MEDINFO 2010
C. Safran et al. (Eds.)
IOS Press, 2010

doi:10.3233/978-1-60750-588-4-289

HL7 CDA Implementation Guide for Structured Anatomic Pathology Reports Methodology and Tools

Haitham Kussaibi[a], François Macary[b], Mary Kennedy [c], David Booker [d], Victor Brodsky[e], Thomas Schrader[f], Marcial Garcia-Rojo[g], Christel Daniel[a,b,h]

[a]INSERM, UMR_S 872, Eq. 20, Paris, F-75006 France; Université René Descartes, Paris, F-75006 France ; [b]ASIP Santé, Paris, France [c]College of American Pathologists, Northfield, Illinois, USA;[d]College of American Pathologists, Medical College of Georgia, Augusta, GA, USA; [e]College of American Pathologists and Weill Cornell Medical College, New York, NY, USA; [f]Department of Pathology, Charite', Berlin, Germany; [g]Hospital General de Ciudad Real, Department of Pathology, Ciudad Real, 13005 Spain ;[h]AP-HP, Hôpital George Pompidou, Département d'Informatique Hospitalière, Paris, F-75015 France

Abstract

Anatomic pathology reports (APR) provide diagnostic and prognostic information crucial to patient care, clinical research and epidemiology. Currently, it is difficult to collect and exchange APR data between different healthcare organizations at an international level. Objective: IHE and HL7 anatomic pathology joint efforts aim at providing a methodology and tools to define an international HL7 "Clinical Document Architecture" (CDA) implementation guide for APRs and especially in the domain of cancer. Methods: A four-step methodology is employed, consisting of comparing existing clinical model of APRs originating from different countries; deriving consensus-based clinical models (Delphi technique); providing the corresponding HL7 CDA implementation guide ("CDA templates") and validating these templates. Results: International experts defined HL7 CDA implementation guides for breast and colon cancer APRs within an IHE content profile. CDA templates include required data elements, as well as optional ones, that can be further specified as required in national extensions. Conclusion: This study demonstrates that it is possible to define an international HL7 CDA implementation guide for cancer APRs. Further efforts are needed to provide CDA templates for approximately 60 other cancer APRs dedicated to different organs, diagnoses, and procedures as well as for APRs of non neoplastic pathologies. The methodology is not confined to APRs and could be applied to clinical documents of any type.

Keywords:

Anatomic pathology reports, Synoptic reports, Cancer checklists, Structured reports, Structured data entry, Delphi method, IHE, HL7, CDA.

Introduction

Anatomic pathology reports (APR) document the pathologic findings in specimens removed from patients for diagnostic or therapeutic reasons. This information can be used for patient care, clinical research and epidemiology. Currently, APRs for cancer patients have generated the greatest need for the collection and exchange of data contained in APRs, and many organizations have created templates in an effort to standardize APR reporting. The heterogeneity of such templates and lack of standards for structuring the relevant data elements in reports, hamper the exchange of this information among different information systems and healthcare organizations. Standardizing and computerizing APRs is necessary to improve the quality of reporting and the exchange of APR information[1].

As part of joint IHE and HL7 anatomic pathology activities, our objective is to provide a methodology and tools that facilitate the development of international clinical models for APR, including cancer APRs, as well as the production of the corresponding HL7 CDA implementation guides (CDA templates).

Several studies provide recommendations that delineate the required, preferred, and optional elements which should be included in any APR, regardless of report types (e.g reporting guidelines in[2].

Several international initiatives intend to define standard clinical models for specific types of APRs. For example, in the cancer domain, in the United States, the CAP (College of American Pathologists) has published 67 cancer checklists and background information[3]. In France, the SFP (French society of pathology) has published 23 minimum data sets for 21 cancer locations[4]. Together, the recommendations for generic and specific APR reporting have become clinical guidelines, the use of which may be required by accrediting bodies. The majority of encoded elements of these clinical models are associated with encoded value sets. The most frequently used coding systems in anatomic pathology domain are SNOMED Clinical Terms®, ICD-O-3 and ADICAP in France [5].

