Designing Clinical Concept Models for a Nationwide Electronic Health Records System For Japan

Shinji Kobayashi[1], Naoto Kume[1], Takahiro Nakahara[2] and Hiroyuki Yoshihara[1]

[1]Department of Electronic Health Record, Graduate School of Medicine, Kyoto University, Kyoto, Japan

[2] Kyushu dental University, Fukuoka, Japan

Abstract

Objectives: We developed an electronic health records (EHR) system for regional healthcare in 2000. This EHR stores the health records of more than 6,300 patients in two regions of Japan; however, clinical updates and improved interoperability with other clinical standards, such as HL7 or others are needed. In 2015, this EHR system was upgraded to create a nationwide-scale healthcare data repository to improve the interoperability of clinical data with openEHR technology.

Methods: The clinical data in our EHR system has 16 components constructed with Medical Markup Language (MML) standards and periodic mass screening for employees and students. Therefore, we constructed mindmaps of the clinical MML and surveillance data to analyse the concept models. Based on mindmap analysis, we designed archetypes of the concepts identified using Ocean Archetype Editor. The artefacts were mainly quoted from the openEHR clinical knowledge manager (CKM). As the archetypes on CKM did not include all MML semantics, the archetypes were newly designed to complement the semantics of the MML

Results: We developed clinical information models by archetypes that semantically equalled the EHR system. Twenty-one MML components/modules and concept models using 99 archetypes were constructed for periodic mass screening services. Most of the archetypes were quoted from CKM; however, 22 archetypes were specialised, and eight archetypes were newly designed. The reasons for specialisation were to adjust the demographics to Japanese and to extend the archetypes to the dental domain.

Conclusion: We constructed concept models with archetypes semantically equivalent to conventional data and developed new archetypes for mass screening by archetype technology. The suggested archetype technology improved the flexibility of the EHR system to cover the existing standards.

Keywords

Archetype; Data modelling; Electronic health records and systems; HL7; Standards adoption; Semantic interoperability

1 Background and Significance

In 2003, we developed an electronic health records (EHR) system called the "Dolphin project" for regional health providers to share clinical records and provide individual patient data to patients [1]. We also developed an XML-based clinical information standard called Medical Markup Language (MML) to share clinical data within and among hospitals [2, 3].

We reported that this EHR system was working successfully, with more than 2,100 patients in two regions in Japan in 2010 [4], and 6,300 patients were registered in

this system in 2017. Consequently, we launched a project to expand this EHR nationwide to collect clinical data from more hospitals, implement mass screening services for the nation, and reuse the data for research and public health purposes. However, the Dolphin system has three problems that limit its ability to meet the requirements of the next EHR system. First, the Dolphin system has not been sufficiently maintained or updated because it was built on a legacy web framework and was written in Perl language. Second, the Dolphin system adopted MML for the data models, but MML does not meet the clinical/technical demands for emerging clinical updates. Therefore, the Dolphin system lacks interoperability with other standards, such as Health

Level Seven (HL7) [5] and others [6, 7, 8]. Interoperability is an emerging issue for the new project to utilise HL7 standardised data available from healthcare systems.

OpenEHR is the basis of the ISO 13606 [9, 10, 11, 12, 13] standard that enables interoperability of clinical concepts [14] and has been widely adopted to build other standards [15, 16]. The Clinical Information Modeling Initiative, an HL7 working group, is developing a common clinical concept model using openEHR archetype technology [17].

2   Objectives

We decided to reconstruct the EHR system using a modern development process, with common clinical concept models based on the ISO 13606 standard with the openEHR specification, to improve flexibility and interoperability for collecting multiple types of formatted data with various standards and catch up requirements of the nationwide EHR project in Japan.

3   Method

The data sources for this EHR project were hospitals that were already involved in the Dolphin project and performed periodic mass screening of employees and students. Healthcare data were collected from healthcare providers and screening services. The data were to be formatted using 16 MML content modules and others. MML data modules were analysed and constructed with semantically equivalent archetypes [18]. The mass screening service for employees adopted the Health Level Seven Clinical Document Architecture Release 2 (HL7 CDA R2), so that the surveillance data were collected by HL7 CDA R2 [19]. Because the mass screening service for students did not adopt any standards, raw data were used. Figure 1 shows all datasets analysed with regard to item, structure, components, and concepts. Ultimately 40 mindmaps were constructed.

