Applying the HL7 Reference Information Model To a Clinical Data Warehouse

Jason Lyman, MD, MS
Sandra Pelletier, PhD
Ken Scully, MS
Department of Health Evaluation Sciences,
University of Virginia,
Charlottesville, VA, USA
lyman@virginia.edu
slp6u@virginia.edu
kws9s@virginia.edu

James Boyd, MD
Department of Pathology
University of Virginia,
Charlottesville, VA, USA
jboyd@virginia.edu

Jason Dalton, MS
Steve Tropello, MS
Department of Systems Engineering,
University of Virginia,
Charlottesville, VA, USA
jrd7y@virginia.edu
spt3a@virginia.edu

Csaba Egyhazy, PhD
Department of Computer Science,
Virginia Polytechnic Institute and State University,
BlacksburgNorthern Virginia Center, Falls Church, VA, USA
cegyhazy@vt.edu

Abstract

Large-scale data integration efforts to support clinical and biologic research can be greatly facilitated by the adoption of standards for the representation and exchange of data. As part of a larger project to design the necessary architecture for multi-institutional sharing of disparate biomedical data, we explored the potential of the HL7 Reference Information Model (RIM) for representing the data stored in a local academic clinical data warehouse. A necessary first step in information exchange with such a warehouse is the development and utilization of tools for transforming between local data schemas and standards-based conceptual data models. We describe our initial efforts at mapping clinical concepts from a relational data warehouse to the HL7 RIM.

Keywords: Data Mining and Management, Decision Support Systems, Informatics, Knowledge Sharing

1 Introduction

The development and adoption of standards to support meaningful information representation and exchange between disparate systems is a fundamental goal within the medical informatics community. Health Level Seven (HL7) is an important ANSI-accredited standards development organization in the healthcare domain, with a specific focus on clinical and administrative data[1]. The HL7 Reference Information Model is a developing standard used to represent the entire information content relevant to the HL7's efforts. We have undertaken a multi-disciplinary, multi-institutional project focused on the high-level design of a system for integrating disparate biomedical data to facilitate research, health promotion, and quality assessment. One critical aspect of such an integrative project is the exploration of issues related to data transfer between systems. Exchanging information in meaningful ways requires a common syntax, shared vocabularies for unambiguous concept representation, and agreement on how concepts are inter-related. The overlapping domains of clinical medicine and biomedical investigation are filled with ambiguous terms, complex information, and rapidly changing technology. A lack of standardization in terminology, knowledge representation, and data structures has impaired integrative efforts in the past, though progress has been made on many of these fronts[2][3][4].

Our approach includes careful consideration of how data can be exported from source systems in a manner that is easy for receiving sources to interpret into their own database schema. The exported data is mapped into the RIM, and sent in XML syntax. The receiving source parses the XML, and using its database schema to RIM mapping transforms the transmitted data into local terms and structures. The RIM, with its flexibility, robustness, and comprehensiveness, offers great potential for serving as the "on the wire" representation of clinical information in the healthcare domain. To explore this potential, we attempted to map from a local clinical data warehouse containing clinical and administrative patient data to RIM-based classes, exporting data in XML format. This represents an ongoing project, and preliminary issues related to the application of the HL7 RIM are presented in this document.

2 Background

2.1 The UVa Clinical Data Repository (CDR)

The CDR is a unique information resource at the University of Virginia Health System that allows researchers, clinicians, administrators and students to perform population-based queries on anonymized UVa patient data[5]. It is a WWW-enabled data warehouse designed to facilitate research, quality assessment, and

0-7803-7952-7/03/$17.00 © 2003 IEEE

medical education by allowing direct access to retrospective administrative, financial, and laboratory data. Data is received in both real-time over the hospital's network (e.g., laboratory data), as well as in periodic batch updates (e.g., administrative data). Its custom-built user interface allows users to set a large variety of conditions by generating ad-hoc SQL to perform dynamic queries on the underlying relational database (Figure 1). Once a population of patient encounters is generated from a query, users can download the data to their own computer for analysis, or select from an assortment of standard reports.

