

# HL7 CDA R2 Based Document Sharing in Biomedical Research: Design of a Severe Adverse Event (SAE) Report Document

Carina Seerainer<sup>1,2</sup>, Hannes Kumpusch<sup>1</sup>, Michael Nitzlnader<sup>1</sup>, Günter Schreier<sup>1</sup>

<sup>1</sup>AIT Austrian Institute of Technology GmbH, Safety & Security Department, Austria

<sup>2</sup>University of Applied Sciences FH Joanneum, Department of eHealth, Austria

## Abstract

**Background:** The IT infrastructure requirements of Medical Research Networks (MRNs) are increasing with time and technological progress. The demand for modularity, interoperability and the support of standards calls for new strategies. In the healthcare domain, the application of Integrating the Healthcare Enterprise (IHE) profiles is a widely established approach. Hence, pursuing this approach may lead to a future-oriented system architecture enabling the integration of biomedical research with the healthcare domain. **Objectives:** The present paper focuses on the document sharing aspect based on HL7 Clinical Document Architecture (CDA), especially on the CDA document design of a Severe Adverse Events (SAE) report document taken from a clinical trial in the field of paediatric oncology. **Methods:** The CDA document design for the pre-existing SAE form was based on the CDA R2 specification and state-of-the-art implementation guidelines issued by the ELGA GmbH and IHE. To facilitate interoperability with third party information systems the focus was on using external vocabulary and code lists as well as predefined CDA structure templates. **Results:** A CDA document design for an SAE report was developed. The document was at least Level 2 and to a large extent Level 3 coded. **Conclusions:** eCRFs can be designed on the basis of CDA R2 while interoperability in large MRNs has to be ensured through validation against strict business rules. Further questions regarding IT infrastructure which result from particular requirements of MRNs have to be dealt with in the future.

## Keywords

Medical Research Network, eCRF, Clinical Trial, Severe Adverse Event Report Form, IHE XDS

## Correspondence to:

### Any Co-Author

AIT Austrian Institute of Technology GmbH, Safety & Security Department, eHealth,  
Reininghausstraße 13/1, 8020 Graz, Austria

EJBI 2012; 8(3):45–50

## 1 Introduction

Today, progress in biomedical research is depending on large, interdisciplinary research projects, involving many people in various roles and specialist disciplines. Adequate information and communication technology (ICT) support is becoming more and more important as the size and dimensionality of the data sets, the regulatory requirements and the complexity of the research processes are increasing [1].

Since custom made ICT systems cannot be built from scratch for each research endeavour, an infrastructure ap-

proach needs to be adopted. This means, ICT systems have to be established in a way primarily characterised by reuse and configuration, rather than by new developments.

Data collection in clinical trials was traditionally based on so called paper Case Report Forms (CRF). Today, with the omnipresence of Electronic Data Capture (EDC) systems, we speak of electronic CRF (eCRF) [2]. In most cases the underlying IT concept is a web-based central database.

The AIT Austrian Institute of Technology GmbH (AIT) is experienced in providing IT infrastructure for

Medical Research Networks (MRNs). With the International Society of Paediatric Oncology European Neuroblastoma Research Network (SIOPEN-R-NET) AIT developed an extensive web-based IT system for interdisciplinary biomedical research in the field of neuroblastoma across Europe and beyond. The infrastructure is based on a multi-tier architecture. Several clinical trials and modules operate on a single central database [3]. Requirements on IT infrastructure solutions for MRNs like SIOPEN-R-NET change over time and with technological progress. Modularity, interoperability, the support of standards as well as the link to healthcare IT infrastructure turn into crucial factors when it comes to providing future oriented solutions for an MRN. In order to serve these needs with predefined approaches that have already been accepted and established in healthcare, an Integrating the Healthcare Enterprise (IHE) based infrastructure for document sharing in biomedical research networks is currently being explored. Besides document sharing, IHE integration profiles also facilitate the implementation of other modules of major importance for MRNs, including management of identity, rights and roles, medical images, patient privacy and informed consent. Standards-based electronic document sharing in an MRN is a prerequisite for comprehensive analysis of data from multiple projects involved in the network.

