHIR<sup>®</sup>)* that aims to combine the focus on implementer needs (as demonstrated in the bottom up evolution of v2) with some of the modeling learned from v3 while leveraging the API implementation experience current developers have from other domains.

### HL7® Version 2.x Artifacts: Reporting Observations (ORU)

Let’s look in detail at the contents of a v2.x HL7® message using the Observation message—for instance, to report a laboratory result. The way the segments are arranged within a message creates a hierarchy of information. There are two types of observation messages used in laboratory reports, (a) the ORU, which is patient centric, meaning the information about the patient comes before anything else, and (b) the OUL, which is sample centric, allowing for grouping under the sample rather than a patient. As an example, the OUL could be used to report a result from testing a water sample for contamination; water usually is not considered a patient. For this illustration, we want to look at the ORU because in healthcare we most often are interested in patient-related observations. The ORU message structure in the standard is defined by its required and optional segments, assembled in a specific order. Required means the segment must be sent, while optional means that it does not have to be sent—but if it is sent, it has to follow the rules of the standard. The standard also specifies whether a segment or a group of segments can be repeated. In order to apply the standard message definition to a specific use case an implementation guide can further constrain the standard, i.e., the underlying segment order must be maintained but changes can be made in the optionality (segments that were optional can be made required) or the number of times a segment or a group of segments can repeat.

Each segment has a specified number of fields that carry specific information related to the general topic of the segment (Fig. 8.2). For example the PID or patient information segment will have fields for name, date of birth, birthplace, address, gender, etc. Each field has a specific format called a data type; data types can be a string of characters (ST, string), while in other cases the format can be more complex and have several components. One such complex data type is Extended Person Name (XPN), which can contain last name, first name, other given names, suffix, etc. Data types may follow



**Fig. 8.2** Generic ORU message structure with the most commonly used segments. Copyright Riki Merrick 2013

a precise order that has meaning, such as in the Date/Time (DTM) data type, used for values like date of birth; DTM values are listed as four digit-year, two digit month, two digit day, and, if available, two digit hours, etc. Another complex data type very important for reporting observations is the Coded with Exceptions (CWE) data type, which is used to carry the codes describing the ordered tests, performed tests, sample types, and results. The ORU message type is used in several v2 implementation guides listed in the Interoperability Standards Advisory published yearly by ONC [27].

### HL7® Version 3 Artifacts: Clinical Document Architecture (CDA)

Along with several HL7® v2.x artifacts, the Health IT Standards Federal Advisory

Committee, in their September 2011 rule about Meaningful Use, named the HL7® Clinical Document Architecture (CDA) standard for use in data exchanges between clinical Electronic Health Record systems (EHRs), as well as from EHRs to public health information systems, called electronic Case Reporting (eCR) [28]. In the US, Meaningful Use required EHRs that create, transmit, receive, and display the Continuity of Care Document (CCD) [29], which uses the CDA as a framework. CDA serves as the basis for the creation of closely-related documents within a document-based health information exchange. CDA defines templates at different levels; documents, which have a header and a body, which in turn has its human readable and structured part comprised of template-based sections and entries. This structure is illustrated in Fig. 8.3, and can also be organized to construct valid public health reports.

MU requires the use of CCD for exchange of discharge information from hospitals to the patient’s primary care provider, for example, core clinical information about the hospital stay and instructions on what to do next. It must also be used to summarize clinical information when

a specialist needs to be involved in the patient care. The CCD described in MU is a collection of CDA templates, which will be combined in a specific order for each specific purpose, but the core information in each template is pre-defined [30].

An example of a CDA-based PH report is that sent for the group of reports about Healthcare-Associated Infections (HAIs), such as bloodstream infections, surgical site infections, urinary tract infections, etc., to the National Healthcare Safety Network at the CDC. CDA parts and specific vocabulary (LOINC®, SNOMED CT®, demographic standards, etc.) are defined in the implementation guide to ensure all required data for this reporting purpose are included. A different implementation guide, also CDA-based, is used to report about cancer patients to cancer registries in the US. Both the way the sections are put together and the vocabulary that is required are specialized to the needs of cancer reporting. In order to reduce the many variations imposed on the data providers (i.e., the EHR systems) the Public Health Reporting Initiative (PHRI) has convened many public health programs at the local, state, and federal level to collaborate and harmonize the format and vocabulary used