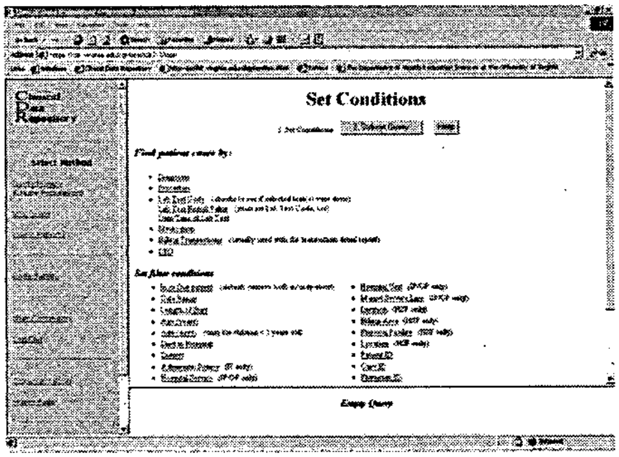


Figure 1: Setting Conditions in the CDR

The CDR currently runs in a Linux environment and uses the Sybase relational database management system (ver. 12.5). Information on approximately ten years of patient visits, both inpatient and outpatient, are included, with over 650,000 patients represented. The data model underlying the CDR is relatively straightforward, with most information centered on the concept of a "visit", "case", or "encounter". These terms will be used interchangeably throughout this document, recognizing their inherent ambiguity. The vast majority of the time, this refers to a direct, face-to-face interaction between a physician (or other healthcare provider) and a patient, during either an inpatient, ambulatory, or emergency department encounter. A simplified entity-relationship model of the CDR is shown in Figure 2. The two relationships between *Visit* and *Diagnosis* (and *Visit* and *Procedure*), depict a visual representation of the following rules:

- Each visit is associated with one and only one *principal* diagnosis (the primary reason for the encounter), but potentially many secondary diagnoses.
- Each visit is associated with one (or zero) *principal* procedure, but potentially many secondary procedures.

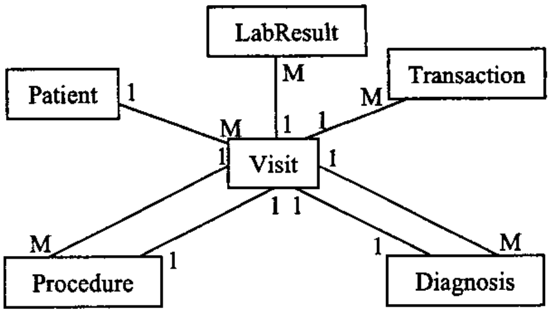


Figure 2: Basic CDR Data Model

The *Transaction* entity shown above includes administered medications, respiratory care services, and other such activities. One over-simplification of the above model (done to reduce complexity of the diagram), is that the CDR also relates the patients directly to diagnoses, procedures, and laboratory results, in a 1-M fashion in each case. This relationship is not explicitly represented in the model above.

As in most data warehouses, the data structures within the CDR are optimized for query performance. The tables tend to be de-normalized (e.g. patient demographic data is stored with each visit) and include many derived attributes (age, number of previous visits, days until next inpatient admission).

2.2 The HL7 Reference Information Model (RIM)

The Health Level Seven (HL7) organization has been responsible for many of the successful standards development efforts within the healthcare domain. HL7's most well known work has been in the development of messaging protocols that are in widespread use by healthcare organizations (including the University of Virginia Health System) for transmitting information between disparate systems (e.g. data exchange between a laboratory information system and a results reporting system). Other HL7 specifications include the Clinical Document Architecture[6] (CDA) for the exchange of documents such as discharge summaries and radiology reports, and the Arden Syntax[7] for representing medical knowledge to assist clinical decision support.

The HL7 Reference Information Model is the central information model around which all HL7 development activities occur. The RIM is maintained under a formal development methodology with consensus building among many technical committees and special interest groups within the HL7 organization along with regularly scheduled balloting sessions. RIM developers have adopted an object-oriented approach, expressing the model using the Unified Modeling Language (UML).

The RIM has four primary components: the *classes* which make up broad categories of information, *relationships* between classes, *attributes* of the classes, and *vocabulary* domains for content representation of the classes and attributes.

