



# HL7 Version 2 Messages

An Introduction

Toolkit for Planning an EHR-based Surveillance Program | [www.phii.org/EHRtoolkit](http://www.phii.org/EHRtoolkit)





## Presentation Outline

- Brief History of HL7 and Version 2 (V2)
- V2 message business process
- Message components and structure
- Vocabulary Constraints
- Role of Profiles in Improving Interoperability Success

The following topics are covered in this presentation:

- A brief review of HL7 as both a standard and a standards development organization
- A brief history of HL7 Version 2
- A high level review of the messaging process or paradigm and where Version 2 messages fit (describe the business process of V2 messaging)
- A review of how HL7 V2 messages are constructed
- The importance and role of vocabulary constraints and structured data
- The role of profiles that constrain the base standard and how that impacts the success of exchanging data between systems

This presentation addresses these aspects of HL7 at an *introductory* level only. **The intent is to give you enough background so you can participate in discussions about HL7 V2. To gain the technical expertise to really work with the standard requires training, which is available from the HL7 organization.**



## Brief History of HL7

- Founded as an international standards development organization in 1987 to promote communication between hospital data systems
- Stated a goal of creating a platform independent method of moving data between different systems
- Developed grammar for messaging and standardized vocabulary

HL7 is an international standards development organization (SDO) that was established to enable interoperability of health care information. Initially it focused on interoperability among information systems within large hospitals. Later, the organization began focusing on interoperability among systems in disparate organizations, including public health.

Early initiatives of the organization included developing grammar for messaging and a standardized vocabulary

It is not the only standard that is used for transmitting health related data. Others include:

- NCPDP (National Council for Prescription Drug Programs) for ordering medications
- EM TEP (OASIS Emergency Management Tracking of Emergency Patients) for tracking health information for patients in transport

While both NCPDP and EM TEP are specialized standards, they can be mapped to HL7 concepts.



## HL7: The Primary Standard for Communicating Health Data

- ANSI Standard for clinical interoperability
- HL7 standards are widely adopted and continue to evolve
- Meaningful Use has identified a number of HL7 standards to support sharing data between systems
- HL7 website: <http://www.hl7.org/>

HL7 is considered *the* standard for communicating health data.

As an ANSI-approved standard, it has gone through rigorous validation and approval process.

HL7 standards for the exchange, management and integration of electronic health care information continue to evolve.

The use of HL7 received a big boost with the enactment of the “Meaningful Use” program. This included use of specific HL7 implementation guides developed by public health for lab results, cancers, syndromes and immunizations.



## HL7 Versions

- HL7 version 2.x messaging
- HL7 version 3 messaging
- HL7 Clinical Document Architecture (CDA)
- HL7 Fast Healthcare Interoperability Resources (FHIR)

Because HL7 continually evolves with use and experience, multiple HL7 versions now exist as you can see by this list.

This presentation focuses on version 2 messaging only. CDA and FHIR are presented in another EHR Toolkit presentation.

The HL7 organization released version 2 messaging decades ago. This version is still being updated and improved and is widely implemented in the U.S. In addition to the current Meaningful Use requirements around version 2 messaging, version 2.x supports unsolicited updates such as new information sent to be added to a case report. Version 2.x also supports query and response—for example, an EHR system requesting and receiving an immunization history from a registry.

HL7 Version 3 messaging has been implemented widely internationally, but not in the U.S.

CDA is widely adopted in the U.S. and is in use with Consolidated-Clinical Document Architecture (C-CDA)

A key distinction between HL7 messages and HL7 CDA documents is that messages are packets of data sent from one system to another, generally for incorporation into the receiving system. In comparison, documents are basically electronic versions of physical documents.

Public health largely relies on version 2.x messages, although CDA has been piloted in some areas, including reporting for cancer, fetal birth and deaths.

FHIR (pronounced “fire”) is just emerging, but appears to be easily implemented and may be the wave of the future. It can support Version 2, Version 3, and CDA paradigms.

For more information about the HL7 CDA version, see the companion presentation in the EHR Toolkit.