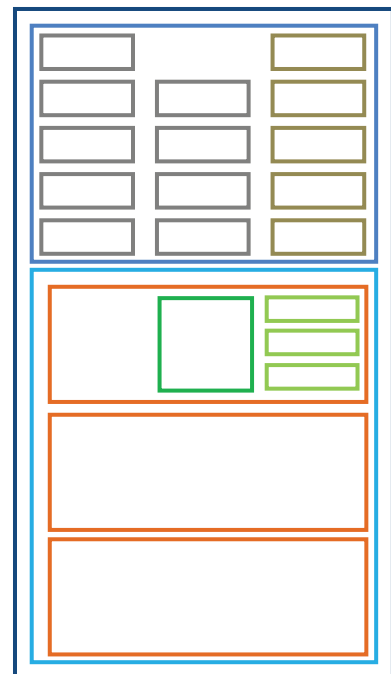
**Fig. 8.3** CDA building blocks. ©2012 Lisa R. Nelson, used with permission

**Document**  
**Header**  
 Header Participants  
 Header Act Relationships

**Body**  
**Sections**  
 Human Readable Text  
 Machine Readable Data

Content Modules are the modular content structures.

Templates define the structure and semantics of a content module.



for data that is needed across many different programs. These harmonized “Common Core” data elements have been incorporated into another Information model to retain the context, by defining format and vocabulary binding in the Federal Health Information Model (FHIM) [31], which is also linked to the underlying HL7<sup>®</sup> RIM.

## HL7<sup>®</sup> FHIR

In 2014 the newest HL7<sup>®</sup> product was published as the first standard for trial use (STU). HL7<sup>®</sup> Fast Healthcare Interoperability Resources (FHIR) builds on previous data format standards from HL7<sup>®</sup>, like version 2.x and version 3, and moves them into the more modern, web-based suite of API technology. Its design rules focus on specific implementation needs by applying the 80/20 rule for inclusion in the core specification. Requirements outside the core can be handled by creating extensions. FHIR<sup>®</sup> defines a set of resources; they are logical building blocks describing clinical elements similar to the segments in V2 or the templates in CDA. These resources are available in a choice of JSON, XML, or RDF for data representation on a FHIR server for discovery and exchange. Each resource includes metadata about itself, to support search and retrieval, and uses specific datatypes to represent dates, text, identifiers, and code-able concepts including the vocabulary binding to the standards needed [32]. Similar to v2 and CDA, implementers create profiles to describe the groups of resources needed to fulfill their use cases. Often v2 or CDA Implementation guides are used as starting points to develop the FHIR IGs, for example the electronic case report also exists as a FHIR IG [33].

---

### A Closer Look: Vocabulary Standards—SNOMED CT<sup>®</sup>

The LOINC<sup>®</sup> system for laboratory observation coding and the SNOMED CT<sup>®</sup> system of laboratory result coding have been discussed in a previous section of this chapter (section “Laboratory

Test and Result Code Standards”). In this section, a more detailed examination of the SNOMED CT<sup>®</sup> system is provided.

As mentioned earlier in this chapter, SNOMED CT<sup>®</sup> is a comprehensive reference terminology that encompasses all areas of healthcare. Its primary use is within EHRs, for the purposes of both meaning-based retrieval, data exchange and use in *clinical decision support* (assistance to health professionals in making choices). The by-product of consistent representation of clinical data is the ability to perform broad-ranging data aggregation, reporting, and analysis. SNOMED CT<sup>®</sup>, at its highest level, is based on three primary structures—concepts, descriptions, and relationships. A brief outline of the content and structure of SNOMED CT<sup>®</sup> is provided below.

## Concepts

Within the SNOMED CT<sup>®</sup> terminology, a concept is a unit of meaning that is assigned a unique “meaningless” identifier in numeric format. These identifiers are meaningless in that it is not possible by simply looking at the identifier to deduce any knowledge of the associated term or its position within the SNOMED CT<sup>®</sup> hierarchy. Each concept is represented by a description called the *fully specified name* (FSN) that uniquely represents the concept; this is accomplished through the combination of the description string and a semantic tag, which represents the top-level category to which the concept belongs. To illustrate this, let’s examine the term “swab,” which has multiple meanings within the healthcare environment. It may represent a physical object, a unit of a product, or a specimen type. Within SNOMED CT<sup>®</sup>, these are represented by uniquely identified concepts: 408098004 identifies swab as a physical object; 420401004, as a unit of product usage (qualifier value); and 257261003, as a specimen.