2.2.1 RIM Classes

The "backbone" of the model is six abstract classes: *Act*, *Entity*, *Role*, *Participation*, *ActRelationship*, and *RoleLink* (Figure 3). An *Act* is any actionable event, including an inpatient administration, the ordering of a medication, and an observation (e.g., a lab result or diagnosis). The *Participation* class allows the explicit representation of how a given *Entity* (a person, place, or thing) functions in a given *Act*. When an *Entity* participates in a given action, it does so acting in a specific *Role*, for example, a patient or healthcare provider. In addition, the *ActRelationship* and *RoleLink* classes allow for explicit representation of how specific instances of *Act* and *Role* inter-relate, respectively. For example, two instances of *Act* might be (1) an outpatient visit with a pediatrician, and (2) the administration of a hepatitis B vaccine. These two events, or acts, are related to each other, in that the latter occurs in the setting of the former. *ActRelationship* allows this relationship to be explicitly represented. As a result of this high degree of abstraction, the RIM has the potential to be extremely robust and comprehensive.

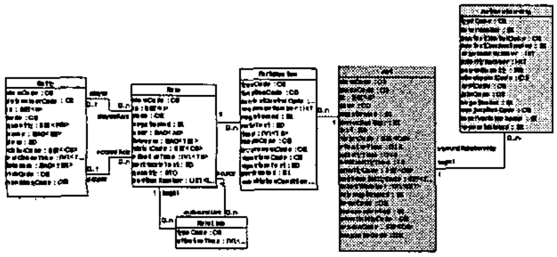


Figure 3: The Six Core Classes of the RIM
©HL7, 2003

Act, *Entity*, and *Role* have subclasses that provide further specialization while inheriting all of the attributes of their "parents". An example of this hierarchy is shown in Figure 4, which focuses on the *Entity* subject area. Specific subclasses for this class include LivingSubject, Organization, Material, among others.

2.2.2 RIM Relationships

Generally speaking, there are two types of relationships in the RIM: generalization relationships and association relationships. Generalization relationships, or "is-a" relationships, occur between subclasses and superclasses. As shown in the diagram, *Person* "is-a" *LivingSubject*, which "is-a" *Entity*.

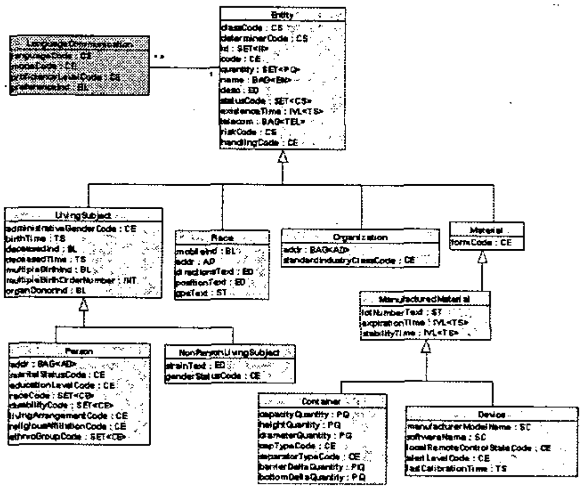


Figure 4: The *Entity* Subject Area in the HL7 RIM
©HL7, 2003

Association relationships, on the other hand, occur "across" the model, between the six core classes described above.

2.2.3 RIM Attributes and Vocabulary Domains

Each class is associated with attributes that allow for specific information to be represented about each instance of a class object. The generalization hierarchy means that subclasses inherit their attributes from their superclasses. For example, the attributes of *Person* include address, marital status, and race, but not date of birth. This last bit of information is an attribute from the superclass *LivingSubject*. The benefits of this include a model that is more "pure"; the alternative is fewer classes but more attributes that aren't always appropriate for a given object. The disadvantages are a model that is more challenging to both read and map to, as will be discussed later.

Three attributes in particular are critical for understanding how concepts are modeled in the RIM: *classCode*, *code*, and *moodCode*. The first two attributes are found in the *Entity*, *Role*, and *Act* classes, and serve to characterize exactly what concept is being represented. The *classCode* provides a broad categorization, while *code* allows greater specialization. Each of these attributes is associated with a particular vocabulary domain depending on the class in which it resides. In the *Entity* class, the value for *classCode* is drawn from the *EntityClass* vocabulary domain, which is a hierarchical terminology with such concepts as "animal", "plant", "public institution", "chemical substance", "place", etc. To further characterize a given object, the *code* attribute in this class draws from the *EntityCode* vocabulary domain, with such values as "package", "strip" "kit", "bed location", etc.

