RJ10433 (A0808-003) August 4, 2008
Computer Science

IBM Research Report

Creating CDA R2 Laboratory Reports to Meet Public Health Surveillance Requirements

S. R. Renly, S. E. Knoop, J. H. Kaufman
IBM Research Division
Almaden Research Center
650 Harry Road
San Jose, CA 95120-6099

R. Ram
IBM Research Division
Haifa Research Laboratory
Mt. Carmel 31905
Haifa, Israel

Research Division
Almaden - Austin - Beijing - Cambridge - Haifa - India - T. J. Watson - Tokyo - Zurich

Creating CDA R2 Laboratory Reports to Meet Public Health Surveillance Requirements

S. R. Renly[1], S. E. Knoop[1], J. H. Kaufman[1], R. Ram[2]
IBM Almaden Research Center[1], 650 Harry Rd, San Jose, CA 95120 USA
IBM Haifa Research Lab[2], University of Haifa, Mount Carmel, Haifa, 31905 Israel

Abstract-**In the absence of world wide public health standards-based networks, the ability to internationally monitor and quickly respond to cross-border outbreaks is often relegated to media correspondents and website technologies[1-3]. HL7 Clinical Document Architecture Release 2 (CDA R2) provides unique benefits to resource poor agencies compared with laboratory messaging. With guidance from the public health community, today's technology can transform international public health surveillance methods into reliable electronic networks. IBM Research successfully incorporated public health laboratory reporting requirements into the Integrating the Healthcare Enterprise (IHE) Cross Enterprise Document Sharing Laboratory Report (XD-Lab) profile for a research surveillance prototype being developed with input from participants in the Middle East Consortium for Infectious Disease Surveillance (MECIDS).**

I. Introduction

Public health agencies are significant consumers of clinical laboratory data. The more standardized electronic clinical data sharing becomes, the more easily and quickly public health surveillance systems can access *and process* data to provide benefit through early detection and warnings.

The HL7 Clinical Document Architecture Release 2 (CDA R2)[4] provides benefits to resource poor agencies over HL7 electronic laboratory messaging. Today, laboratory information management systems (LIMS) have highly complex and expensive software for the determination of report copies, report routing, report preferences, report rendering, and report archival. Once a lab has installed a LIMS system, one is often confined to the abilities of that system or by the LIMS vendor fees for additional features (e.g. a custom HL7 message interface). The generation of HL7 CDA R2 laboratory reports allows for more flexible management of these important laboratory requirements. An archival report can be generated once, either from within a LIMS or by a data import into a CDA generation system. Add-on technologies can then be responsible for the importing, transforming, routing, rendering, and auditing functions. Such technologies can more easily tailor their functions to the simpler requirements and limited network availability of locations in resource poor areas. Other benefits our team has observed include the closer approximation to current paper-based reporting methods and human readability provided by XML.

With these benefits in mind, IBM Research[5] started a First of a Kind (FOAK) project to create a research prototype for any Public Health Information Affinity Domain (PHIAD)[6]. The standards-based public health laboratory surveillance system uses HL7 CDA R2 as its principle data exchange format and was developed with input from participants in the Middle East Consortium for Infectious Disease Surveillance (MECIDS)[7].

This paper highlights our success at partnering with Integrating the Healthcare Enterprise (IHE)[8] to build public health requirements into document-based laboratory reports. We outline the changes we found necessary for the MECIDS implementation and conclude with outstanding issues that will be necessary to meet additional laboratory needs.

II. Integrating The Healthcare Enterprise

A strong foundation for our project began with the work published through the IHE Laboratory Domain. The Laboratory Technical Framework Volume 3: Content[9] document provides detailed specifications for creating a CDA R2 compliant clinical laboratory report. This IHE profile, called Cross Enterprise Document Sharing Laboratory Report (XD-Lab) is currently in trial implementation with finalization expected in the fall of 2008.

Our first task was to outline public health surveillance use cases and requirements. We documented the hierarchical paradigm of public health reporting, an information flow process from local laboratories to regional surveillance networks and mandatory national government reporting. We also documented common data elements that are frequently reported.

