

Public Health Data Standards

Improving How Public Health Collects, Exchanges and Uses Data

Introduction

The health care and public health communities are transforming how they manage and use health information. Rapid advances in information technology, coupled with national calls to improve the efficiency, quality and safety of health care, have contributed to a bold and historic initiative to create a national health information infrastructure.

Public health agencies are necessarily a part of this monumental transformation. Much of public health data comes from hospitals, laboratories and private providers. And public health has considerable data that has health policy, research and clinical value.

A key component necessary for the success of this nationwide transformation is the widespread adoption of data standards.

What Are Data Standards?

Data standards are an agreed-upon, common and consistent way to record information. They allow data to be exchanged among different information systems, and for that data to have consistent meaning from system to system, program to program, and agency to agency.

Data standards are important in almost every aspect of our lives. They are what make it possible for us to consistently measure distances and time, get the same results from a recipe as our next door neighbor, place phone calls across the globe, and withdraw money from almost

any ATM in the world. Without standards, the electronic exchange of information that occurs every second of every day across countless businesses and organizations would grind to a near halt. Without a consistent way to denote a piece of data, the communication, interpretation and translation of that data would become time-consuming at best, and totally erroneous at worst.

With public health data, standards make it possible for us to, for example, collect client names in the same way (for example, one field for first name, a second for middle initial, another for last name, and a final field for suffixes). When information systems collect and store client names in the same way, it is much more efficient and accurate for one system to send that data to another, or to compare and match names from two different systems so the data can be exchanged or merged.

Another example is how we agree to denote vaccine products. Td, DT, DTaP and TdaP are all different vaccine formulations. Without agreement on how to standardize the abbreviations, a public health nurse couldn't be sure what vaccine product to give today or whether a dose has to be repeated or not

There are different types of standards, each serving a particular purpose. For instance, there are standard ways to code nursing functions (the Omaha system used in PH-DOC, CHAMPS, and CareFacts), for diagnostic codes (ICD-9), to bill for medical services

(CMS 1500), to send health data between different information systems (HL7), and to code lab results (LOINC). In every case, standards enable computers to send data back and forth, usually in the same format and meaning the same thing. Being able to exchange data from information system to information system, without having to translate it into a new format and being able to retain the same meaning, is what is meant by the term *interoperability*.

So standards are basically universally agreed upon ways to handle data in ways that ensure interoperability.

What standards are most important to public health?

There are 2,100 different standards being used in health care today—an unwieldy number that highlights how standards have historically arisen to meet very specific needs in specific types of

settings. Reducing these to a manageable number that health care organizations, public health agencies, and venders can reasonably work with is the focus of considerable work nationally. For our purposes, we can group relevant standards for public health into four categories;:

- Terminology
- Messaging
- Transactions/claims
- Data content

Terminology standards are ways to define and classify individual health and other terms so that they are easily and consistently understood for one organization to another. Table 1 highlights the most important content standards for public health.

Table 1. Examples of Terminology Standards¹

Type of Content Standard	Definition	Example	Link
Code Sets	A list of codes, each code being associated with a particular result, product or term.	<p>LOINC (Logical Observations, Identifiers, Names, and Codes): Widely used by public health and clinical laboratories for electronic reporting of lab results.</p> <p>CVX: Code set developed by CDC to uniquely identify each vaccine product.</p>	<p>LOINC</p> <p>CVX</p>
Classification systems	A method for classifying data into terms that can be easily and consistently reported, understood, retrieved and analyzed	ICD-9-CM: The International Statistical Classification of Diseases and Health Related Problems – Clinical Modification is widely used by hospitals for both billing and statistical analyses, such as studies using hospital discharge data. ICD-9 is the classification used to code and classify mortality data from death certificates.	<p>ICD-9-CM</p> <p>ICD-9</p>

¹ Adapted from the [Public Health Data Standards Consortium's](#) tutorial module on data standards, 2006.

Nomenclature	Specialized terms that are given standardized, precise and unambiguous definitions, which makes meaningful exchange of data between providers possible.	<p>SNOMED: The Systemized Nomenclature of Medicine is a robust classification system used in human and veterinary medicine.</p> <p>Omaha System: A system for standardizing terminology used in nursing. Used in PH-Doc, CHAMPS and other integrated public health information systems.</p>	<p>SNOMED</p> <p>Omaha</p>
--------------	---	---	--

Creating consistent ways to classify data is critical but you also have to have a consistent way to send data back and forth between organizations. That is the role of **messaging standards**. The most widely used messaging standard in public health and health care is [HL7](#) (Health Level 7)².