The *moodClass* attribute, found only in the *Act* class, reveals much of the true flexibility of the model. Since the HL7 standards development efforts seek to facilitate both the exchange and documentation of healthcare information, it is important that concepts be represented at many different stages of completion. As an example, a physician might document in a patient chart the fact that the goal for that patient's systolic blood pressure is 130mm Hg or below. That same provider might document during that same visit that the measured systolic blood pressure is, in fact, 138 mm Hg. Both concepts involve the same type of measurement, the same patient, the same provider, and the same encounter, yet mean very different things. The *moodCode* allows the distinction between these concepts by allowing the explicit representation of the stage of an event, from consideration to planning to ordering to completion. Note that each of these is its own instance of an *Act*. This attribute draws its values from the *ActMood* vocabulary domain, with example values of "order", "promise", "intent", "appointment", and "event".

The vocabulary domains utilized in the RIM are a combination of externally derived terminologies (such as LOINC for laboratory results and other clinical observations[2], or ICD9-CM, primarily used for clinical diagnoses), as well as internally developed domains.

HL7 RIM developers have been quick to state that the RIM is an *information model*, not a data model. It lacks the detail required for direct implementation, and is not intended for such purposes. As our current effort began, it quickly became evident that it was not simply a matter of considering each RIM class as a relational table and using the model directly to build a relational database. Instead, it seemed more appropriate to ensure that a given database could map its core content to the standard model. If it could, then any data exported by the local system could be exported in RIM-compliant format.

3 Mapping to the RIM

To explore the potential for mapping from our local data warehouse to the RIM, we adopted two concurrent approaches. The first was characterized by a table-by-table, field-by-field examination of the CDR to identify where analogous concepts resided in the RIM. The second involved creating selected clinical scenarios to explore how they were represented in both systems. Based on our initial results, one of the authors (JD) developed an XML-based tool for exporting data from the CDR in RIM format.

Mapping efforts were primarily performed during group working sessions where consensus could be developed. Participants included all of the authors, whose expertise spanned clinical informatics, bioinformatics, laboratory medicine and standards, systems engineering, data fusion, and computer science. Two of the authors (JL, CE) are members of the HL7 organization, and two authors (JL, KS) are directly involved in the CDR.

3.1 From CDR *Visit* Table to the RIM

The *Visits* table in the CDR is one of the core tables in the system, containing one row per patient visit, including both inpatient admissions as well as ambulatory encounters. The *Visit* table contains almost seventy variables, combining unique identifiers for the visit, patient, and providers, demographic information about the patient in question (e.g. age, gender, race, zip code), administrative / clinical data (e.g. principal diagnosis and procedure), utilization data (e.g. length of stay, hours spent in an ICU), charge / cost data, and date / time data. As mentioned earlier, many of the fields are derived (e.g. patient age) or de-normalized (date, calendar year, calendar quarter) to improve performance. The *Visit* table was a reasonable place to begin our mapping, since it represents a core table in the CDR and houses a fairly broad range of data.

3.1.1 Mapping a Visit

The representation of the central concept of a visit itself, uniquely identified with our *caseid* field, requires an instance of the *Act* class in the RIM. The instance of *Act* required to represent a visit in the CDR uses a *classCode* with value 'ENC', short for encounter. The optional use of the *code* attribute could provide additional classification, such as 'inpatient encounter', 'emergency', or 'oncology' as needed for a given visit. The *Act* class contains about eighteen other attributes, allowing the representation of priority status, time period, reason for the visit, and several other concepts. Many of these were not utilized in our mapping effort because the information was not available in the CDR (e.g. priority status). Attributes that were clearly appropriate for our mapping included *id* (the unique identifier for a given *Act*) and *effectiveTime* (which is obtained from the admission and discharge date for a given visit).

The subclass *PatientEncounter* offers the ability to represent additional details about the visit, such as referral source, length of stay, and discharge disposition. The latter two attributes were used to map directly from identical concepts in the CDR. Other attributes, like *preAdmitTestInd*, to allow for the indication that preadmission tests are required, were not utilized since this information is not contained within our data warehouse.