## 2 Objectives

The overall objective is to design and develop an IHE-compliant MRN IT infrastructure to bridge the gap between different research groups in terms of specialisation and tumour type. The European Network for Cancer Research in Children and Adolescents (ENCCA), a Network of Excellence funded by the European Union's 7th Framework Programme (FP7) can be seen as a potential field of application. Together with other European partners, AIT is responsible for the development of the ENCCA IT infrastructure [4].

The present paper focuses on the document sharing aspect, thus the objective was to develop syntax and semantics of a Clinical Document Architecture (CDA) document based on the structure and content of a pre-existing eCRF from a clinical trial carried out within SIOPEN-R-NET. After a workflow analysis of one of the clinical trials carried out within SIOPEN-R-NET, the Severe Adverse Event (SAE) form was chosen as an example report to be designed according to the CDA specification issued by Health Level Seven International (HL7).

## 3 Methods

To collect all data elements relevant for an SAE report CDA document, the corresponding eCRF from the Low and Intermediate Risk NBL European Study (LINES), which is carried out within SIOPEN-R-NET, was analysed.

The CDA document was designed based on the CDA Release 2 (R2) specification and several state-of-the-art CDA implementation guides (IGs). The guiding documents included current CDA IGs for the Austrian Electronic Health Record (ELGA Discharge Letter Full Support and Laboratory Report) [5, 6] and the IHE Patient Care Coordination (PCC) Technical Framework [7].

The structure of the CDA document was developed using Eclipse Instance Editor [8]. The Eclipse Instance Editor offered W3C and HL7 CDA schema as well as HL7 Model Interchange Format (MIF) validation functionality while editing the CDA document. Additional validation was carried out using the IHE Gazelle CDA validation service [9].

In order to facilitate interoperability with third party information systems the focus was set on using external vocabulary and codes including Logical Observation Identifiers Names and Codes (LOINC) as well as predefined section and entry structure templates.

## 4 Results

Based on a pre-existing eCRF, a CDA document for the reporting of SAE was developed which is at least Level 2 and to a large extent Level 3 coded. The external code lists were partly drawn from the supporting CDA IGs. The Electronic Transmission of Individual Case Safety Reports (ICSR) IG, issued by the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) in cooperation with the HL7 Regulated Clinical Research Information Management (RCRIM) was chosen as a resource for code lists issued by the U.S. Food and Drug Administration (FDA), the National Cancer Institute (NCI) and the ICH [10]. Additionally, a manual search for LOINC and NCI Thesaurus terms was performed [11, 12]. The structure of sections and entries was mainly based on templates that were drawn from ELGA IGs and the IHE PCC Technical Framework. Internally developed code lists were identified by means of Object Identifiers (OIDs) from an OID sub tree held by AIT.

The header of the SAE report was based on the required header elements in the CDA R2 specification. Additionally, metadata were included, that were required to register a document in an IHE XDS (Cross-Enterprise Document Sharing) infrastructure. In the header, the document was identified using a unique ID and a code element taken from the NCI Thesaurus. No demographic patient data were included in the header, the patient was only identified using a unique ID.

The structured body held the actual medical report data in a structured way. It contained Level 2 and Level 3 sections. In Level 3 sections, there was an observation element designed for each medical finding. The arrangement of the sections was based on the corresponding main parts of the eCRF.

Table 1 shows details on structure and content of the

Table 1: Details on structure and content of the structured body of the CDA document "SAE report"

Title	Content	CDA Level	Applied External Code Lists	Applied Structure and Codes to identify a section	%
Treatment at Time of SAE	medical treatment at the time the SAE occurred	2	LOINC	section code from ELGA IG discharge letter	2
Seriousness Information	severity of SAE	3	NCI Thesaurus	severity actCode from IHE PCC IG	<1
Death Report	cause and date of death	3	LOINC	LOINC section code	2
SAE Diagnosis	date and name of diagnosis, CTCAE grade, outcome, possible explanation for SAE	3	SNOMED CT ICD-10 ICH	entry structures and codes partly from ELGA IG discharge letter and IHE PCC IG  The observations for each diagnosis are grouped within an organizer element	11
SAE Description and Classification	detailed description of the SAE as well as the information whether the event was expected or unexpected	2-3	-	-	<1
Treatment of SAE	type and details of treatment	2	NCI Thesaurus	-	2
[SAE Medication]	The structure of the SAE Medication section can be used for study chemotherapy, other study treatment and relevant concomitant medication.	3	NCI Thesaurus, ICH code lists	structure from IHE PCC  structure of "consumable" from ELGA IG discharge letter  The observations for each drug reported are grouped within an organizer element	79
Relevant Medical History	date and name of diagnosis including causal relationship to the SAE	3	LOINC	LOINC section code from ELGA IG	2
Comments	further issues	2	-	-	<1