Health Level 7 is a way to package data so that the receiving computer knows precisely what data is coming in, and where each data element occurs in the electronic file. For instance, HL7 will tell the receiving computer, “The next data you read will be Patient Identification Information.” It does this by using a specific HL7 code (in this case, ‘PID’) and putting it immediately in front of the string of relevant data, like so:

```
PID||0493575^^^2^ID1||DOE^JOHN^M^^|DOE^JOHN^M^^|19480203|M||B|254E38ST^^DULUTH^MN^55802^USA||(218)625-4359||
```

While this may look confusing at first, one of the advantages of an HL7 message is that it is, without too much effort, fairly readable by humans.

The HL7 coding scheme is used for a very wide range of clinical and

² For the curious, Health Level 7 refers to the top layer (Level 7) of the Open Systems Interconnection (OSI) layer protocol for the health environment.

demographic data, any of which may need to be exchanged between health care organizations. The beauty of HL7 is that the two organizations do not need to be using the same information systems—HL7 makes it possible for the computer in the receiving organization to make sense of the incoming data without staff having to manually sort the data into the appropriate fields.

Transaction or claims standards provide a uniform method for sending bills and getting reimbursed, as well as for exchanging other types of administrative data. Prior to these standards, processing claims was a very expensive task for both providers and insurance companies. Every insurance company had their own forms and requirements that providers had to learn. For years, the standard paper claims form was the Uniform Bill-92 or [UB92](#). With the enactment of the Health Insurance Portability and Accountability Act ([HIPAA](#)), the paper UB-92 form was replaced by the electronic standard [ASC X12N 837](#). Any provider that bills electronically must produce the claim using this format. By adopting this standard, the HIPAA requirements seek to improve administrative efficiency by reducing administrative costs across all health and health care settings.

Data content standards is a broad term that covers a wide range of data

standards, mostly around establishing a consistent, uniform way to capture, record and exchange data. For instance, when immunization registries first emerged in the early-1990's, CDC created a standard known as the [Core Data Set](#) that every registry could use not only as the basis for building or buying their application, but also for establishing what data needed to be reported by providers.

A well-known example of a data content standard is how we collect race and ethnicity data. If we did not have a standard way to collect this information (established by the [National Committee on Vital and Health Statistics](#)), we could not readily merge, compare, exchange or analyze different reports/data sets that included race and ethnicity data. (When the Census Bureau changed their taxonomy for collecting race and ethnicity data, it created monumental problems for comparability of data, both across time and across different information systems, since the old and new data content standard did not match, the new being much more complicated and nuanced (although arguably more accurate.))

The definition of terms used in the CHS Performance and Practice Measurement System (PPMRS) and for the Environmental Health Knowledge Management project are two more recent examples. As a state, we cannot have comparable—and so meaningful—statewide data if every agency defines “visit” or “inspection” differently.

Though not listed above, HIPAA has established a national floor for **privacy and security standards** which are of vital importance to health care and

public health but which are beyond the scope of this paper. More information can be found at

<http://www.hhs.gov/ocr/hipaa/>.

Why do we need data standards in public health?

The critical need for data standards rises from factors both internal and external to public health. Chief among them are:

- The demand for a more efficient and responsive public health system that uses its data as a resource to improve community health and public health practice.
- The increasing need to exchange data across public health information systems in order to create more complete and integrated profiles of clients, families, and communities.³
- The increasing need to exchange data with hospitals and private providers, as well as with jails, state agencies, other local health departments, long term care facilities, and others.
- The frustration of working with silo information systems that cannot readily exchange data, often don't support quality care/services or improvements in public health practice, are inefficient, and make comprehensive community assessments difficult.
- The fact that the health care industry is moving rapidly, through both mandates and market forces, toward an increased adoption of standards. Since they are the source of much of public health's data, we need to

³ It should be noted here that integrating data from several sources does not imply integrating those data into a single database. But an integrated *view* of data is possible when data standards are used to merge data from different sources into a single report or profile.

ensure our information systems can readily accept and exchange that data.