Concepts are arranged hierarchically within SNOMED CT<sup>®</sup>, such that less granular (more general) concepts are assigned as “parents” to more granular (more detailed) “children” through explicitly defined “is a” relationships (i.e., a granular concept “is a” child to the more general par-

ent concept). In some cases, concepts may have many parents depending on the types of defining relationships assigned to them. There are a number of important considerations when assessing these parent-child relationships:

1. For a concept to be a child of another concept, all of the defining attributes for the purported parent must be always and necessarily true for the child.
2. One cannot, by looking at the children of a concept, deduce the meaning of a parent concept: i.e., a parent defines the children, children do not define the parent.
3. Not all levels of intermediate granularity of meaning are represented by the terminology: i.e., there may be perceived “gaps” in the hierarchies.
4. It may not be possible to ascertain the full meaning of a concept without looking at all of the parents.

What this means in practice is that in some cases it may be difficult to ascertain the full meaning of the “words” in the concept description without looking at the surrounding content in SNOMED CT® to gain the full context of the term.

## Concept Identifiers

Concept identifiers are assigned permanently to any concept that is incorporated into the terminology. This means that once an identifier is assigned, it is never reused. Extensions to SNOMED CT® are assigned namespace identifiers that allow for the unique assignment of extension concepts, descriptions, or relationships that augment the content of the International release in order to meet specific needs of the extension owner. In the US, the National Library of Medicine has been assigned the extension namespace identifier for the official US extension to SNOMED CT®. This extension is designed to support the specific needs of US healthcare as identified by legislative mandates such as Meaningful Use. In general, SNOMED CT® identifiers have the structure demonstrated below.

- SNOMED CT® Identifier (SCTID): 101291009, is comprised of an item identifier (101291), a partition identifier (00), and a check digit (9).
- SCTID: 99999999 0989121 104, is comprised of an extension item identifier (99999999), a namespace identifier (0989121), a partition identifier (10), and a check digit (4).

SNOMED CT® differs from most other clinical terminologies in that it provides a multi-hierarchical representation of distinct clinical concepts as well as a set of defining relationships that allow systems to perform reasoning against the terminology. For example, a disease concept in SNOMED CT® may have assigned relationships to a causative agent (e.g., a particular species of bacteria) and a finding site (a specific anatomic structure). This would allow a reasoning system to classify the concept as a bacterial disease, an infectious disease, a disease affecting a particular part of the body, etc. This allows one to analyze SNOMED CT® encoded content from a variety of perspectives, based on the associated defining relationships attached to a concept.

## SNOMED CT® Descriptions

In addition to fully specified names, SNOMED CT® allows for a variety of alternative descriptions to represent the intended meaning of the concept. The primary purpose of these alternative descriptions is to provide users with assistance in searching the terminology for the proper concept that meets their particular need. While these alternative descriptions have historically been called “synonyms,” in many cases they are not true synonyms; the meaning of these terms could be more general, or in some cases, ambiguous. Looking back at the example provided at the beginning of this section, all three of the concepts related to “swabs” have alternate descriptions of “swab.” Without the knowledge of the concept’s fully specified name, one could not determine the full meaning of the descriptive term “swab.” Thus, it might be possible for the term to be used

incorrectly if a user did not have access both to the descriptive term and the associated FSN.

SNOMED CT<sup>®</sup> does not attempt to be exhaustive in the inclusion of alternative descriptions for a concept. Additional descriptions that may be of use locally can be added to approved extensions and treated as “regular” SNOMED terms. The lack of an exhaustive list of SNOMED CT<sup>®</sup> descriptions often causes users some frustration because SNOMED CT<sup>®</sup> “does not have my words.” Because the ways in which users might want to have terms displayed by their own EHRs is nearly limitless and often colloquial, SNOMED CT<sup>®</sup> has chosen to focus on adding descriptions that are of general clinical use and are as close to true synonymy as possible.

