

Facility and Patient Identifiers in regulatory agencies

Draft proposal, David John Marotta, DT7 Software

Summary

This provides a suggested convention for patient identifiers storage and transmission between various sources and state regulatory agencies.

Patient Identifier Background

The HL7 patient identifier found in PID-3, is a repeating field of the following format:

3.4.2.3 PID-3 Patient identifier list (CX) 00106

Components:

<ID (ST)> ^
<check digit (ST)> ^
<code identifying the check digit scheme employed (ID)> ^
<assigning authority (HD)> ^
<identifier type code (ID)> ^
<assigning facility (HD)> ^
<effective date (DT)> ^
<expiration date (DT)>

Subcomponents of assigning authority:

<namespace ID (IS)> & <universal ID (ST)> & <universal ID type (ID)>

Subcomponents of assigning facility:

<namespace ID (IS)> & <universal ID (ST)> & <universal ID type (ID)>

Most HL7 2.x messages only populate the following components at most: ID and identifier type code. Obviously, facilities that use a check digit also populate check digit and code identifying the check digit scheme employed. Occasionally the component assigning authority or assigning facility will be populated in a multi system facility.

HL7 2.x messages within a hospital environment can use the degenerate form of patient identifier (id + identifier type code) because within the hospital these two values form a unique key. The identifier type code is understood to imply the only system that generates that type of code, and the id itself is therefore unique.

Most state and federal regulatory agencies have a coding system they use to identify the clinics and hospitals from which they collect information. This facility code plus the normal patient identifier information in the HL7 message should uniquely identify the patient.

Since each regulatory agency uses a different coding scheme for facilities, it is important to add a universal identifier for each agency to identify which regulatory agency's coding scheme is being used to identify the facility. The universal identifier for each agency should be an ISO object identifier of the format 2.16.840.X.X.X.

The top level tree for ISO object identifiers can be found at <http://www.alvestrand.no/objectid/top.html>. The ISO Object identifier for Health Level Seven, Inc. (HL7) is 2.16.840.1.113883. Since each state

regulator agency assigns identifiers to hospital and clinic sources, each agency should have its own object identifier.

Examples

Example #1

Assume that the Virginia Department of Health is collecting information from various hospitals and clinics. It collects some information from the University of Virginia on a patient (John Doe) whose HL7 PID-3 looks like this:

```
PID|||1277586^^^MR~857677^^^AN~228273^^^VN|
```

Showing that according to the University of Virginia, John Doe has a medical record number of 1277586, an account number of 857677 and a visit number of 228273.

The Virginia Department of Health has a unique set of identifiers for the hospitals and clinics that it collects data from. In its table, the University of Virginia's identifier is UVAHOSP. Furthermore, the Virginia Department of Health has applied for and received an ISO object identifier of the format 2.16.840.X.X.X.

Given these values, John Doe would have the following patient identifiers:

```
1277586^^^MR^UVAHOSP&2.16.840.X.X.X&ISO  
857677^^^AN^UVAHOSP&2.16.840.X.X.X&ISO  
228273^^^VN^UVAHOSP&2.16.840.X.X.X&ISO
```

Example #2

Assume that Metropolis Hospital Conglomerate (MHC) in California is sending information to the California Department of Health. MHC's PID information looks like this:

```
PID|||28734^4^M11^REGSYS^MR^MHCFAC1~76857^4^M11^REGSYS^MR^MHCFAC2~  
5464765^^^PATCARE^MR^MHCFAC1|
```

The same software (REGSYS) at two different facilities within MHC (MHCFAC1 and MHCFAC2) assigned this patient different medical record numbers. Another system (PATCARE) at one of the facilities (MHCFAC1) assigned yet another medical record number. These complex patient identifiers usually only happen when several hospitals are combined into a single system. In this multi-facility environment it is best if the California Department of Health assigns different facility ids to each of the two facilities (MHCFAC1 and MHCFAC2). Assuming that the ISO object identifier for the California Department of Health is 2.16.840.X.X and they have assigned facility ids to these facilities of CA0345 and CA0722 then it would result in the following patient identifier information:

```
28734^4^M11^REGSYS^MR^CA0345^2.16.840.X.X^ISO  
76857^4^M11^REGSYS^MR^CA0722^2.16.840.X.X^ISO  
5464765^^^PATCARE^MR^CA0345^2.16.840.X.X^ISO
```

In this information, the unofficial and local designation of facility (MHCFAC1 and MHCFAC2) with the facility ID assigned by an organization with a universal ISO object identifier (in this case the California Department of Health).

Recommended database capabilities

In order to utilize this encoding to its fullest, regulatory agencies must be able to store multiple patient identifiers for each patient. Use of this encoding will facilitate, but not solve problems of deduplication. Once two patient identifiers are found to represent the same patient, they can be linked in the agency database as synonyms. Storage capabilities should allow merging and unmerging patient identifiers, and the ability to tie result reporting to a specific patient id in case incorrect patient identifier alias are later separated again.

Additionally, storage of facilities should allow facility id aliases as well. A facility providing information to one state may also provide information other states as well. Each state may have its own designation for the facility. In the same way, sending and receiving facility (MSH-4 and MSH-6) should use a fully populated hierarchical designator from an ISO object identifier source (probably in our case the CDC).

Example

If the ISO identifier for the CDC is 2.16.840.X.X, then the MSH for a message being sent from the George State registry to the California State registry would have the form:

```
MSH|^~\&||GA0000^^2.16.840.X.X^ISO||CA0000^^2.16.840.X.X^ISO|...
```

Benefits

Short of a universal patient identifier or a universal facility identifier this methodology would uniquely and universally identify patients. It would allow regulatory agencies from different states to identify patients uniquely and universally such that deduplication processes could be used to identify patient identifier aliases and mark them so that future results reporting would be handled automatically. It would allow the same form of aliases to be constructed for facilities as well.