We then analyzed the IHE XD-Lab profile for suitability. This work was initially done within the IHE Patient Care Coordination (PCC) Domain because of meeting proximity (in the USA), timing with the IHE IT Infrastructure (ITI) meetings where we were already active, and mentoring availability. IHE PCC had already established a number of CDA-based content profiles. IHE content profiles are specifications that define the data format and clinical information exchanged in a particular use case (eg. patient referral). A CDA-based IHE content profile is one that leverages the CDA R2 standard to satisfy a particular use case, often imposing constraints. By addressing key use cases in the exchange of healthcare content, PCC has developed a wealth of CDA-based content profiles in which concepts common to more than one use case are uniquely represented and reused. These 'pieces' of either level 1, level 2 or level 3 CDA information are called content modules and are identified by a unique template id. HL7 CDA templates are analogous to these content modules.

It was this CDA R2 content module approach in PCC that we elected to adopt to address our public health data represen-

tation needs and, by extension and in coordination with the IHE Lab Committee, applied to the whole of the XD-Lab content profile. By working with both Lab and PCC domains, we created new communication channels that have improved documentation consistency and created an ongoing harmonization effort to improve re-usability of laboratory components in PCC.

Two key changes resulted from our public health needs analysis, neither requiring schema extensions in CDA. The first is our ability to represent laboratory reports for human only, subject only, and human/subject use cases. The second is our ability to document public health notifications which can include a notifiable condition, a case number, and an outbreak number.

III. Building Public Health Laboratory Reports

PHIAD is a hierarchical standards-based public health laboratory reporting surveillance system focusing initially on notifiable microbiology and virology isolate reporting (ex Salmonella, Influenza). The XD-Lab profile provided excellent guidance for the challenges in isolate reporting. We were able to quickly complete clinical reports that included single isolate culture, susceptibility, and serotyping results.

We identified three unique document varieties in public health reporting: human patient reports, non-human subject reports, and human patient with non-human subject reports. The human patient report is a traditional CDA report for a patient who provides a specimen for testing. The non-human subject report includes results from specimens collected from non human material (e.g. water, food, animal). The human patient with non-human subject report includes results from a non-human subject specimen where the result affects the patient. For example, food a person ate may be tested for Salmonella or an animal that bit a person may be tested for Rabies. In the cases of reports with non-human subjects and patients with non-human subjects, we created IHE content modules in the XD-Lab profile shown in Table I. We now describe these in more detail.

Our CDA R2 representation of these document varieties represents the human patient as specified in the standard. If there is no human patient, we set the nullFlavor attribute on the recordTarget/patient to 'OTH' (Other). This, along with a templateId, provides an indication at the time of processing or rendering the content that the information pertaining to the non-human subject is documented elsewhere. In the document body under a Subject element, a code and qualifier is used to describe the non-human subject along with an address to indicate subject location. In addition, the subject id is able to be captured under recordTarget/patientRole/id. In this way, the recordTarget/patient in the CDA header and the subject elements work in conjunction to best document the identifying information for the non-human subject, see Fig 1. When both a human patient and non-human subject are included in the report, we have a problem in representing the id of the subject.

The Subject type in CDA R2 currently does not have an id sub-element. We work around this in our previous case where only a non-human subject is documented in the report by using the recordTarget/patientRole/id for the non-human subject id. When a human patient is present in the documentation, this id is necessarily reclaimed for identification of the patient. Again as an indicator, the templateId element is used for identifying the recordTarget as a human patient directly impacted by a non-human subject, see Fig 2.

To report public health notifications, we specified that the principle Act in the Entry for laboratory result data shall have an entryRelationship sub-element with an organizer to encapsulate Notifications. Report Notifications, represented as content modules, include the Notifiable Condition (e.g. Salmonella, Escherichia coli O157:H7), Case Report identification number(s), and Outbreak identification number(s) and are listed in Table II. Additionally, we allow for the qualification of the notification code to distinguish human patient, animal, or food reporting. Each component – the Notification Organizer, the Notifiable Condition, the Case Report identifier, and Outbreak identifier each have an assigned templateId and accompanying IHE content module specification added to XD-Lab. See Fig 3 for an example.