## SNOMED CT<sup>®</sup> Relationships

The relationships defined by SNOMED CT<sup>®</sup> are at the heart of the true value of the terminology, to provide enhanced usefulness for a variety of analytical needs. Through the explicit relationships, it is possible to easily select concepts based on particular attributes such as infectious disease, neoplastic disease, location on the body, or clinical manifestation. Because all concepts are related to one or more “parents,” it is possible to computationally aggregate highly specific terms into more general categories for trend reporting and analysis. This value allows data recorders to be as specific as possible with their entries, without having to worry about how their entries will be categorized during analysis.

## SNOMED CT<sup>®</sup> Browsers

Generally, users will first become familiar with the content of SNOMED CT<sup>®</sup> through exposure to one of a growing population of SNOMED CT<sup>®</sup> specific “browsers.” These tools, whether stand-alone or web-based, provide mechanisms to search for specific concept, descriptions, or identifiers within the entire SNOMED CT<sup>®</sup> terminology and then traverse associated hierarchies to view the terms located within the same “vicinity”

of the searched term. Currently, the US National Library of Medicine maintains a listing of available SNOMED CT<sup>®</sup> browsers [34]. These browsers include the NLM SNOMED CT<sup>®</sup> Browser, which differs from all the others in that it leverages the NLM Unified Medical Language System (UMLS) to find terms within SNOMED CT<sup>®</sup>. As mentioned above, SNOMED CT<sup>®</sup> does not attempt to include all possible descriptions that might be applicable for a particular SNOMED CT<sup>®</sup> concept. The NLM browser, by utilizing the power of the UMLS<sup>®</sup> Metathesaurus<sup>®</sup> (a multi-lingual collection of biomedical and health-related concepts, synonyms, and relationships), can use descriptions that originate from any of its over 150 source terminologies. Of these sources, 15–20 are updated annually. Thus, the NLM SNOMED CT<sup>®</sup> browser allows users to search for concepts in SNOMED CT<sup>®</sup> using descriptions that do not actually exist in the full SNOMED CT<sup>®</sup> terminology. This additional power provides more comprehensive retrieval of concepts than can be accomplished through the use of SNOMED-only browsers.

---

## Summary

There is a well-worn saying among standards aficionados, to the effect that the nice thing about standards is that there are so many from which to choose. Entertaining as that phrasing may (or may not) be, it actually may be true that the variety of standards really is a positive development. The incredible complexity of situations and data inherent in public health and healthcare EDI demands a similar complexity in standards.

---

## Future Directions

New technologies such as artificial intelligence, “big data” analytics, and biomedical ontologies will play an increasingly important role in the delivery of healthcare. The role of artificial intelligence in processing and interpreting health data has become important enough that a scientific journal, *Artificial Intelligence in Medicine* [35],

has been created to disseminate the ever-growing base of knowledge being applied to improving the performance of the healthcare provider and the healthcare system as a whole.

Already, there are new technology start-ups leveraging advances in machine-learning, image analysis, and access to large amounts of data to provide enhanced diagnostic capabilities in radiology, pathology, antimicrobial susceptibility, and many others. Technological advances have allowed the manufacture of a new breed of intelligent medical devices; wearables that can track heart rate, blood pressure, and glucose levels, and medical devices that allow for continuous monitoring and detection and early warning of health conditions. These devices create large amounts of highly granular data, requiring sophisticated algorithms to identify actionable events and communicate these to both the patient and the healthcare provider. The rapid growth in the field of genomics has resulted in an exponential growth in the amount of data that requires extensive and elegant computational power to provide meaningful applications in the field of precision or personalized medicine.

Artificial intelligence in pathology has been touted as a “third revolution” of pathology [36], coming on the heels of other advancement such as immunohistochemistry, genomic analysis, and digital pathology. The volume of data that can be generated from detailed image analysis and molecular diagnostics requires new approaches to gain an understanding of the molecular basis of disease states represented by the structural changes seen in histopathology.

In the area of medical terminology, the Open Biomedical Ontologies program is “... an effort to create controlled vocabularies for shared use across different biological and medical domains” [37]. All contributions to the project strive to conform to a set of guiding principles that have been accepted as “best practices.” While these ontologies have been mostly used within the research domain, efforts to integrate them with other terminology standards are underway, which will leverage the detailed knowledge represented in ontologies with the terminology used in clinical practice. This will ultimately result in more

robust clinical decision support systems that will be needed as precision medicine becomes more widely adopted.