Based on mindmap analysis, archetypes and templates were designed using Ocean Archetype Editor [20] and Ocean Template Designer [21]. The artefacts were mainly quoted from the openEHR clinical knowledge manager (CKM) [22] with or without modifications to maintain interoperability with regard to other standards. As the archetypes on CKM could not cover all case semantics, the archetypes were newly designed to complement the semantics used.

4   Results

All datasets were constructed on 18 templates and mapped with each corresponding data source (Table 1). Archetypes were reused for multiple templates, and Table 2 shows common archetypes in representative cases.

Figure 2 displays an example of the mindmap for a dental examination of teeth in the mass screening for students.

Table 1 shows each case and the number of archetypes used to construct semantically equivalent data schema. All components were designed with 1–50 archetypes and were fully conformant with openEHR specifications [23]. The most complex model was the mass screening for students, which required 50 archetypes. Mass screening for students requires a dental examination (Table 1). Figure 3 provides an overview of the archetype hierarchy in for mass screening of students. The report information model was designed for use with general reporting documents, but it had to be redesigned for use with specific reports, such as imaging reports, pathological diagnosis reports, and other diagnostic reports.

Table 3 shows the details of archetypes used to construct logical data models. A total of 92 archetypes were needed to construct the models, and 62 archetypes were quoted from CKM without any changes. However, 22 archetypes had to be
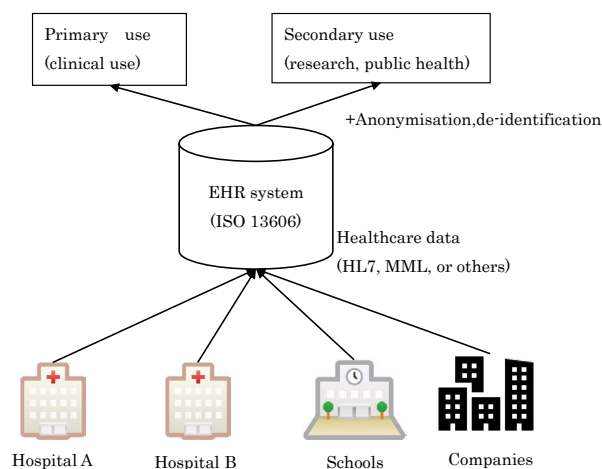

Figure 1: Nationwide electronic health records (EHR) system. Healthcare data will be collected from hospitals, schools, and companies by HL7, MML, original format for mass screening, or others. The EHR system stores healthcare data in the ISO 13606/openEHR repository for primary/secondary use with/without anonymisation or de-identification.

Table 1: Number of archetypes used for each concept.

Use case	Number of archetypes
Personal information	8
Creator information	7
Patient information	9
Health insurance	10
Diagnosis record	2
Lifestyle information	5
Basic clinical information	5
Initial consultation	14
Progress course	15
Surgical operation record	13
Clinical summary	19
Laboratory test	11
Report information	18
Referral letter	22
Mass screening for employee	36
Mass screening for students	50

Each case was analysed and constructed as an archetype to a template.

Table 2: Frequently reused archetypes in multiple cases.

Archetypes	Total	Diagnostic report	Progress notes	Clinical summary	Surgical operation record	Flow sheet	Referral	Mass screening for employee	Mass screening for students
CLUSTER.organisation-mml	26	6		6	5	3	6		
CLUSTER.citation	21	1	8	7	1		4		
EVALUATION.problem_diagnosis	15	1	1		1		1	2	9
EVALUATION.citation	13	1	4	5			3		
CLUSTER.person_name-mml	12	3		3	4		2		
CLUSTER.individual_professional-mml	10	3		3	2		2		
EVALUATION.clinical_synopsis	10	1	1	3	1	2	2		
CLUSTER.device	9				4				5
EVALUATION.health_risk	8							8	
OBSERVATION.progress_note	8		3	2			3		
INSTRUCTION.medication_order	7		4	1			2		
CLUSTER.address-japan	7			3	2		2		
OBSERVATION.exam	6		1	1				2	2
CLUSTER.telecom_details-mml	6			2	2		2		
OBSERVATION.story	5		1	1			2	1	
OBSERVATION.body_weight	4					1		2	1
INSTRUCTION.request	4	1	3						
CLUSTER.medication_preparation	4		2	1			1		
CLUSTER.daily_timing	4		2	1			1		

The common archetypes listed above were used for multiple cases. Organisation archetypes are commonly used to describe the facility, department, or ward. Citation archetypes are used to point to external links to images or other relative resources. Problem/diagnosis archetypes are also frequently reused.

modified, and eight archetypes were created from scratch to conform to the concept models. Reasons for modifying archetypes had to do mainly with localising for demographics, insurance claims, and the inclusion of specific information for dentistry (Table 4).