3.1.2 Mapping a Patient

The *Visits* table also contains several fields that characterize the patient involved in the encounter. To represent this information in the RIM, the *Entity* and *Role* subject areas are required, as shown in Figures 4 and 5. Three classes are used to represent the relevant information associated with a human being (*Person*, *LivingSubject*, and *Entity*). Since a person can act in many different capacities during a particular healthcare "event", the *Role* subject area allows the representation of a variety of 'roles', including patient (Figure 5).

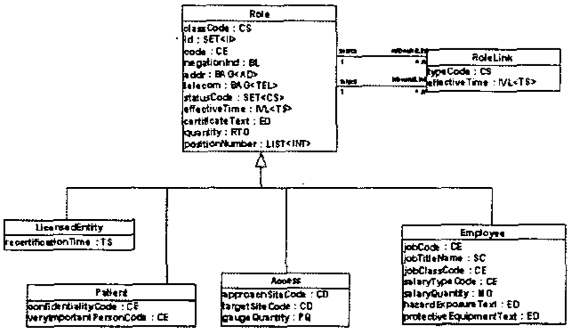


Figure 5: The *Role* Subject Area in the HL7 RIM
©HL7, 2003

The RIM requires that a given instance of a *Role* class have relationships with two members of the *Entity* class: one that is the "player" of the role, and one that is the "scoper". In the example of a patient role, the player is the individual acting as patient, while the scoper is the health care organization proving care.

In the *Visits* table, patient age, gender, race, zip code are stored, as well the *ptid*, the unique identifier for the patient. Representing these fields requires all of the classes mentioned above. The *Entity* class allows the representation of a unique identifier (*id*), and other attributes relevant to the patient but not stored in the *Visits* table (e.g. name, phone number or other telecommunications address). The *LivingSubject* class includes attributes necessary for identifying gender and date of birth, while the *Person* class allows the designation of address and race.

The *Role* class, necessary for designating our person as a patient at UVaHS, allows the representation of our patient identifier (different from the *Entity* identifier which should be constant for a person for their entire life – a person might have a different medical record number for every organization from which they receive care). The *classCode* in *Role* adopts a value of "patient" for our mapping. Interestingly, our mapping effort does not require using the *Patient* subclass, since the only added attributes in that class are *confidentialityCode* and *veryImportantPersonCode*, neither of which are concepts in our data warehouse.

3.1.3 Mapping a Visit Diagnosis

Clinical information within the CDR *Visits* table, such as principal diagnosis and principal procedure (there is one each per encounter, required for reimbursement purposes), requires additional instances of the *Act* class. The representation of a diagnosis in the RIM involves an *Act* instantiation with *classCode* value of "Observation" and *code* value of "admitting diagnosis", "discharge diagnosis", or "intermediate diagnosis". The diagnosis itself is represented in the *Observation* subclass, using the *value* attribute.

3.1.4 Putting it Together

To connect our patient information to the visit information, the *Participation* class is used. This object forms a bridge between *Role* and *Act*, resolving the many-to-many relationship between those two classes into two one-to-many relationships. Each instance of a *Participation* object represents a given entity playing a specific role engaging in an act. There are multiple different types of participation (subject, performer, donor, admitter, beneficiary, etc.). In the CDR *Visits* table, multiple "entities" exist, including the patient, one institution (by implication), an admitting physician, a referring physician, a discharging physician, a payor, and others. In mapping to the RIM, each of these "players" would warrant its own instance of a *Participation* object linking their specific *Role* to the *Act* (the visit) in question.