## HL7 Version 2 History

- HL7 Version 2 first published in 1989
- Each version of Version 2 builds on the previous versions with a better data model and stronger vocabulary standards
- The Normative Version specifications have evolved
- The current Normative Version is Version 2.7
- New versions are intended to allow backward compatibility



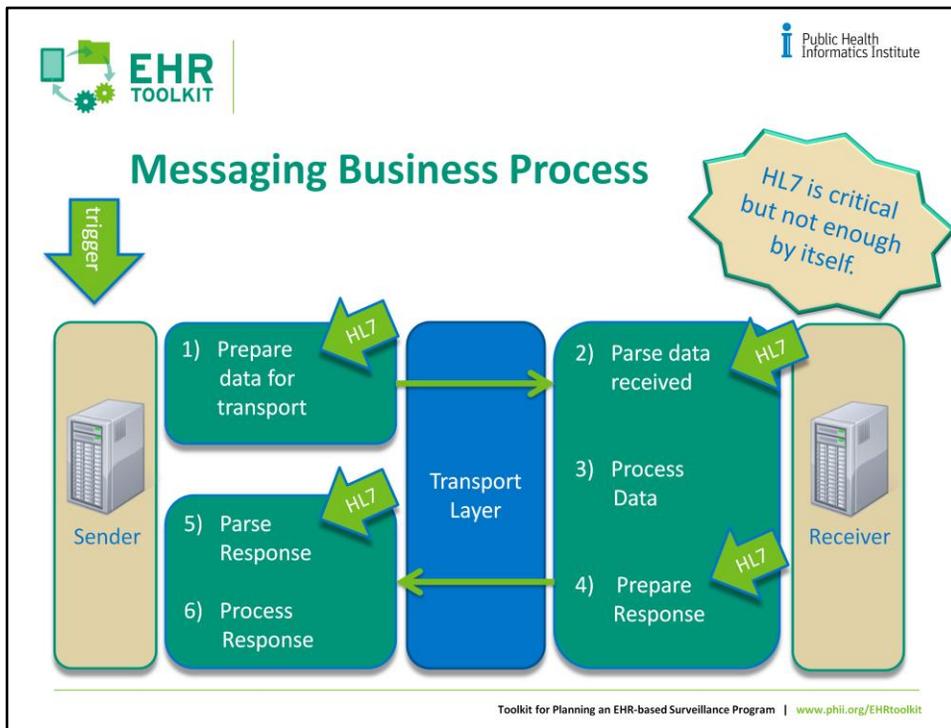
The first messages between systems tended to be sent and received within the same organization. For example, messages sent between patient administration and the lab.

The standard evolved by accretion; that is, upon being identified, a new field would be appended to the existing segment. As the standard has matured, several fields have been deprecated. For example, in the past the standard had fields for patient ID, social security number, drivers license number and others in the patient identification segment (PID). Now the standard has a field for patient identifier that may repeat and hold a list of various identifiers.

As the HL7 community has grown more experienced and technology has progressed, a better organized data model has also evolved. The comprehensive nature of the work undertaken to develop Version 3, largely in parts of the world other than the U.S., has also influenced the evolution of Version 2.

Backward compatibility, such as a Version 2.3 message being readable by a machine set up to send and receive Version 2.5 messages, is a goal, but is not always accomplished.

Many current implementations in the public health space are based on Version 2.5.1, but they also adopt, in advance, some features of Version 2.7.



This picture provides a simplified graphical overview of what interoperability entails, showing the possible connection between sender and receiver systems. Interoperability can be more complex—for example, when one system talks to another through a health information exchange, or HIE.

And as noted on this slide, HL7 is critical to sharing messages between sender and receiver systems, but is *not* enough on its own.

HL7 is responsible for both semantic and syntactic interoperability (see the tools under the Analyzing Technical Options section of this presentation). Semantic standards refer to the vocabularies used to denote data elements such as events, lab results, and conditions. Syntactic standards refer to how the sender system packages those data and associated vocabularies before sending.

In addition, the transport layer refers to the protocol for connecting one system to another. Transport standards are different than semantic and syntactic standards, and the transport layer is “agnostic” to the semantics and syntax (the content and format) of the data being transported.