Because Health Level Seven version 2.5 (HL7 2.5) messages were also mapped to the constructed archetypes, this EHR system obtained the HL7 2.5 interface. Although the interface of the old EHR system was only MML, the newly developed EHR system used the following interfaces mapped to ISO 13606/openEHR archetypes:

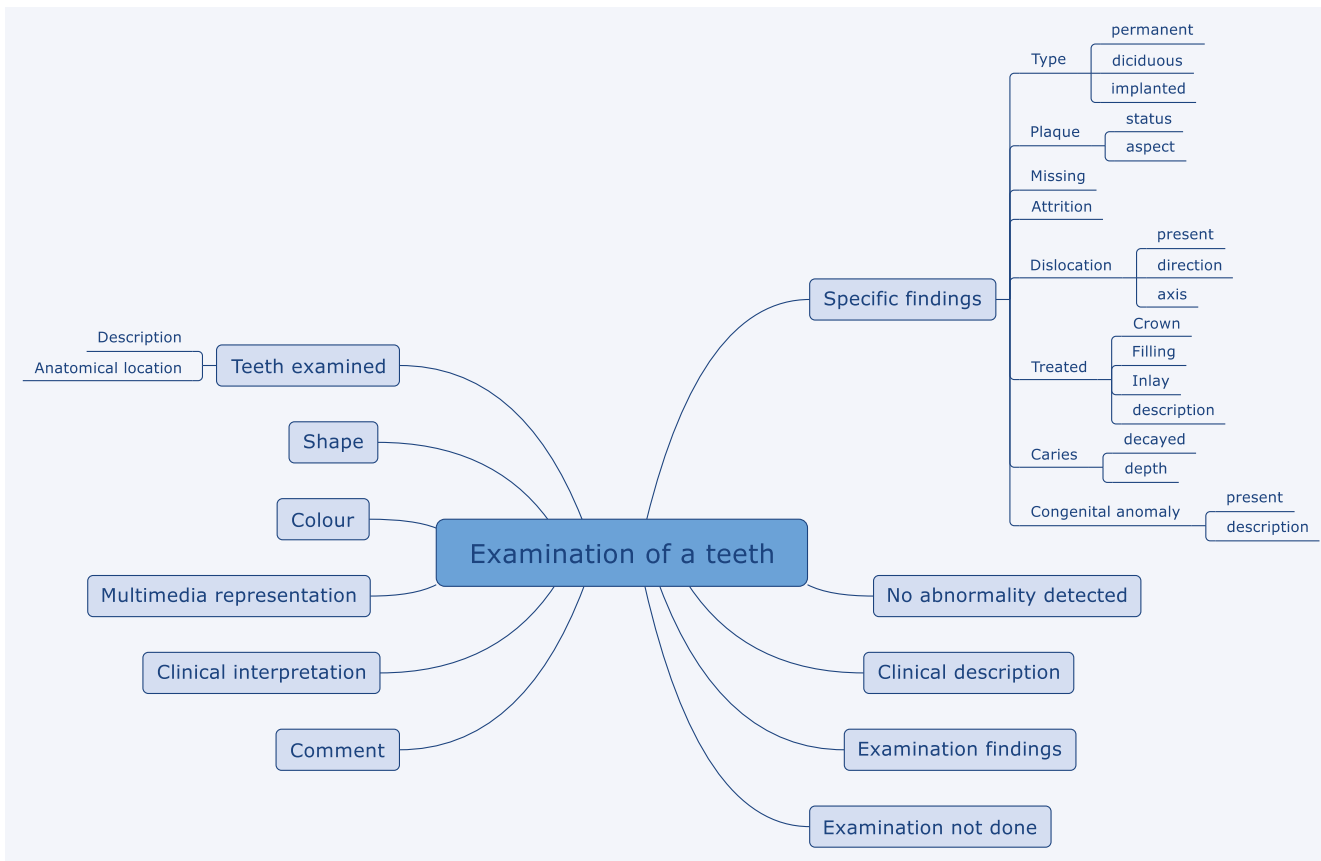• ISO 13606/openEHR native API

• MML

Figure 2: Mindmap for the examination of teeth.