structured body of the SAE report. The percentage in the rightmost column indicates the fraction of data elements in each section compared to an average fully populated SAE form within the LINES trial. In connection with the column entitled "CDA Level" it illustrates that more than 90% of all data elements are Level 3 coded. Figure 1 shows a snippet of the CDA document.

HL7 CDA schema and MIF validation was successfully carried out using Eclipse Instance Editor. Additional

schema validation using the IHE Gazelle testing service confirmed the validity of the SAE report.

## 5 Discussion and Lessons Learned

Today's translational research projects are often characterised by an increasing number of research team members, sometimes from different research institutions and diverse regulatory and organisational frameworks, who

```

<section classCode="DO CSECT" moodCode="EVN">

  <code code="SERINF"
  codeSystem="1.3.6.1.4.1.231.54.3.1.130.2.0.1.1.2"
  codeSystemName="Lines document section"
  displayName="Seriousness Information"/>

  <title>Seriousness Information</title>
  <text>results in persistent of significant disability/in capacity</text>

  <entry>

    <observation classCode="OBS" moodCode="EVN">

      <code code='SEV' displayName='Severity' codeSystem='2.16.840.1.113883.5.4'
      codeSystemName='ActCode' />

      <value xsi:type="CD" code="C21007"
      displayName="DISABILITY"
      codeSystem="2.16.840.1.113883.3.26.1.1"
      codeSystemName="NCI Thesaurus" />

    </observation>

  </entry>

</section>

```

Figure 1: CDA document snippet (SAE Seriousness Information)

deal with the various forms or segments of the whole treatment process. This is particularly applicable for research dealing with complex diseases and treatments, e.g. in the field of paediatric oncology. An increasing part of the research is carried out on the “basic science level”, e.g. involving biomolecular aspects. In fact, the whole data collection task is more and more broken down into diverse actors, belonging to different departments or even institutions. Interoperability is, therefore, crucial.

The form based concept is still alive and attractive since it segments the entire data space into different parts characterised by structuring aspects like:

- Parts of the human body or anatomical regions (e.g. the heart)
- Particular procedures (e.g. surgery)
- Results from contributing disciplines (e.g. laboratory data, imaging studies)
- Core concepts of medical care (e.g. diagnosis, treatment)
- Regulatory concepts (e.g. adverse events)
- Phases of the healthcare process (e.g. medical history, annual follow-up)

To stay in line with these concepts, a standardised document-based research infrastructure approach seems to be well suited for future research applications in order to capture the interdisciplinary aspect of the process and to provide modularity.

There have been attempts on interoperable research infrastructures in the past [1]. Although, so far, health-care as such was the focus of all major Standardisation Developing Organisations (SDO) including HL7 and interoperability initiatives like IHE, they have also started to deal with biomedical research as an important field for the future. Since the value of standardisation (or the costs of a lack thereof) has also been recognised for the research domain, new standardisation initiatives specific to the research domain have been created, e.g. the Clinical Data Interchange Standards Consortium (CDISC). Based on these activities a number of studies and pilots have been performed to design and to showcase interoperability in ICT for research systems.

In the current phase of the project, HL7 CDA R2 was mainly chosen because of its flexibility. While CDISC Operational Data Model (ODM) specialises in data exchange in regulated biomedical research, it does not deal with broader aspects of clinical research [2]. To serve the needs of strictly regulated biomedical research, ODM follows a trial protocol driven approach. On the other hand, CDA follows an event driven approach. CDA doc-

uments can therefore also be used in a less regulated environment characteristic to interdisciplinary, investigator driven academic biomedical research networks [13]. Since CDISC standards are already widely supported by EDC system vendors, the document sharing architecture has to be adapted to facilitate a coexistence of both standards.