TABLE I
IHE Content Modules for Subject Based Reports

Description	IHE Content Module
Non-Human Subject	1.3.6.1.4.1.19376.1.3.3.1.2.1
Human Patient with Non-Human Subject	1.3.6.1.4.1.19376.1.3.3.1.3.1

```
<!-- CDA Header -->
<recordTarget typeCode="RCT">
  <templateId root="1.3.6.1.4.1.19376.1.3.3.1.2"/>
  <patientRole classCode="PAT">
    <id extension="66373839" root="1.3.6.1.4.1.19376.1.3.4"/>
    <patient nullFlavor="OTH">
  </patientRole>
</recordTarget>

<!-- CDA Body -->
<subject>
  <templateId root="1.3.6.1.4.1.19376.1.3.3.1.2.1"/>
  <relatedSubject>
    <code code="18998007" codeSystem="2.16.840.1.113883.6.96"
      codeSystemName="SNOMED-CT" displayName="Ferret species">
      <qualifier>
        <name code="105590001" codeSystem="2.16.840.1.113883.6.96"
          codeSystemName="SNOMED-CT" displayName="Substance"/>
        <value code="39866004" codeSystem="2.16.840.1.113883.6.96"
          codeSystemName="SNOMED-CT" displayName="Animal"/>
      </qualifier>
    </code>
    <addr><country>USA</country></addr>
  </relatedSubject>
</subject>
```

Figure 1. Representation of a Non-human Subject

```
<!-- CDA Header -->
<recordTarget typeCode="RCT">
 <patientRole classCode="PAT">
  <id extension="sw54321" root="1.3.6.1.4.1.19376.1.3.4"/>
  <addr>
   <city>Janesville</city><state>WI</state>
   <postalCode>53545</postalCode><country>USA</country>
  </addr>
  <telecom nullFlavor="UNK"/>
  <patient classCode="PSN">
   <name>
    <family>Winters</family><given>Shelly</given>
   </name>
  <administrativeGenderCode code="F"/>
  <birthTime value="19401213000000.0000-0500"/>
  </patient>
 </patientRole>
</recordTarget>

<!-- CDA Body -->
<subject>
 <templateId root="1.3.6.1.4.1.19376.1.3.3.1.2.1"/>
  <relatedSubject>
   <code code="18998007" codeSystem="2.16.840.1.113883.6.96"
    codeSystemName="SNOMED-CT" displayName="Ferret species">
    <qualifier>
     <name code="105590001" codeSystem="2.16.840.1.113883.6.96"
      codeSystemName="SNOMED-CT" displayName="Substance"/>
     <value code="39866004" codeSystem="2.16.840.1.113883.6.96"
      codeSystemName="SNOMED-CT" displayName="Animal"/>
    </qualifier>
   </code>
  <addr><country>USA</country></addr>
 </relatedSubject>
</subject>
```

Figure 2. Representation of a Patient and Non-human Subject

When sharing public health laboratory reports, we made extensive use the existent CDA nullFlavor attribute and 'MSK' (Masked) value. This type of nullFlavor indicates that there is a known proper value, but it will not be released in this context. This feature of CDA is especially useful in public heath for appropriately de-identifying certain information for cross jurisdiction sharing. In our work with MECIDS, we applied the 'MSK'-valued nullFlavor attribute to reflect data sharing policies which principally focused on de-identification. In our work with MECIDS, configurable sharing policies were created and applied throughout the international sharing hierarchy and ranged from full patient identification, full authorship, and all the tests results to limited patient identification, partial authorship information, and only the notifiable isolate test result.

With these additions, the IHE XD-Lab profile has been recommended by the US Health Information Technology Standards Panel (HITSP)[10] within the biosurveillance use case.