A concept model for clinically examining teeth was developed into a mindmap. The concept was taken from the openEHR Clinical Knowledge Manager and specialised for dental examination-specific items.
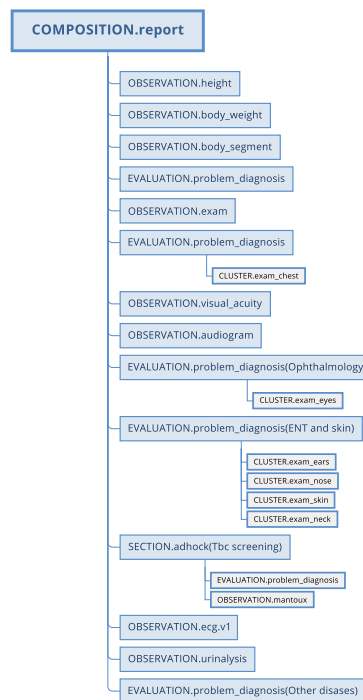

Figure 3: Overview of the hierarchy of archetypes in mass screening for students.

Table 3: Archetype construction using the case information model.

Archetype	Count of use
No change	62
Specialised	22
Newly created	8

Archetypes were mainly derived from the openEHR clinical knowledge manager. Sixty-two archetypes were used as-is, ten required specialisation, and eight archetypes were newly created.

Table 4: Example of rationales for specialised archetypes.

Concept model	Rationale
Patient name	To record ideographic (Kanji), syllabic (Furigana), and alphabetic names
Address	To record addresses in ideographic, syllabic, and alphabetic form
	To add the field for prefecture code
Telecoms	To adjust the item hierarchy to suit Japanese telecoms
Procedure	To match claim information to insurance systems
Medication	To record Japanese-style preparation and claims information for insurance purposes
Report	To record the report date, name of requester, responder to context field
Dental examination	To record specific expression for dentists

- HL7 2.5

- HL7 CDA R2

- Original data format for mass screening

5   Discussion

An EHR system requires semantic interoperability for multiple use cases, such as point of care, clinical research, and public health [24]. To achieve interoperability, EHR should have a commonly agreed upon logical model [25]. Therefore, the openEHR project has developed the CKM to share archetype concept models [26] to improve the interoperability of EHR.

The CKM is also a repository for international collaborative development of clinical concepts, and the CKM reported that 679 archetypes have been developed (cited on 19 December 2017). Archetypes on CKM are designed for general international use; however, in our case, the archetypes had to be localised mainly to express local demographics as well as the inclusion of dental examinations. As demographic expressions in CKM are widely diverse from MML, it was not possible to cover all of them. Specialising the existing archetypes was a reasonable way to localise them. Thus, although the existing HL7 and MML data models captured in the archetypes had an impedance mismatch in terms of information granularity, it was possible to devise equivalent forms with 92 archetypes; in fact, 67% (62/92) of them were adopted with no modifications. Although 1,821 users are registered with CKM, there are only 34 dentists (CKM statistics page, retrieved on 19 December 2017). Thus, there are few archetypes for dentistry information on CKM, and we

had to modify the archetypes to record data from dental examinations.

Moreover, the artefacts can be modified with other archetypes to describe other cases immediately, suggesting that creating a standardised clinical concept in the openEHR project would have more potential than making an EHR system with existing data models. However, we needed more clinical concepts to meet the unique demand of local clinicians and specialists.

Because archetype concept models have fine-grained items, they are capable of covering multiple standardised formats, such as HL7 2.5 and HL7 CDA, among others. As archetypes must be interoperable, the concept models in this study could be used with other standards. In addition, the old EHR system had only an MML interface, but the new EHR system uses five interfaces with archetype technology.

We are currently implementing an EHR system with these archetype concept models and, in the near future, plan to demonstrate interoperability for secondary use.