Since these standardization efforts are conducted at a national level there are some discrepancies between clinical models across countries and even some heterogeneity between clinical models within the same national initiative. There is a need to propose a methodology and tools to achieve better consistency of clinical models at an international level. There are several methods to achieve consensus-based agreement among experts. One such method is based on Delphi technique: a sys-

tematic, interactive forecasting method that relies on a panel of independent experts[6].

In addition to standardizing the cancer APR contents, it is necessary to computerize them. Several studies have focused on defining an appropriate IT standard comprising the structured and encoded clinical documents (e.g. CAP eCC). HL7 CDA is one of the most reliable standards that can support these needs[7]. CDA allows the clinical data to be both human and machine-readable and provides a framework for incremental growth in the granularity of structured, codes-bound clinical information. However, there are currently very few national initiatives of CDA implementation guides for the APR, one example being developed at the National IT Institute for Healthcare in the Netherlands, another one by HL7 Germany[8].

Materials and Methods

We followed a 4-steps methodology to define an HL7 CDA implementation guide for APR:

Step 1: Defining clinical models for APR (structuring and standardizing APRs medical content)

Clinical models for APR should address all Anatomic Pathology reporting domains such as surgical pathology, cytology, autopsy and even research (e.g. molecular biology or tissue micro arrays (TMA)). In order to ensure consistency among clinical models, we first defined a set of constraints that apply across all APRs regardless of domain. We then further identified the set of constraints that apply across all **cancer** APRs.

Generic clinical model for APR

Based on analysis the available recommendations that outline elements which should be included in an APR regardless of report types [2], we identified the sections of the generic clinical model for APR.

Generic clinical model for cancer APR

Based on the recommendations[2], specific to cancer APR, we defined sub-sections specific to the generic clinical model for cancer APR [2]. Then, based on comparison of the existing organ/diagnosis/procedure specific cancer checklists as defined by the CAP and by the SFP, we identified the elements that were most frequently present in the various organ/procedure specific checklists.

Organ/diagnosis/procedure specific clinical models for cancer APR

Based on the generic clinical model for cancer APR, we created organ/diagnosis/procedure specific clinical models. We merged the CAP and SFP checklists keeping a single occurrence of each common data element, and flagged the other elements with the name of the source template.

Step 2: Validating clinical models for APRs

Consensus sessions were organized by IHE and HL7 Anatomic Pathology workgroups and European COST action IC0604 in order to validate the clinical models for cancer APRs. A panel of experts first agreed on the sections of the generic clinical model for cancer APR during two face-to-face meetings. In France, two online questionnaires were published in order to evaluate discrepancies between CAP and SFP cancer checklists for breast cancer and colon.

According to the Delphi method, after the first survey round, a facilitator provides an anonymous summary of the experts' responses with their comments, in order to decrease the range of answers in the second round. After achieving the consensus or stability of results the mean or median scores of the final rounds determine the results.

Step 3: HL7 CDA implementation guide for APRs

HL7 CDA Release 2.0 provides a general architecture for designing and implementing clinical documents in an electronic format that is both human and machine-readable. Because of the architectural nature of the CDA standard, individual implementations are always associated with an implementation guide (also called "HL7 CDA template"), i.e. a document that describes how the CDA standard should be implemented for a particular type of document used in a specific context. A CDA document begins with a header that states the context of care in which the document was produced, identifies the various participants involved (patient, care providers, devices, etc) and states the responsibilities regarding the content of the document. The body of the document can be organized as a hierarchy of sections. Each section lays out its text for the reader, and may in addition carry fine-grained coded machine-readable data, corresponding to that text. We mapped the various roles of professionals involved in an APR to header elements. We then defined body sections, and assigned each section a unique code, a title and a text block. Finally, we coded the fine-grained machine-readable data into entries attached to the sections. Codes have been assigned to sections and to the various entry elements (acts (observations, procedures, etc), entities (specimen)) carried within the entries.