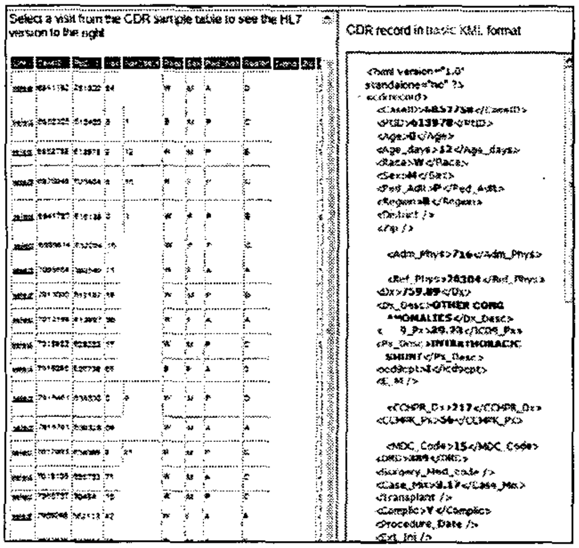
Connecting the visit to the diagnosis requires the *ActRelationship* class. This class allows the explicit representation of how two *Act* instances relate to each other. For example the principal diagnosis described earlier is, in effect, the diagnosis that caused the encounter to occur. This can explicitly designated by the *typeCode* attribute in the *ActRelationship* class, by giving the value of "CAUS", short for "is cause for".

Thus, the inherent flexibility, comprehensiveness, and robustness of the RIM leads to increased complexity when mapping from a relatively simplistic data model. Mapping individual patient data in one row of the CDR's "Visit" table requires instances of at least eight RIM classes: three for a person (*Entity*, *LivingSubject*, *Person*), two for an organization providing care (*Entity*, *Organization*), one for the representation of patient status (*Role*), one for the visit itself (*Act*), and one to connect the patient to the visit (*Participation*).

3.2 An XML Mapping Tool

Our approach also included the development of a prototype for converting CDR records into RIM-compatible structures. This tool contains three parts. The first is a database connection object. In this case, the testing data was housed in a Microsoft Access database in the same format at the native CDR data warehouse. An

ADO connection to the database was used so that the conversion code would be independent of any database system. ADO database objects connect to all major database systems and present a consistent interface to those sources. From the database, the XML Document Object Model (XML-DOM) implemented on a Windows 2000 server was used as the XML engine. For each record in the database, the fields were converted into an equivalent XML element. In this way, a generic script can be used to transform any database record into an XML format that is equivalent to the fields in that record (Figure 6a). From this point, an eXtensible Stylesheet Language (XSL) transformation was developed to modify the XML record into the HL7 RIM format (Figure 6b). This process generates a RIM compatible structure and can then forward it to a receiving service. We intend that this service will then use the incoming message to populate a cross facility database to store and accurately link records among different biomedical data sources.



Select a visit from the CDR sample table to see the HL7 version to the right

CDR record in basic XML format

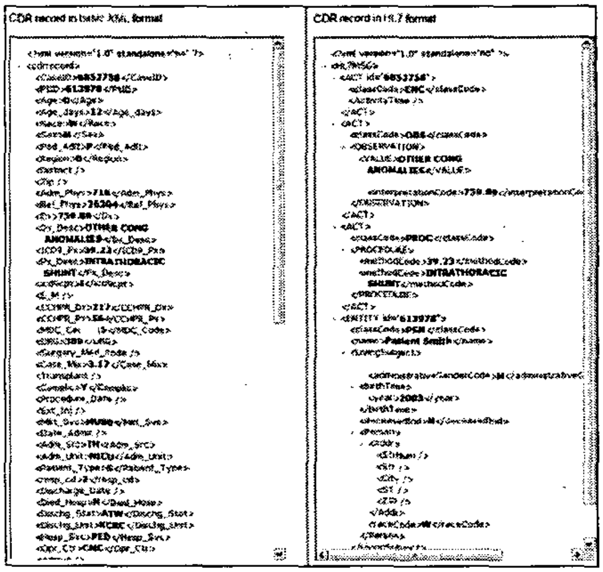
```
<?xml version="1.0"
standalone="no" ?>
- <cdrrecord>
    <CaseID>6853758</CaseID>
    <PtID>613978</PtID>
    <Age>0</Age>
    <Age_days>12</Age_days>
    <Race>W</Race>
    <Sex>M</Sex>
    <Ped_Adlt>P</Ped_Adlt>
    <Region>8</Region>
    <District />
    <Zip />
      <Adm_Phys>716</Adm_Phys>
      <Ref_Phys>26304</Ref_Phys>
    <Dx>759.89</Dx>
    <Dx_Desc>OTHER CONG
      ANOMALIES</Dx_Desc>
    <ICD9_Px>39.23</ICD9_Px>
    <Px_Desc>INTRATHORACIC
      SHUNT</Px_Desc>
    <Icdxcept>1</Icdxcept>
    <E_M />
      <CCHPR_Dx>217</CCHPR_Dx>
    <CCHPR_Px>56</CCHPR_Px>
      <MDC_Code>15</MDC_Code>
    <DRG>389</DRG>
    <Surgery_Med_code />
    <Case_Mix>3.17</Case_Mix>
    <Transplant />
    <Complic>Y</Complic>
    <Procedure_Date />
    <Ext_Inj />
```