6   Conclusion

We constructed concept models with archetypes for a nationwide EHR system, which is semantically equivalent to the legacy EHR. Mass screening of the newly developed archetypes was conducted, and 11 dental archetypes were modified based on existing archetypes. The suggested archetype technology improved flexibility for collecting health data from multiple existing standards.

7   Clinical Relevance Statement

The developed clinical concept models can be used to describe clinical contents. Using these models would clarify the semantics of clinical descriptions and assessments and improve machine processing for clinical data. Thus, this project should contribute to clinical data management and the aid in the discovery of novel findings from EHRs.

8   Conflict of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

9   Human Subjects Protections

The study was performed in compliance with the World Medical Association Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects and Japanese laws and guidelines for privacy and security.

10   Acknowledgements

References

[1] Takada A, Guo SJ, Tanaka K, Sato J, Suzuki M, Suenaga T. et al. Dolphin project--cooperative regional clinical system centered on clinical information center. J Med Syst. 2005; 29: 391-400.

[2] Guo J, Araki K, Tanaka K, Sato J, Suzuki M, Takada A. et al. The latest MML (Medical Markup Language) Version 2.3 - XML-based standard for medical data exchange/storage. J Med Syst 2003; 27: 357–66.

[3] Guo J, Takada A, Tanaka K, Sato J, Suzuki M, Suzuki T. et al. The development of MML (Medical Markup Language) version 3.0 as a medical document exchange format for HL7 messages. J Med Syst 2004; 28: 523-33.

[4] Takemura T, Araki K, Arita K, Suzuki T, Okamoto K, Kume N. et al. Development of fundamental infrastructure for nationwide EHR in Japan. J Med Syst. 2012; 36: 2213–8.

[5] Health Level Seven® International. http://www.hl7.org/

[6] Integrating the Healthcare Enterprise (IHE). https://www.ihe.net/

[7] Clinical Data Interchange Standards Consortium. https://www.cdisc.org/

[8] DICOM® (Digital Imaging and Communications in Medicine). http://www.dicomstandard.org/

[9] ISO 13606-1:2008. Health informatics -- Electronic health record communication -- Part 1: Reference model.

[10] ISO 13606-2 :2008. Health Informatics – Electronic health record communication – Part 2: Archetype interchange specification.

[11] ISO 13606-3:2009. Health informatics -- Electronic health record communication -- Part 3: Reference archetypes and term lists.

[12] ISO/TS 13606-4:2009. Health informatics -- Electronic health record communication -- Part 4: Security.

[13] ISO 13606-5:2010. Health informatics -- Electronic health record communication -- Part 5: Interface specification.

[14] Beale T. Archetypes: Constraint-based domain models for future-proof information systems. Proceedings of Eleventh OOPSLA Workshop on Behavioral Semantics: Serving the Customer. 2002: 16-32.

[15] ISO 13940:2015. Health informatics -- System of concepts to support continuity of care.

[16] Goossen WTF. Detailed clinical models: Representing knowledge, data and semantics in healthcare information technology. Healthc Inform Res. 2014; 20: 163–72.

[17] Policies: Selected CIMI policies and decisions. http://opencimi.org/policies_cimi

[18] Kobayashi S, Bosca D, Kume N, Yoshihara H. Reforming MML (Medical Markup Language) standard with archetype technology. Indian J Med Inform. 2014; 8: 57–60.

[19] Dolin RH, Alschuler L, Boyer S, Beebe C, Behlen FM, Biron P V. et al. HL7 Clinical Document Architecture, Release 2. J Am Med Inform. 2006; 13: 30–9.

[20] Informatics O. Ocean Archetype Editor. http://wiki.oceaninformatics.com/confluence/display/TTL/Archetype+Editor+Releases

[21] Informatics O. Ocean Template Designer. http://www.openehr.org/downloads/modellingtools

[22] openEHR Clinical Knowledge Manager. http://www.openehr.org/ckm/

[23] The openEHR. openEHR specification program; 2015. http://www.openehr.org/programs/specification/latestreleases

[24] ISO/TS 18308: 2011. Health informatics — Requirements for an electronic health record architecture.

[25] ISO/TR 20514: 2005. Health informatics —Electronic health record —Definition, scope and context.

[26] Garde S, Chen R, Leslie H, Beale T, McNicoll I, Heard S. Archetype-based knowledge management for semantic interoperability of electronic health records. Stud Health Technol Inform. 2009; 150: 1007–11.