In looking at the **seven** steps depicted in this slide, you see where HL7 contributes and where it doesn't.

In addition, the transport layer refers to the protocol for connecting one system to another. Transport standards differ from, and are content and format agnostic to, the semantic and syntactic standards of the data being transported.

The steps depicted in this graphic point out where HL7 does and does not contribute:

0) Something in the real world causes the process to start. It is called a *trigger* event. The trigger could be any number of activities, such as a person clicking a button on the sending system that requests a query of another system, or entering a new immunization into the system.

- 1) First, the sender prepares the data for transport. This means that they extract the needed data and package it for transport. They apply the specified HL7 standard to create this package of data in a specific and predictable way.
- 2) Next, the sender connects to the receiver through the transport layer. This step includes authenticating that the sender may send data.
- 3) Third, the receiver gets the package of data and parses it (translates it from HL7 to an internal format).
- 4) In the fourth step, the receiver processes the data, applying local business rules and data hygiene. While this is not a part of the standard used, it is an important part of the process and can cause issues if not clearly documented by the receiver and communicated to the sender.
- 5) Then, the receiver acknowledges the receipt of the package of data and indicates whether it was successful or not.
- 6) Finally, and crucially, regardless of the number of **steps** between sender and receiver, the response must return to the sender. For instance, if the data pass through an HIE, the receiver must return the acknowledgement through the HIE to the initiating system. This allows the sender to be informed of the outcome, and is especially important for helping detect problems on the receiving side.

## Version 2.x: Messages

- Each message has a job and one or more triggers
  - Send vaccination history (VXU)
  - Send demographic update or transfer patient (ADT)
  - Send acknowledgement (ACK)
  - Request immunization history (QBP)
  - Return immunization history (RSP)
- The “job” and structure are identified by the name and the trigger.

The name of an HL7 Version 2 message is captured as a 3 character string, for example, “ACK.” The trigger is typically an event, like a patient being transferred from one facility to another or a patient receiving an immunization.

The list in this slide includes a few of the many possible examples of Version 2 messages:

Some messages, like VXU, have one trigger and one format. Other messages, like ADT, have more than one trigger, and each trigger may have a unique format. This format may include different types of data and have a different structure. We will discuss how messages are composed later in this presentation.

Unlike flat file structures, in which fields have fixed positions and fixed sizes, the Version 2 message permits the size and content of the message to expand or contract to carry only the data required. This is accomplished by allowing:

- Some segments to be optional, repeat, or both.
- Some fields to be optional, repeat, or both. For example, the NK1 (next of kin) segment is usually optional and repeating. When no information about a next of kin is known or relevant, the message includes no NK1 segment. When the Mother, Father, and the Grandmother are important content in the message, each would have an NK1 segment.

- A person may have more than one identifier, such as SSN or medical record number. The field holding the person identifier allows repeats, but requires that the identifier type and assigning authority be included.



## Version 2.x Messages Are Composed of Segments and Segments Are Composed of Fields

- Segments are the “sentences” of the message
- Each has a job (PID is the personal identifier segment)
- Segment name (e.g., PID)
- Each segment is composed of fields
  - PID|field1|field2|person identifiers|
- Fields are separated by delimiters (ASCII characters or XML tags)
- Fields may be optional, repeat, or both
- Empty fields have place holders
- Each field has a data type specification

Segments are like sentences in a message, or rail cars in a train.

Each segment has a name that identifies what it does, or its “job.”

Each segment also has a position in the message. Each instance of a segment is composed of fields, which carry the data.

The fields have fixed positions, but some may be optional and some may repeat. This is one way the message may flexibly carry varying amounts of information, depending on a specific situation.

Most systems use a standard set of separators.

Each field starts with a “pipe” denoted by a straight line “|” (that looks like a long bar).

Each field has a fixed position. If a field is empty, the pipe maintains its position.

Each field has a data type. A data type specifies the structure (to be discussed below).