TABLE II
IHE CONTENT MODULES FOR PUBLIC HEALTH NOTIFICATIONS

Description	IHE Content Module
Notification Organizer	1.3.6.1.4.1.19376.1.3.1.1
Notifiable Condition	1.3.6.1.4.1.19376.1.3.1.1.1
Case Identification	1.3.6.1.4.1.19376.1.3.1.1.2
Outbreak Identification	1.3.6.1.4.1.19376.1.3.1.1.3

IV. OTHER PUBLIC HEALTH WORK

Important parallel efforts are underway in both HL7 and IHE to address public health use cases when building new electronic data exchanges between clinical and public health organizations. Many of these efforts are evaluating HL7v3 and CDA.

The US Centers for Disease Control and Prevention (CDC) has worked closely with HL7 over the last two years to complete a CDA Draft Standard for Trail Use (DSTU) for the electronic submission of Healthcare Associated Infection (HAI)

```
<entryRelationship typeCode="COMP">
 <organizer classCode="CLUSTER" moodCode="EVN">
  <templateId root="1.3.6.1.4.1.19376.1.3.1.1"/>
  <statusCode code="completed"/>
  <component>
   <observation classCode="COND" moodCode="EVN">
    <templateId root="1.3.6.1.4.1.19376.1.3.1.1.1"/>
    <id extension="SALM" root="1.3.6.1.4.1.19376.1.3.4"/>
    <code code="170516003" codeSystem="2.16.840.1.113883.6.96"
     codeSystemName="SNOMED-CT"
     displayName="Notification of Disease">
     <qualifier>
      <name code="246087005"
       codeSystem="2.16.840.1.113883.6.96"
       codeSystemName="SNOMED-CT"
       displayName="Source of Specimen"/>
      <value code="116154003"
       codeSystem="2.16.840.1.113883.6.96"
       codeSystemName="SNOMED-CT" displayName="Patient"/>
     </qualifier>
    </code>
    <statusCode code="completed"/>
    <effectiveTime value="20080408000000.0000-0400"/>
    <value xsi:type="CE" code="27268008"
     codeSystem="2.16.840.1.113883.6.96"
     codeSystemName="SNOMED-CT"
     displayName="Salmonella"/>
   </observation>
  </component>
 </organizer>
</entryRelationship>
```

Figure 3. Representation of a Public Health Notifiable Condition

Reports[11]. Currently the DSTU includes four CDA R2 Report types. The reports are Blood Stream Infection Numerator Report, Surgical Site Infection Numerator Report, Procedure Denominator Report, and the Denominator for Intensive Care Unit Report. The Blood Stream Infection Numerator Report and Surgical Site Infection Numerator Report contain specific instances of infection and observation details. Denominator Reports provide summarized data which provide patient days denominator data for a given location and period of time. When numerator and denominator reports are evaluated together, risks are better understood and metrics to reduce infections can be derived. Future work is slated to include numerator reporting on Urinary Tract Infections, Pneumonia, and Multi-Drug Resistant Organisms. Denominator reporting will include Specialty Care Areas and Neonatal Intensive Care Units.

It is common for public health epidemiologists to request forms that provide similar numerator and denominator requests tallying laboratory orders received based on clinical presentation including flu-like or diarrhea symptoms. Thus, this group's expanded use of CDA to creating reports containing summarized data for many patients over a period of time offers new relevance for CDA in public health laboratory reporting. We are eager to adapt their work to the laboratory use case.

Also in the US, the Healthcare Information Technology Standards Panel (HITSP) has completed a public comment phase on Requirements, Design, and Standards for a Public Health Case Report[12]. They are initially focused on Tuberculosis, Hepatitis B, Tularemia, and Anthrax case reports. CDA R2 is listed as one of the implementation options under consideration. We will encourage the selection of CDA R2 and would like to add public health case report documents into our PHIAD prototype.

The HL7 Public Health and Emergency Response (PHER) Working Group[13] is active in several areas. Their stated areas of interest are very broad and include population health monitoring, disease and event detection, contact tracing, outbreak investigation, immunization, vital records, public health registries, food safety, emergency preparedness, and reporting to monitoring/regulatory agencies. Recent activity includes work in Vital Records, Immunization, and Tuberculosis data standards. We presented our work to this group in January 2008 and see greater future collaboration as beneficial.