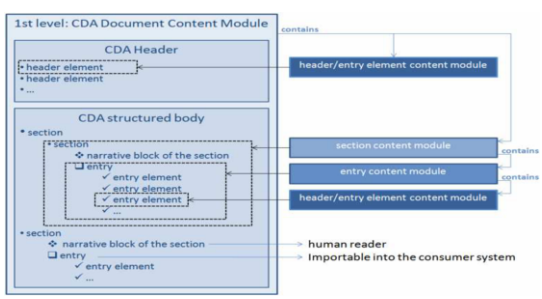

Figure 1- Hierarchical organization of CDA R2 content modules

Vocabulary domains provide value sets for coded CDA components. Some of these vocabulary domains are internally defined by HL7 V3. Others are drawn from external coding systems such as LOINC or SNOMED CT. Whether internal or external, every vocabulary domain has a unique HL7-assigned identifier (HL7 OID which is an ISO object identifier), and every concept within a vocabulary domain has a unique code and an associated display name in a given language. For some of the CDA components, the vocabulary domain is imposed by the standard. For others, the implementer is free to choose

from any relevant external source, such as LOINC, SNOMED CT or some other realm-specific vocabulary. For example, the possible values for the observation "histological type" in CAP cancer checklists are encoded using SNOMED CT values sets, while in France, SFP cancer checklists these values are encoded using ADICAP or ICD-O-3 values sets.

All the operations described above have been eased by the reuse whenever possible of relevant templates for CDA elements that had formerly been defined by the Patient Care Coordination (PCC) and Laboratory (LAB) domains of IHE.

Step 4: Evaluating HL7 CDA implementation guides

The good practice in building HL7 CDA implementation guides is driven by this key principle of the standard – "*one single xml schema CDA.xsd for all types of clinical documents*" – Therefore, like all other CDA implementation guides, this APR CDA implementation guide relies on the original CDA.xsd schema, unchanged. The set of templates that compose the implementation guide express constraints restraining the options allowed by the original CDA.xsd schema, and binding its coded elements to predefined value sets.

These constraints are expressed in formal language within the implementation guide, and will also be translated as assertions into a schematron file[9]

From that point, the process of evaluation of the implementation guide will consist in three steps:

a) Build a collection of APR instances conforming to the implementation guide, with a clinically relevant content provided by the domain experts (pathologists and clinicians).

b) Validate each APR instance against the standard CDA.xsd schema.

c) Validate each APR instance against the templates, applying the schematron file of assertions expressing the constraints of these templates.

Results

Clinical model for Anatomic Pathology Report (APR)

The generic clinical model for APR is structured into six sections; three required sections (Clinical Information, Macroscopic Observation and Final Diagnosis section), one conditional section (Intra-operative Observation) and two optional sections (Microscopic Observation, Tissue Dissection and Ancillary Tests). We defined additional constraints for the generic clinical model for cancer APR.

Table 1 summarizes the six sections and the corresponding sub-sections and entries of the generic clinical model for cancer APR (entries specific to cancer as designated by the * symbol).

In addition, two clinical models for organ-procedure specific cancer APRs (breast and colon) constraining the generic cancer model were defined. Following the consensus sessions conducted in France, SFP cancer checklists were aligned to the corresponding CAP cancer checklists in order to be defined as French national extensions of these CAP checklists.

Table 1- Sections, subsections and entries of the generic clinical model for cancer APR.