Figure 6a: XML HL7 Mapping Tool Part I

The prototype shown in Figures 6a and 6b allows a user to select a given record in the CDR *Visits* table shown at far left, and then transforms that record, first into generic XML, then into an XML document which uses RIM objects.

4 Discussion

Based on our progress to date, it is evident that the vast majority of our local data can be properly mapped to the RIM. This is not surprising, since the goal of the HL7 is to provide an information model that acts as a reference for HL7 development of other message communication tools. The content of these messages includes standard healthcare concepts and values that, by and large, parallel the information stored in our clinical data warehouse. Because the primary customers of the CDR are clinical researchers, administrators interested in quality assessment, and students, there are many areas of the RIM that are not relevant for this mapping effort. Information relevant to patient financial account management, facilities planning, and dietary services are beyond the scope of the CDR.



CDR record in basic XML format

```
<?xml version="1.0" standalone="no" ?>
- <cdrrecord>
    <CaseID>6853758</CaseID>
    <PtID>613978</PtID>
    <Age>0</Age>
    <Age_days>12</Age_days>
    <Race>W</Race>
    <Sex>M</Sex>
    <Ped_Adlt>P</Ped_Adlt>
    <Region>8</Region>
    <District />
    <Zip />
    <Adm_Phys>716</Adm_Phys>
    <Ref_Phys>26304</Ref_Phys>
    <Dx>759.89</Dx>
    <Dx_Desc>OTHER CONG
      ANOMALIES</Dx_Desc>
    <ICD9_Px>39.23</ICD9_Px>
    <Px_Desc>INTRATHORACIC
      SHUNT</Px_Desc>
    <Icdxcept>1</Icdxcept>
    <E_M />
    <CCHPR_Dx>217</CCHPR_Dx>
    <CCHPR_Px>56</CCHPR_Px>
    <MDC_Code>15</MDC_Code>
    <DRG>389</DRG>
    <Surgery_Med_code />
    <Case_Mix>3.17</Case_Mix>
    <Transplant />
    <Complic>Y</Complic>
    <Procedure_Date />
    <Ext_Inj />
    <Att_Svc>NURB</Att_Svc>
    <Date_Admit />
    <Adm_Src>TH</Adm_Src>
    <Adm_Unit>NICU</Adm_Unit>
    <Patient_Type>I</Patient_Type>
    <Hosp_cd>2</Hosp_cd>
    <Discharge_Date />
    <Died_Hosp>N</Died_Hosp>
    <Dischg_Stat>ATW</Dischg_Stat>
    <Dischg_Unit>NCBC</Dischg_Unit>
    <Hosp_Svc>PED</Hosp_Svc>
    <Opr_Ct>CMC</Opr_Ct>
```

CDR record in HL7 format

```
<?xml version="1.0" standalone="no" ?>
- <HL7MSG>
  - <ACT id="6853758">
      <classCode>ENC</classCode>
      <ActivityTime />
    </ACT>
  - <ACT>
      <classCode>OBS</classCode>
    - <OBSERVATION>
        <VALUE>OTHER CONG
          ANOMALIES</VALUE>
          <interpretationCode>759.89</interpretationCo
      </OBSERVATION>
    </ACT>
  - <ACT>
      <classCode>PROC</classCode>
    - <PROCEDURE>
        <methodCode>39.23</methodCode>
        <methodCode>INTRATHORACIC
          SHUNT</methodCode>
      </PROCEDURE>
    </ACT>
  - <ENTITY id="613978">
      <classCode>PSN</classCode>
      <name>Patient Smith</name>
    - <LivingSubject>
          <administrativeGenderCode>M</administrativeG
      - <birthTime>
          <year>2003</year>
        </birthTime>
        <deceasedInd>N</deceasedInd>
      - <Person>
        - <Addr>
            <Street />
            <Str />
            <City />
            <ST />
            <ZIP />
          </Addr>
          <raceCode>W</raceCode>
        </Person>
```