Joint with IHE, the US Public Health Data Standards Consortium (PHDSC)[14] has completed a significant white paper "Building a Roadmap for Health Information Systems Interoperability for Public Health" [15]. This paper brought together a diverse public health community to define who they are, their missions, and their challenges. The goal to create a standalone IHE Public Health Domain to begin addressing some of these challenges was partially realized by the creation of an IHE Quality, Research, and Public Health Domain in 2007. This new domain's goal is to identify and begin addressing shared challenges, including how to realize aggregate reporting needs not previously expressed in patient centric clinical workflows. We are grateful for the expertise provided within this new domain and their feedback supporting public health reports within the XD-Lab specification.

V. FUTURE WORK

We are encouraged that HL7 CDA R2 met the majority of implementation requirements for our prototype Public Health Information Affinity Domain system. It is equally noteworthy that our implementation of XD-Lab, initially focused only on the clinical laboratory setting, needed minimal refinements to encompass our public health use cases. Even with these advancements, there are additional public health and general laboratory requirements that remain outstanding. We have listed these here for future work, in no particular order.

First, epidemiologists describe a desire to receive summary laboratory reports that contain counts of tests ordered, performed, and the frequency of positives. Adapting what has been done in HAI, we believe we can achieve additional report specifications for these types of laboratory summaries.

Second, many human diseases of public health concern are transmitted from animal to human. Over 200 zoonoses[16] are known; including Salmonellosis, Taeniasis, West Nile, and Avian Influenza. Collection and testing of veterinary samples requires additional coding that is not currently available in CDA R2. IHE has identified the need to support much more information at the subject level and subject owner level than is currently available.

Third, as described earlier, we identified a human patient with non-human subject report as a unique instance of a public health laboratory report. This report describes a non-human subject that has an affect on the patient. When a non-human subject affects several patients at the same time, such a scenario could be accompanied with a number of lab reports, one for each patient. These reports have identical sections that describe the subject. A potential useful solution to avoid this duplication, while protecting patient privacy, would be to add the ability to associate the human patient reports with a single non-human subject report. The ability to leverage CDA R2 in this scenario is an open question.

Fourth, we have requests to document additional specimen storage information for specimens that will be retained for long periods and may have future uses. Specimen storage information can include: current quantity, quality, a storage type (e.g. paraffin block, frozen vial), expiration, and a location. Currently, this level of granular mark-up is unavailable in CDA R2.

Fifth, the XD-Lab specification has intentionally left out of scope details for how to report appended, edited, and corrected reports; frequent document operations in the laboratory setting. In some situations it may be sufficient to document in a comment what has changed between document versions, but policies and regulations can require greater documentation such as entire previous result documentation accompanying a new, updated result. We, in conjunction with the IHE Lab Committee, have identified the following use cases for future evaluation as we have not identified solutions in the current CDA R2:

- Use Case 1 – A laboratory report was issued on the wrong patient. Today two documents will be issued. For the correct patient, a new laboratory report shall be approved. For the wrong patient, the specification to represent the error in the CDA document and negate content has been left for future work.

- Use Case 2 – A laboratory report was issued with incomplete or incorrect non-result data, defined as the information found in the CDA Header or Specimen, such as the collection date and time. The results and result interpretation are unchanged by the addition, edit, or correction of non-result data. The specification to represent the change has been left for future work.

- Use Case 3 – A laboratory report was issued with incomplete or incorrect non-result data. In this case, the results and interpretation, defined as the information found in an observation, organizer, or media, are changed by the addition, edit, or correction of non-result data. The specification to represent the change and the interpretation impact has been left for future work.

- Use Case 4 – A laboratory report was issued with incomplete or incorrect result data. The specification to represent the change and the interpretation impact has been left for future work.

- Use Case 5 – A laboratory report was issued with incomplete or incorrect non-result and result data. The specification to represent the change and the interpretation impact has been left for future work.