Generic cancer APR sections	Generic cancer APR subsections/entries
CLINICAL INFORMATION Clinical information provided by the ordering physician: reason for anatomic pathology procedure, active problems (preoperative and/or postoperative diagnosis, lab data), collection procedure(s) and specimen description(s) for all delivered specimen(s) reported separately.	**Reason for AP procedure** **History of present illness** **Active Problems** **Specimen clinical information entry** • Specimen Information Organizer o Collection procedure o Specimen(s) type o Specimen location
INTRAOPERATIVE EXAMINATION (conditional) Intraoperative diagnoses +/- images for all delivered specimen(s) reported separately.	**Intraoperative entry** • Specimen Information Organizer o Diagnostic observation o Link(s) to images
MACROSCOPIC OBSERVATION Collection procedure(s) and specimen description(s) (if not provided by the ordering physician), +/- gross findings +/- images for all delivered specimen(s) reported separately.	**Macroscopic observation entry** • Specimen Information Organizer(s) o Collection procedure o Specimen type o Diagnostic observation(s) o Link(s) to images
MICROSCOPIC EXAMINATION (Optional) Histopathologic findings (e.g results of histo-chemical and immunohistochemical stains) +/- images for some delivered specimen(s) reported separately.	**Microscopic observation entry** • Specimen Information Organizer(s) o Diagnostic observation (s) o Link(s) to images
DIAGNOSTIC FINDINGS Diagnoses +/- additional pathologic finding(s) +/- results of ancillary studi(es) (=cancer checklist(s), in case of cancer) +/- images for all specimens delivered, reported separately.	**Diagnostic findings entry** • Specimen Information Organizer(s) o Tumor location* o Tumor histologic type and grade* o Tumor extension (including pT, pN)* o Treatment effect o Additional findings o Results of ancillary techniques o Link(s) to images
TISSUE DISSECTION AND ANCILLARY TESTS (Optional) Tissue dissection (representative specimens and derived specimens dissected for other ancillary procedures (flow cytometry, cytogenetics, molecular studies, electron microscopy, etc) or biorepository) for all specimens delivered, reported separately.	**Tissue dissection and ancillary tests entry** • Specimen Information Organizer(s) o Dissection technique o Specimen type o Ancillary technique

HL7 CDA implementation guide for APR

HL7 CDA implementation guide for APRs consists in a set of CDA templates described within the IHE Anatomic Pathology content profile "Anatomic Pathology Structured Report" available on the IHE web site[10].

With regards to the header, we have defined the content modules representing the various participants involved in the documented act, and/or in the production or stewardship of the APR.

With regards to the body, we first defined CDA templates for the six sections (Clinical Information, Intra-operative Observation, Macroscopic Observation, Microscopic Observation, Final Diagnosis section, Tissue Dissection and Ancillary Tests). Each section (e.g "Clinical information") is provided with a unique code, a title, a free text zone, the description of sub-sections (e.g "Reason for anatomic pathology procedure") and entries (e.g "Specimen Clinical Information") as shown in Figure 2.

We then defined CDA templates for entries. Each entry (e.g "Specimen Clinical Information") contains a unique code and the description of the embedded entries or entry elements (e.g "specimen collection procedure", "effectiveTime" or "targetSiteCode"). One or more organizer may be used according to the number of specimens to which information is attached to. Each organizer allows identifying one specimen and describing its related acts (observation, procedure…).

```
<component>
 <section>
   <templateId root='1.3.6.1.4.1.19376.1.8.1.2.1'/>
   <code code='22636-5' displayName=Pathology report
relevant history' codeSystem='2.16.840.1.113883.6.'
codeSystemName='LOINC'/>
   <title>Clinical information</title>
   <text>
   Excision biopsy of the breast. Nodule of 1cm,
upper inner quadrant of right breast.
   </text>
    <entry>
     <!--Specimen Clinical Information-->
     <templateID root='1.3.6.1.4.1.19376.1.8.1.3.1'/>
      <!--Information related to 1st specimen-->
       <organizer classCode="CLUSTER">
       <templateId root='1.3.6.1.4.1.19376.1.8.1.4.4'/>
        <!-- Specimen collection procedure -->
        <component>
         <procedure classCode="PROC"moodCode="EVN">
           <templateId
root='2.16.840.1.113883.10.20.15.3.2'/>
            <code
code='277261002'displayName='Excision Biopsy'
codeSystem='2.16.840.1.113883.6.96'/>
          <effectiveTime><!--collection date&time-->
          <high value="201012150935"/>
          </effectiveTime>
          <targetSiteCode code='76752008'
displayName='Breast'codeSystem='2.16.840.1.113883.6.
96'/>
         </procedure>
       </component>
      </organizer>
       ...
    </entry>
    <component>
       <section>
          <templateId root='1.3.6.1.4.1.19376.1.8.1.2.1'/>
          <code code='34122-2' displayName='Reason
for referral' codeSystem='2.16.840.1.113883.6.1'
codeSystemName='LOINC'/>
          <title> Reason for anatomic pathology pro
cedure</title>
          <text>
Nodule of 1cm, upper inner quadrant of right breast.
         </text>
       </section>
    </component>
  </section>
</component>
```