Standards will be as critical to these future efforts as they are in the present. Harmonizing the use of standards in research, clinical care, and public health is paramount in order to fully utilize the power of all the collected data and create feedback loops among these sectors of the health care system.

### Review Questions

1. Discuss the two categories of standards used in this chapter, process standards and data or content standards. What are some examples of these standards that are used in public health?
2. Describe the general process for standards development. How can public health participate in standards development, and why should it do so?
3. Select two of the standards reviewed in this chapter and describe in detail (a) the benefits accrued to public health from use of the standard, and (b) the barriers to implementing this standard in public health.
4. If you had to choose between process standards and data or content standards, which category do you think is more important for public health to utilize?
5. Based on the two standards presented in depth in this chapter, HL7® and SNOMED CT®, compare and contrast their role in public health and their importance to population health as a whole. Use examples to strengthen your key points.

### References

1. HealthIT.gov. Introduction to the ISA. <https://www.healthit.gov/isa/>. Accessed 6 Oct 2019.
2. HealthIT.gov. Policymaking, regulation, & strategy. <https://www.healthit.gov/topic/meaningful-use-and-macra/meaningful-use-and-macra>. Accessed 6 Oct 2019.

3. Jernigan DB, Davies J, Sim A. Data standards in public health informatics. In: O'Carroll PW, Yasnoff WA, Ward ME, Ripp LH, Martin EL, editors. *Public health informatics and information systems*. Cham: Springer; 2003. p. 213–38.
4. Centers for Disease Control and Prevention. HIV surveillance reports. <https://www.cdc.gov/hiv/library/reports/hiv-surveillance.html>. Accessed 2 Nov 2019.
5. International Organization for Standardization. ISO. <https://www.iso.org/home.html>. Accessed 6 Oct 2019.
6. American National Standards Institute. About ANSI. [https://www.ansi.org/about\\_ansi/overview/overview](https://www.ansi.org/about_ansi/overview/overview). Accessed 6 Oct 2019.
7. International Organization for Standardization. Developing standards. Available from <https://www.iso.org/developing-standards.html>. Accessed 6 Oct 2019.
8. Panackal AA, M'ikanatha NM, Tsui F, McMahon J, Wagner MM, Dixon BW, et al. Automatic electronic laboratory-based reporting of notifiable infectious diseases at a large health system. *Emerg Infect Dis*. 2002;8(7):685–91.
9. Health Information Technology Standards Panel (HITSP). <http://www.hitsp.org>. Accessed 6 Oct 2019.
10. Public Health Data Standards Consortium (PHDSC). Health IT standards. <http://www.phdsc.org/standards/health-information-tech-standards.asp>. Accessed 27 Mar 2013.
11. HL7<sup>®</sup> International. HL7<sup>®</sup> version 3: reference information model (RIM). 2019. [https://www.hl7.org/implement/standards/product\\_brief.cfm?product\\_id=77](https://www.hl7.org/implement/standards/product_brief.cfm?product_id=77). Accessed 16 Oct 2019.
12. Centers for Medicare and Medicaid Services, US Department of Health and Human Services; National Provider Identifier Standard. <https://www.cms.gov/Regulations-and-Guidance/Administrative-Simplification/NationalProvIdentStand/index.html>. Accessed 6 Oct 2019.
13. U.S. Department of Health & Human Services, Summary of the HIPAA Security Rule. <https://www.hhs.gov/hipaa/for-professionals/security/laws-regulations/index.html>. Accessed 6 Oct 2019.
14. National Information Standards Organization, Understanding Metadata. [http://groups.niso.org/apps/group\\_public/download.php/17446/Understanding%20Metadata.pdf](http://groups.niso.org/apps/group_public/download.php/17446/Understanding%20Metadata.pdf). Accessed 6 Oct 2019.
15. The Office of the National Coordinator for Health Information Technology. Health IT playbook. Health information exchange. <https://www.healthit.gov/playbook/health-information-exchange/>. Accessed 6 Oct 2019.