## Delimiters Separate Components

Type	Character
Segment Name	3 character code
Segment Terminator	Carriage Return (ASCII 13)
Field Separator	(ASCII 124)
Component Separator	^ (ASCII 94)
Repeat Separator	~ (ASCII 126)
Escape Separator	\ (ASCII 92)
Subcomponent Separator	& (ASCII 38)

Each segment begins with a *segment name* and is terminated by a *carriage return*. Segment lines may be very long.

Each field begins with a *field separator*.

Each component in a field is separated by a *component separator*.

Fields that may repeat have a *repeat separator* between the repetitions.

The *escape* character delimits a reserved character (any of the separator characters). For example, if a name includes an ampersand, it would need to have the ampersand “escaped”.

Some data types are composed of other complex data types and therefore have subcomponents. The ampersand character separates such subcomponents from the component.

## ACK Message Specification

### HL7's Definition:

“The simple general acknowledgment (ACK) can be used where the application does not define a special application-level acknowledgment message or where there has been an error that precludes application processing.”

Segment	Usage	Cardinality	Notes
MSH (Message Header Segment)	Required	(1..1)	
SFT (Software Segment)	Optional	(0..1)	
MSA (Message Acknowledgement Segment)	Required	(1..1)	
ERR (Error Segment)	Required but may be empty	(0..*)	

Toolkit for Planning an EHR-based Surveillance Program | [www.phii.org/EHRtoolkit](http://www.phii.org/EHRtoolkit)

To better understand the message, let's look at an example of a simple message type:

Every message has a name and definition. In this case, the name is the ACK, or “Send Acknowledgement” message.

The definition included in the slide is quoted directly from the HL7 standard.

ACK messages inform the initiator of a message exchange that the message was received and indicate if it was successfully handled. An ACK message has four segments:

- First, is the Message Header (MSH) and is a required segment. It indicates what type of message it is, where it came from and where it is going, among other things.
- Second, is the Software Segment (SFT), which is optional and rarely used. This segment indicates what software generated the message.
- Third is the Message Acknowledgement Segment (MSA), which is required in an ACK message type. It indicates the overall success of the message. In other words, was it rejected for specific HL7 reasons, did it have errors, or was it accepted with no issues?
- Fourth, is the the Error Segment (ERR), which is optional and repeating. This

segment carries a list of the errors that were encountered and their severity level.

Cardinality represents the minimum (the first number in the table above) and maximum (the second number) number of values that can exist in a given element of a HL7 message. The cardinality of the MSH segment means that exactly 1 MSH segment must be present in the message. The cardinality of the SFT segment means that it is optional, and that if present, no more than one instance is permitted. Finally, the ERR segment may have zero or an unlimited number of instances. If a system would only accept 5 or fewer values for a given element, the cardinality would be (0..5).



## Example ACK Message

```
MSH|^~\&||DCS|MYIIS|PHS||20090604||ACK^V04^ACK|  
9299381|P|2.5.1||NE|NE<CR>
```

```
MSA|AA|400586704<CR>
```

This message indicates that the VXU (unsolicited immunization history) message was successfully received and processed. You will note that the optional segments (SFT and ERR ) are not present. Note that a <CR> is inserted to show where a carriage return would occur. This helps with human readability, but is not part of the real message and would NOT be sent.

Note that the MSH segment identifies the field separator (|) and the other separator characters (^~\&). Although the standard permits other characters to be identified and used, *no one* does this.

Next we see a code for the sending organization followed by the receiving application and receiving organization.

The date the message was created and sent follows and has a format of YYYYMMDD.

The ACK^V04^ACK tells the receiver the name of the message (ACK), the trigger event (Received VXU^V04), and the structure of the message (ACK).

The unique id of this ACK message is 9299381. It is sent as a production (real) message and is using HL7 Version 2.5.1

This message expects no acknowledgments (NE|NE)

The MSA segment indicates that the transaction was a success with the “AA” in the first field. The id of the original VXU message is returned in the MSA (400586704)

To understand the requirements for each segment and component, you must reference the standard or an implementation guide or profile.