Figure 2- CDA template of the section "Clinical Information"

With regards to the encoding process, in order to maintain different value sets derived from different coding systems, we have used the capacity of HL7 CDA to express any encoded element as two or more equivalent codes derived from different vocabularies (coding systems).

HL7 CDA implementation guide evaluation

A collection of APR instances of cancer APRs, including breast and colon cancer APRs, was built conforming to the implementation guide and validated against the standard CDA.xsd schema.

Discussion

Based on different initiatives for standardizing Anatomic Pathology structured Reports (APR) and as part of joint IHE and HL7 Anatomic Pathology activities, international experts defined an HL7 CDA implementation guide for structured Anatomic Pathology reports and in particular cancer structured reports. In order to ensure consistency, the organ specific HL7 CDA templates are based on a common generic template, including the constraints that apply across all APR regardless of the reporting activity, procedure, diagnosis or organ. Subsequent templates would aim to maximize the reuse of the data elements across templates whenever possible.

HL7 CDA templates support semantic interoperability of clinical data that are both human readable and machine-processable and that can be stored in databases to be further queried or mined. Furthermore, filtering algorithms may be applied on machine-processable data elements of HL7 CDA templates in order to exploit different clinical information according to different contexts of use (patient care or research).

Medical consensus is not easy to achieve at regional, national, and international levels on important features that should be reported, as well as the vocabulary or coding system to use. Although there were many similarities between, for example, the CAP and SFP cancer checklists for breast and colorectal cancer, it was impossible to achieve exact mapping between them. We found discrepancies in comparing value sets of encoded elements. Furthermore, common data elements may be encoded using different reference terminologies (e.g some data elements are encoded by CAP using SNOMED CT, while they are encoded using ADICAP or ICD-O-3 by SFP in France).

More generally speaking, due to frequent changes in the value sets, the use of ISO/IEC 11179 (Information Technology -

Metadata registries) has been proposed to allow report senders to reference externally accessible metadata dictionaries for each data element. Given the current development of various vocabulary server projects (such as the Distributed Annotation System (DAS) Server[11], LexGrid [12] and Common Terminology Services (CTS2); an additional attribute specifying the communication method with the referenced server would be necessary.

Given this issue of value set variability across countries and over time, it may be necessary to provide access to the entire current available value set for each data element. Ideally, the upcoming versions of the CDA standard would provide an attribute for directly referencing an externally accessible local (or a global) vocabulary server for each coded data element, thereby allowing the recipient of the report to query the sender's vocabulary server for a data element description and a value set belonging to the unique id of the given data element. Additionally, standardized methods for capturing form rules governing dynamic data element availability and values within the current template instance (for example using XForms) would need to be investigated and agreed upon by experts for each template, allowing for consistency across various laboratory information systems. The most appropriate method to ensure the intended layout and appearance of the transmitted report remains to be identified.

This study demonstrated that it is possible to define international IHE content profiles for both a generic cancer APR and organ/procedure specific cancer APR and to derive from these HL7 CDA international implementation guides, national extensions taking into consideration national or local constraints (e.g. local coding systems).