The flexibility of CDA entails its generic characteristics. A successful schema validation carried out with the designed SAE document alone cannot be seen as an absolute indicator for interoperability. Detailed business rules, including supported code lists and structure templates have to be converted into schematron rules for deeper interoperability verification.

The SAE form was chosen for several reasons. As opposed to a disease evaluation form previously designed as a Level 2 CDA document, the SAE report is a standard eCRF present in almost all clinical trials. Current regulatory developments in the realm of pharmacovigilance are leading to standardised reporting of adverse events. The corresponding eCRFs are therefore similar in every biomedical trial. Another reason is the process of data capture dealing with adverse events which is a multi-step process, likely to involve a number of different parties [14]. Besides study centres, concerned parties may also include pharmacovigilance centres, the sponsor of the trial and regulatory authorities.

Developing a common set of codes for an eCRF was one of the greatest challenges in this work. To increase interoperability between systems and machine-readability of data, the percentage of Level 3 coded sections still has to be increased. At the same time, the number of different supported code systems has to be decreased in order to enhance maintainability and alleviate version management.

Another challenge not dealt with in this paper is the process of anonymisation and the provision of data to third parties for meta analyses and other research purposes that also face major interoperability challenges [15].

## 6 Next Steps

Further tasks at this stage of the project include developing a CDA generator tool and creating a sample implementation guide for the SAE report. The generator tool will support the workflow from a standardised non-CDA input format to the IHE XDS document repository via a source adapter [16]. With this step completed, an HL7 CDA schema validation service will be provided.

Having shown the feasibility of designing CDA documents for pre-existing eCRFs, the future perspective involves the design of CDA documents for a broader set of eCRFs and the development of schematron file sets (templates) as well as their integration into a validation service. Similar to the architecture of the RE-USE project the Cross Enterprise Template Sharing (XTS) integration profile may be used to provide respective templates of the target system [17]. Additionally, research has to be done

on how to plan and define the features of eCRFs optimised for the generation of CDA documents.

The ultimate goal should be to focus the design efforts on specific aspects of a given biomedical research problem (for example how a certain biomarker relates to the long-term survival of cancer patients) and to deal with related aspects in a standardised and therefore efficient way; for example to define or utilise “standard forms” for standardised aspects including SAE management, toxicity assessments, etc.). This requires deconstructing the trials of the given domain into well-segmented and self-contained parts (forms) which can be retrieved from repositories like building blocks.

Standards have to be developed to foster interoperability on all three levels, i.e. syntax (data formats and communication protocols), semantics (metadata, code lists, vocabularies, ontologies) and pragmatics (processes).

Finally, the ideas and processes of standardisation in electronic data capture constantly have to be communicated to the users so as to gain acceptance and to ensure long-term success of this architectural approach.

## 7 Acknowledgement

This work was partly funded by the European Commission under FP7-HEALTH-F2-2011 Contract no. 261474 (ENCCA - European Network for Cancer Research in Children and Adolescents).

## References

- [1] Ohmann C, Kuchinke W. Future Developments of Medical Informatics from the Viewpoint of Networked Clinical Research. *C. Methods Inf Med* 2009;48:45–54.
- [2] Richesson R L, Nadkarni P. Data Standards for Clinical Research Data Collection Forms: Current Status and Challenges. *J Am Med Inform Assoc* 2011;18:341e346.
- [3] Schreier G, Messmer J, Rauchegger G, Modre-Osprian R, Ladenstein R. A Web-based Platform for Interdisciplinary Biomedical Research. *Front Biosci.* 2009 Jan 1;14:2738-46.
- [4] ENCCA [Internet]. European Network for Cancer Research in Children and Adolescents[cited 2012 Jun 05]. Available from: <http://www.encca.eu>.
- [5] ELGA [Internet]. CDA Implementation Guide Discharge Letter Version 2.0[updated 2011 Oct 10; cited 2012 Jun 04]. Available from: [http://www.elga.gv.at/fileadmin/user\\_upload/uploads/-download\\_Papers/Harmonisierungsarbeit/-ELGA\\_CDA\\_Entlassungsinformation\\_AErztlich\\_2.00\\_FWGD.pdf](http://www.elga.gv.at/fileadmin/user_upload/uploads/-download_Papers/Harmonisierungsarbeit/-ELGA_CDA_Entlassungsinformation_AErztlich_2.00_FWGD.pdf).
- [6] ELGA [Internet]. CDA Implementation Guide Laboratory Report Version 2.0[updated 2011 Oct 10; cited 2012 Jun 04]. Available from: [http://www.elga.gv.at/fileadmin/user\\_upload/uploads/-download\\_Papers/Harmonisierungsarbeit/-ELGA\\_CDA\\_Laborbefund\\_2.00\\_FWGD.pdf](http://www.elga.gv.at/fileadmin/user_upload/uploads/-download_Papers/Harmonisierungsarbeit/-ELGA_CDA_Laborbefund_2.00_FWGD.pdf).
- [7] IHE [Internet]. IHE Patient Care Coordination Technical Framework Volume 2 Revision 7.0 - Final Text [updated 2011 Sep 09; cited