Figure 6b: XML HL7 Mapping Tool Part II

Similarly, there are specific concepts in the CDR that do not require representation in the RIM. Derived attributes stored in the CDR to optimize performance do not require mapping to external systems; in fact, the associated redundancy in doing so might have more disadvantages than benefits.

While the information content in the RIM was generally very satisfactory for our mapping effort, the largest barrier appears to be the steep learning curve associated with the model. This is arguably unavoidable for an information model that purports to comprehensively represent all information content pertinent to the healthcare domain. In addition, the RIM is still under active development and primarily used internally for other standards development rather than being intended for "public consumption".

One essential element in data integration that was purposefully set aside in our mapping effort deals is the fundamental challenge of healthcare terminology[8]. With some notable exceptions (ICD9-CM, CPT), healthcare organizations have not adopted emerging terminology standards. Being able to export information in a common model does little good if two organizations use completely different codes for laboratory results, for example. Such terminologies are slowly being adopted at both the local and national levels, and offer great promise for sharing information in meaningful ways and unlocking the tremendous potential of large volumes of clinical and biomedical data.

5 Conclusions and Next Steps

Several concomitant forces are driving the need for the further adoption of information storage and communication standards in the healthcare field. The sheer volume of information related to clinical care, medical knowledge (in the form of academic literature), clinical and biomedical research, and quality assessment has led to the development of a nearly infinite number of disparate information systems ranging from small desktop databases to large, shared national information resources. Integrating this data to support a wide variety of secondary uses requires the investigation and adoption of appropriate standards. We have begun to explore issues related to mapping from a local data warehouse to a developing standard for modeling health care information, and are optimistic about our findings to date. As the RIM undergoes further development, testing, and refinement, its ability to represent critical information components in the healthcare domain will no doubt increase.

We are currently conducting a data integration project between our institution and another large academic health center in the state. This effort will include not only clinical data resulting from patient care activities, but biomedical research data such as gene expression array results and proteomic studies. As a result of this project, we will be able to explore the generalizability of our findings to another institution, as well as test the ability of the RIM to handle these additional types of data associated with the basic science domain.

Acknowledgements

Information about the HL7 RIM and its development, along with Figures 3 – 5, are from the HL7V3 Ballot 4. The images are used with the permission of the HL7 organization.

References

[1] Health Level 7, http://www.hl7.org, July 2003.

[2] McDonald CJ. Huff SM. Suico JG. Hill G. Leavelle D. Aller R. Forrey A. Mercer K. DeMoor G. Hook J. Williams W. Case J. Maloney P. LOINC, a universal standard for identifying laboratory observations: a 5-year update. Clinical Chemistry 2003, 49(4):624-33.

[3] Unified Medical Language System, National Library of Medicine, http://www.nlm.nih.gov/research/umls/, July 2003.

[4] Stearns MQ. Price C. Spackman KA. Wang AY. SNOMED clinical terms: overview of the development process and project status. Proc AMIA Annual Symp 2001, 662-6.

[5] Scully KW. Pates RD. Desper GS. Connors AF. Harrell FE Jr. Pieper KS. Hannan RL. Reynolds RE. Development of an enterprise-wide clinical data repository: merging multiple legacy databases. Proc AMIA Annual Symp 1997, 32-6.

[6] Dolin RH. Alschuler L. Beebe C. Biron PV. Boyer SL. Essin D. Kimber E. Lincoln T. Mattison JE. The HL7 Clinical Document Architecture. JAMIA 2001, 8(6):552-69.

[7] Hripcsak G. Ludemann P. Pryor TA. Wigertz OB. Clayton PD. Rationale for the Arden Syntax. Computers & Biomedical Research 1994, 27(4):291-324.

[8] Cimino JJ. Review paper: coding systems in health care. Methods of Information in Medicine 1996, 35(4-5):273-84.