16. Health Level Seven International. HL7<sup>®</sup> International. [www.hl7.org](http://www.hl7.org). Accessed 6 Oct 2019.
17. LOINC<sup>®</sup> from Regenstrief. Logical observation identifiers names and codes (LOINC<sup>®</sup>). <http://loinc.org/>. Accessed 6 Oct 2019.
18. Regenstrief Institute. Implementation LOINC<sup>®</sup>. <https://www.regenstrief.org/implementation/loinc/>. Accessed 6 Oct 2019.
19. SNOMED International International Health Terminology Standards Development Organisation. <http://www.snomed.org>. Accessed 6 Oct 2019.
20. American Medical Association. Practice Management CPT<sup>®</sup>. <https://www.ama-assn.org/practice-management/cpt>. Accessed 6 Oct 2019.
21. USGS. US Board on Geographic Names. <https://www.usgs.gov/core-science-systems/ngp/board-on-geographic-names>. Accessed 6 Oct 2019.
22. National Institute of Standards and Technology (NIST). Information Technology Laboratory. <https://www.nist.gov/itl/itl-publications/federal-information-processing-standards-fips>. Accessed 6 Oct 2019.
23. More about Census Tracts FCC Form 477. 2015. [https://transition.fcc.gov/form477/Geo/more\\_about\\_census\\_tracts.pdf](https://transition.fcc.gov/form477/Geo/more_about_census_tracts.pdf). Accessed 21 Oct 2016.
24. Bureau of Labor Statistics. Standard occupational classification. <http://www.bls.gov/soc/home.htm>. Accessed 3 Oct 2019.
25. US Census Bureau, North American Industry Classification System. <http://www.census.gov/eos/www/naics/>. Accessed 3 Oct 2019.
26. US Health Information Knowledgebase, Health Information Technology Standards Panel. <http://ushik.ahrq.gov/ViewItemDetails?system=hitsp&itemKey=86685000>. Accessed 3 Oct 2019.
27. Office of the National Coordinator for Health IT. 2019 Interoperability Standards Advisory. [www.healthit.gov/isa/sites/isa/files/inline-files/2019ISAReferenceEdition.pdf](http://www.healthit.gov/isa/sites/isa/files/inline-files/2019ISAReferenceEdition.pdf). Accessed 4 Nov 2019.
28. Health Level Seven International. HL7 CDA<sup>®</sup> R2 Implementation Guide: Public Health Case Report, Release 2 - US Realm - the Electronic Initial Case Report (eICR). 2018. [http://www.hl7.org/implement/standards/product\\_brief.cfm](http://www.hl7.org/implement/standards/product_brief.cfm). Accessed Jul 21 2019.
29. US Department of Health and Human Services Washington (DC): Office of the National Coordinator for Health Information Technology (US). Health information technology: initial set of standards, implementation specifications, and certification criteria for electronic health record technology. Regulation Identification Number 0991-AB58. 2010.
30. Lantana Consulting Group for NIST. Templated CDA: key concept for interoperability. 2011. [https://www.nist.gov/sites/default/files/templated\\_cda-key\\_concept\\_for\\_interoperability\\_0.pdf](https://www.nist.gov/sites/default/files/templated_cda-key_concept_for_interoperability_0.pdf). Accessed 4 Oct 2019.
31. FHIMS.org. Background on the FHIM. <http://www.fhims.org/background.html>. Accessed 3 Oct 2019.
32. HL7<sup>®</sup> FHIR. Introducing HL7<sup>®</sup> FHIR. <https://www.hl7.org/fhir/summary.html>. Accessed 4 Nov 2019.
33. HL7<sup>®</sup> FHIR. Electronic Case Reporting (eCR) Implementation Guide [current available ballot ver-



- sion]. <http://hl7.org/fhir/uv/ecr/2018Jan/index.html>. Accessed 4 Nov 2019.
34. U.S. National Library of Medicine. SNOMED CT® Browsers. [https://www.nlm.nih.gov/research/umls/Snomed/snomed\\_browsers.html](https://www.nlm.nih.gov/research/umls/Snomed/snomed_browsers.html). Accessed 16 Oct 2019.
35. Elsevier. Artificial intelligence in medicine. <https://www.journals.elsevier.com/artificial-intelligence-in-medicine/>. Accessed 16 Oct 2019.
36. Salto-Tellez M, Maxwell P, Hamilton PW. Artificial intelligence - the third revolution in pathology. *Histopathology*. 2019;74(3):372–6.
37. The OBO Foundry. [https://en.wikipedia.org/wiki/Open\\_Biomedical\\_Ontologies](https://en.wikipedia.org/wiki/Open_Biomedical_Ontologies). Accessed 23 Oct 2019.