## PID Segment (Partial)

Element Name	Data Type	Usage	Cardinality
Set ID	SI	Required	[1..1]
Patient ID	CX	Deprecated	[0..0]
Patient ID List	CX	Required, but may be empty	[0..*]
Alternate Patient ID	CX	Deprecated	[0..0]
Patient Name	XPN	Required	[1..*]
Mother's Maiden Name	XPN	Optional	[0..1]
Date of birth	TS	Required	[1..1]
Many more fields			



## Version 2.x: Data Types

- Atomic units of structured data
  - Think of these as fields in a database
- NM: Numeric (primitive data type)
  - For instance, 27.3
- CQ: Composite Quantity with units
  - Quantity (NM)
  - Units (CE)
  - Example: 27.3<sup>mL</sup>milliliters<sup>UCUM</sup>

Data types are the building blocks of messages. They structure the data so that both sides of a transaction can read/write using the same syntax. It is important that a data type supports the required data elements. A field that requires a number will not support a data element that is an alphanumeric code.

Data types structure the format of the data elements. For instance, a person name could be structured as one continuous string: "John R Johnson." Alternatively, it could be "John," "R," "Johnson." By specifying a common data type for each data element, sender and receiver share a common syntax in the message and can more readily process the data.

All data types are basically formatted strings of characters.

Primitive data types are the basic building blocks. They include strings (a group of alphanumeric characters), and numeric data types (a string of numeric characters).

Complex data may be represented by composite data types, which are composed of a group of subcomponents. These subcomponents have a data type in their own right. For example, the CQ data type has two subcomponents: the quantity (NM, a numeric data type) and the "Units" (CE, a Coded Element data type). The CE data type is composed of subcomponents that allow coded data to be included that are identified by a value set identifier. (See the next slide for more information.)

## Coded Element (CE) Data Type

Position	Requirements	Description
1	Required	Code
2	Required but may be empty	Explanatory text
3	Required	Code system identifier
4	Optional	Alternative code
5	Optional	Alternative explanatory text
6	Required if position 4 is populated	Alternative code system identifier

The CE data type is commonly used to carry coded entries.

The first position in the data type is a code that carries the meaning for the field. The explanatory text helps humans read and understand it. This text is not commonly used by computers. To ensure the code is correctly interpreted, it must be associated with a code system. Imagine a code in position 1 of 'Y'. What does this mean? It could mean 'Yes' or it could be the Yth value in a code table and mean a symptom or an observation. It's important to know to which code system "Y" refers.

It is crucial that if the alternative code is populated, it must mean essentially the same thing as the code in position 1.



## Profiles put Constraints on the Base Standard

- Base standard supports many needs
- Each need has different requirements
- Profiles are constraints that remove optionality and bind to vocabulary
- The profile in use is identified in the MSH segment of the message

The base standard includes everything that someone may want or need, but it must function as needed in different countries and health care systems. This means that many fields and components are optional. For example, the patient identifier has a field for Breed and for Species, which are important in veterinary medicine but have little meaning for human patients.

Profiles are built on top of the standard and can add constraints to the standard. This means that a field that is optional in the base standard may be required by a profile. It does NOT allow relaxing of existing constraints. For example, a required field in the standard for a message may NOT be relaxed to optional by a profile built on that standard.

Profiles may also specify the code system or value set that must be used for a particular field. For example, while many identifier type codes exist, only a few may be appropriate in a given situation. The profile can identify the subset to use.



## Additional Resources

- HL7 (hl7.org) has a variety of materials including tutorials where you can get the specifics for each of its standards.
- HL7's website is at: <http://www.hl7.org/>
- CDC's Vocabulary Access and Distribution Systems (PHIN VADS) can be found at <https://phinvads.cdc.gov/>

This presentation addressed these aspects of HL7 at an introductory level. For additional information, HL7 has a very robust website with many free offerings. For example, they have an Education Portal with free webinar recordings(<http://www.hl7.org/implement/courseList.cfm>). Some of these recordings include an introductory webinar for HL7 and an overview of the healthcare connection with HL7. These and other resources, such as tutorials where you can get the specifics for each of the standards, are available as well.

You can also get information from the CDC Public Health Information Network (PHIN) web site, which includes vocabulary and other standards related to public health CDAs

We encourage you to check out these and other HL7 resources!