Finally, IHE identification of open internationally accepted terminologies, for example, Logical Observation Identifiers Names and Codes (LOINC)[17], is important to support consistent representation of laboratory concepts in CDA exchanges. XD-Lab uses LOINC codes to document the Laboratory Specialty, Specimen Collection Date/Time, and Annotation Comment. Additional open licensed codes are desired for concepts such as Specimen Receive Date/Time, Notifiable Condition Code, Notifiable Condition Value, Subject Code, Subject Type Code, and Organism Code. Vocabularies for these concepts are available under Systematized Nomenclature of Medicine – Clinical Terms (SNOMED-CT)[18] for use only within countries that are licensed. Finding appropriate, global terminology for these key concepts is key to consistent, international interpretation of the data carried in one of our XD-Lab public health reports.

VI. CONCLUSION

In conclusion, we are pleased with the current capabilities in CDA R2 for meeting public health surveillance requirements in laboratory reporting. The IHE XD-Lab profile provided a strong foundation for laboratory document constraints and by working with IHE we successfully created one content profile that addresses both clinical and public health workflows, ensuring bi-directional exchange capability between these entities. Some people consider public health needs a small niche in standards implementation. We strive to demonstrate that collaborative discussions can lead to a single standards implementation that meets clinical data exchange needs as well as clinical and public health exchanges. We look forward to continuing our work in HL7 and IHE to address items discussed within the future work section as well as continuing collaboration with public health organizations to bring their needs forward.

REFERENCES

[1] Mawudeku, A. and Blench, M., "Global Public Health Intelligence Network (GPHIN)", 7th Conference of the Association for Machine Translation in the Americas (2006)

[2] Samaan, G. et al., "Rumor surveillance and avian influenza H5N1", Emerging Infectious Diseases, 11, 3 pp 463-466 (2005)

[3] Madoff, L.C., "ProMED-mail: An Early Warning System for Emerging Diseases", Clinical Infectious Diseases, 39,2 pp 227-232, (2004)

[4] R.H. Dolin, L. Alschuler, S. Boyer, C. Beebe, F. M. Behlen, P. V. Biron, A. Shabo, (Editors). HL7 Clinical Document Architecture, Release 2.0. ANSI-approved HL7 Standard; May 2005. Ann Arbor, Mich.: Health Level Seven, Inc.

[5] IBM Research – http://www.ibm.com

[6] B. Carmeli, et al. "Public Health Affinity Domain: A Standards-Based Surveillance System Solution.", In "Intelligence and Security Informatics: Biosurveillance", Proceedings of the second NSF Workshop on Biosurveillance 2007, Lecture Notes in Computer Science, vol 4506, 147-158, 2007, Springer Berlin / Heidelberg.

[7] Middle East Consortium for Infectious Disease Surveillance – http://www.mecids.net

[8] Integrating the Healthcare Enterprise – http://www.ihe.net

[9] Laboratory Technical Framework Volume 3 Content Revision 2.1. August 2008. Integrating the Healthcare Enterprise.

[10] Health Information Technology Standards Panel – http://www.ansi.org/hitsp

[11] HL7 Implementation Guide for CDA Release 2: NHSN Healthcare Associated Infection (HAI) Reports Draft Standard for Trial Use Release 1, February 2008. Ann Arbor, Mich.: Health Level Seven, Inc.

[12] HITSP Public Health Case Reporting Use Case Requirements, Design and Standards Selection Version 1.0, June 2008.

[13] Public Health and Emergency Response Working Group – http://www.hl7.org

[14] Public Health Data Standards Consortium – http://www.phdsc.org

[15] Building a Roadmap for Health Information Systems Interoperability for Public Health, 2008. Integrating the Healthcare Enterprise.

[16] World Health Organization – http://www.who.org

[17] Logical Observation Identifiers Names and Codes, Regenstrief Institute – http://www.regenstrief.org/medinformatics/loinc/

[18] Systematized Nomenclature of Medicine – Clinical Terms, IHTSDO – http://www.ihtsdo.org