- 2012 Jun 04]. Available from: [http://www.ihe.net/-Technical\\_Framework/upload/IHE\\_PCC\\_TF\\_Rev7-0\\_Vol\\_2\\_2011-09-09.pdf](http://www.ihe.net/-Technical_Framework/upload/IHE_PCC_TF_Rev7-0_Vol_2_2011-09-09.pdf).
- [8] hl7book.net [Internet]. Eclipse Instance Editor [updated 2009 Jul 19; cited 2012 Jun 04]. Available from: [http://hl7book.net/-index.php?title=Eclipse\\_Instance\\_Editor](http://hl7book.net/-index.php?title=Eclipse_Instance_Editor).
- [9] gazelle.ihe.net [Internet]. CDA External Validation Service Front-end; c2010-11 [cited 2012 Jun 04]. Available from: <http://gazelle.ihe.net/-EVSClient/cda/validator.seam?cid=467>.
- [10] ICH [Internet]. Electronic Transmission of Individual Case Safety Reports (ICSR) Implementation Guide Version 3.01 [updated 2011 Jun 16; cited 2012 Jun 04]. Available from: [http://www.pmda.go.jp/-ich/e/step3\\_e2br3\\_guideline\\_draft\\_e.pdf](http://www.pmda.go.jp/-ich/e/step3_e2br3_guideline_draft_e.pdf).
- [11] LOINC [Internet] LOINC Search Browser. [cited 2012 Jun 16]. Available from: <http://search.loinc.org/>.
- [12] NCI [Internet] NCI Thesaurus Browser. [cited 2012 Jun 16]. Available from: <http://ncit.nci.nih.gov/>.
- [13] El Fadly A, Daniel C, Bousquet C, Dart T, Lastic P-Y, Degoulet P. Electronic Healthcare Record and Clinical Research in Cardiovascular Radiology. HL7 CDA and CDISC ODM Interoperability. AMIA AnnuSymp Proc. 2007; 2007:216-220.
- [14] London J W, Smalley K J, Conner K, Smith J B. The Automation of Clinical Trial Serious Adverse Event Reporting Workflow. Clin Trials. 2009 Oct;6(5):446-54. Epub 2009 Sep 8.
- [15] Kubick W R. Big Data, Information and Meaning. Applied Clinical Trials, 2012; 21:26-26. [cited 2012 Jun 21]. Available from: <http://www.appliedclinicaltrialsonline.com/-appliedclinicaltrials/article/articleDetail.jsp?id=757976>.
- [16] IHE [Internet]. IT Infrastructure Technical Framework Revision 8.0. [updated 2011 Aug 19; cited 2012 Jun 05]. Available from: [http://www.ihe.net/Technical\\_Framework/-upload/IHE\\_ITI\\_TF\\_Rev8-0\\_Vol1\\_FT\\_2011-08-19.pdf](http://www.ihe.net/Technical_Framework/-upload/IHE_ITI_TF_Rev8-0_Vol1_FT_2011-08-19.pdf).
- [17] El Fadly A, Rance B, Lucas N, Mead C, Chatellier G, Lastic P-Y, et al. Integrating Clinical Research with the Healthcare Enterprise: From the RE-USE Project to the EHR4CR Platform. J Biomed Inform. 2011;44 Suppl 1:S94-S102.