With regards to the evaluation step, our perspective is to make a broader use of schematrons. This rule-based validation language for making assertions about the presence or absence of patterns in XML trees is currently used to validate each APR instance only against the CDA schema. We plan to validate instances against specific APR templates, applying the schematron file of assertions expressing the constraints of these templates. Furthermore, the successful tests of the content profile at IHE Connectathons could attest the quality of the CDA templates. At last, only their adoption in real world implementation will attest their relevancy.

Further efforts are needed to provide implementation guides for the remaining organ-specific cancer APRs and also for APRs of non-neoplastic pathologies.

It is necessary to develop a tool that automates and supports the modeling process in order to cover all cancer domains in an acceptable period of time.

One of the crucial issues is not only to guarantee the consistency between APR templates at an international level but also to ensure their consistency with other clinical domains such as IHE Laboratory or Patient Care Coordination. The on-going international effort aimed at providing world-wide template repositories will allow the reuse of previously defined templates in order to avoid duplication of data elements across domains and to save time and effort[13].

The modeling methodology in its three phases of consensus achievement, modeling, and evaluation could be applied not only to the other organ/procedure specific APRs, but also to clinical documents of any type. The topic has now emerged as an important area of standards development, and a useful focus for international cooperation[14].

Acknowledgments

This study was supported by the following grants: FISCAM BR-CCM-2006/03, COST Action IC0604 Euro-Telepath, ADICAP.

References

[1] Nakhleh RE. Patient safety and error reduction in surgical pathology. Arch Pathol Lab Med. 2008 Feb;132(2):181-5.

[2] Goldsmith JD, Siegal GP, Suster S, Wheeler TM, Brown RW. Reporting guidelines for clinical laboratory reports in surgical pathology. Arch Pathol Lab Med. 2008 Oct;132(10):1608-16.

[3] CAP. Cancer Protocols and Checklists. http://www.cap.org.

[4] SFP. Standardized patholoy reports. http://sfpathol.org.

[5] Bailly L, Marine-Barjoan E, Ambrosetti D, Roussel JF, Caissotti C, Ettore F, et al. [Data quality of cancer registration by Adicap codes, used by French pathologists from Paca, 2005-2006]. Ann Pathol. 2009 Apr;29(2):74-9.

[6] Linstone H. The Delphi Method: Techniques and Applications. http://wwwisnjitedu/pubs/delphibook.

[7] Dolin RH, Alschuler L, Boyer S, Beebe C, Behlen FM, Biron PV, et al. HL7 Clinical Document Architecture, Release 2. JAMIA. 2006 Jan-Feb;13(1):30-9.

[8] Germany H. HL7 implementation guide for diagnoses. http://www.hl7.de/download/documents/diagnosen/Diagnoseleitfaden-v0.99b.pdf

[9] ISO/IEC 19757 - Document Schema Definition Languages (DSDL) - Part 3: Rule-based validation – Schematron: http://standards.iso.org/ittf/PubliclyAvailableStandards/index.html.

[10] IHE. IHE Anatomic Pathology content profile. http://www.ihe.net/.

[11] Dowell RD, Jokerst RM, Day A, Eddy SR, Stein L. The distributed annotation system. BMC Bioinformatics. 2001;2:7.

[12] Pathak J, Solbrig HR, Buntrock JD, Johnson TM, Chute CG. LexGrid: a framework for representing, storing, and querying biomedical terminologies from simple to sublime. JAMIA. 2009 May-Jun;16(3):305-15.

[13] HL7 template registry project. http://gforge.hl7.org/gf/project/templatesreg/docman/.

[14] Hoy D, Hardiker NR, McNicoll IT, Westwell P, Bryans A. Collaborative development of clinical templates as a national resource. Int J Med Inform. 2009 Apr;78 Suppl 1:S95-100.

Address for correspondence

Dr. Haitham KUSSAIBI
h.kussaibi@gmail.com - http://anapath.org