What do data standards mean to me as a public health professional?

As we increasingly move toward broader use of data standards, public health agencies will benefit in a number of ways:

- Greater continuity of care because you will be able to exchange clinical data with private providers on clients you are both serving. This also means that public health can participate in the regional health information exchanges that are beginning to emerge around the state.
- The ability to receive data from others without having to manually translate the data into a form and format that works for your information systems.
- More meaningful reports, because there will be more consistency in how data gets entered, merged and shared.
- More complete profiles of clients, families and communities because data from different information systems can be consolidated and integrated view.
- Less need for double data entry, because data can be exchanged between information systems that include records on the same client. Because the information systems can ‘interoperate,’ the data from one system can be used to populate the other, saving data entry time (and reducing the chances of data entry errors).

How do I know if the applications I rely on use data standards?

It is an unfortunate reality that public health information systems do not historically rely on data standards to any great extent. Partly this is because many standards arose out of the need to process claims, and many public health services and their associated data are not reimbursable in the traditional sense.

The result is that we do not have codes for data such as client risk factors and symptoms, community coalition building activities, and many health promotion and prevention services.

How can I begin?

There are places you can start within your program or agency to move toward standardizing your information systems.

- Verify that the demographic fields in your applications match standards set by the [National Committee on Vital and Health Statistics](#). You may decide to begin changing any systems that do not meet this standard; for instance, by moving from a single name field to having separate fields for first, middle and last names. This is not a trivial task, and involves either writing a script (a short program written in computer language to perform a defined task) to move the last word in the single name field into Last Name, and any single letters into Middle Initial, or manually moving/reentering the data. Using an automated script requires careful review to ensure accuracy. You will want to ensure any new applications you develop or purchase match these demographic standards.
- Make sure your staff are entering data in a standardized way. It’s not uncommon for different staff to use the same fields in different ways or

to enter the same data in different ways. Exporting select data fields into a spreadsheet enables you to easily scan down each column to identify unacceptable variations in data entry, either by how data is entered or by inconsistent uses of a field. (Sometimes this is done intentionally because an application doesn't have a field for data the agency wants to collect, so they use an otherwise unused field. These are good ones to check for consistent use across all users.)

- Whether developing or purchasing an information system, seek to minimize the number of free text fields. Because there are few controls and ways to standardize what data gets entered how, the data is unlikely to ever be useful for exchange or reports. For instance, if you allow the hematocrit test to be entered as free text, you may get 'Hematocrit,' 'Crit,' or 'PCV'—data that is not easily used in creating a report on, say, the number of hematocrits run in the last month. Use picklists wherever possible to standardize data entry and minimize data entry errors.
- Ensure that any [content standards](#) used in your purchased applications are maintained and routinely updated by the vendor. Standards are generally driven by the user community, so changes are not uncommon. Verify that your vendor is using, or shortly plans to release, the latest version of a standard such as the [Omaha System](#).
- Ensure your lab and clinical data match the appropriate [content standards](#). For instance, your immunization screen should include the core data set established by CDC and adopted by the Minnesota

Immunization Information Connection. Lab results should match [LOINC](#) codes.

Who needs to care most about standards?

While in truth, adherence to data standards are the business of every public health staff person that uses information systems, some staff clearly have more responsibility around standards than others. These are:

- Anyone who enters data, to make sure the same data from different people is entered in a consistent way, so that any reports using that data—and exchanges of that data with others—are consistently meaningful.
- Anyone developing an application or information system of any size. Since you can never be sure what sort of life even a small and seemingly short-lived application is going to have, make sure to develop it using whatever standards are most appropriate.
- IT managers responsible for the operations, interoperability and security of agency information systems.

Where can I find out more?

1. The Public Health Data Standards Consortium (<http://phdatastandards.info/>)
2. The Center for Disease Control and Prevention's site on the Public Health Information Network (www.cdc.gov/phinfo)
3. Healthcare Information Technology Standards Panel: A national initiative to harmonize health care standards (http://www.ansi.org/standards_activities/standards_boards_panels/hisb/hitsp.aspx?menuid=3)
4. MDH Center for Health Informatics (www.health.state.mn.us/e-health)