HIR

Healthcare Informatics Research

Contribution of Clinical Archetypes, and the Challenges, towards Achieving Semantic Interoperability for EHRs

Archana Tapuria, DCH, M.Tech[1], Dipak Kalra, PhD[1], Shinji Kobayashi, PhD[2]

[1]Centre for Health Informatics and Multiprofessional Education, University College London, London, UK; [2]Department of Bioregulatory Medicine, Ehime University Graduate School of Medicine, Ehime, Japan

Objectives: The objective is to introduce 'clinical archetype' which is a formal and agreed way of representing clinical information to ensure interoperability across and within Electronic Health Records (EHRs). The paper also aims at presenting the challenges building quality labeled clinical archetypes and the challenges towards achieving semantic interoperability between EHRs. **Methods:** Twenty years of international research, various European healthcare informatics projects and the pioneering work of the openEHR Foundation have led to the following results. **Results:** The requirements for EHR information architectures have been consolidated within ISO 18308 and adopted within the ISO 13606 EHR interoperability standard. However, a generic EHR architecture cannot ensure that the clinical meaning of information from heterogeneous sources can be reliably interpreted by receiving systems and services. Therefore, clinical models called 'clinical archetypes' are required to formalize the representation of clinical information within the EHR. Part 2 of ISO 13606 defines how archetypes should be formally represented. The current challenge is to grow clinical communities to build a library of clinical archetypes and to identify how evidence of best practice and multi-professional clinical consensus should best be combined to define archetypes at the optimal level of granularity and specificity and quality label them for wide adoption. Standardizing clinical terms within EHRs using clinical terminology like Systematized Nomenclature of Medicine Clinical Terms is also a challenge. **Conclusions:** Clinical archetypes would play an important role in achieving semantic interoperability within EHRs. Attempts are being made in exploring the design and adoption challenges for clinical archetypes.

Keywords: Electronic Health Records, Semantics, Health Information Management

I. Introduction

1. The Need for Interoperability between EHRs

The communication, interoperability and analysis of Electronic Health Records (EHRs) is of growing global importance as the functionality and use of EHR systems increases. Longitudinal EHRs can improve the quality and safety of care to individuals, provide the knowledge needed to improve the efficiency of healthcare services and population health programmes, and accelerate clinical research.

Clinical information about individual patients is inevitably collected across multiple care settings and within diverse

heterogeneous EHR repositories. Integrating this information is a recognised health informatics challenge, and has been the subject of over 20 years of international research.

The requirements for EHR information architectures have been consolidated within ISO 18308 [1] and adopted within the ISO 13606 EHR interoperability standard. This five-part standard defines a generic information model for representing part all of an individual's EHR [2], vocabularies for some of its information properties [3], a security policy model for representing the consent and permissions for access to the EHR information being communicated [4], and an interface specification for requesting and providing EHR information [5].

However, a generic EHR architecture cannot ensure that the clinical meaning of information from heterogeneous sources can be reliably interpreted by receiving systems and services. Therefore, the organization (hierarchy) of clinical information within the EHR also needs to be formalized in the form of clinical archetypes. These clinical model specifications are commonly known as archetypes. Part 2 of ISO 13606 defines how archetypes should be formally represented for interoperability [6], drawing on pioneering work of the openEHR

Foundation [7]. SemanticHealthNet, an EU Network of Excellence, is exploring many of these design and adoption challenges [8].

There are many clinical and health service drivers for integrated EHRs, in which the cumulative health information of an individual patient can be accessed from any point of care delivery [9]. As shown in Table 1, all of drivers in each perspective, to a greater or lesser extent, contribute to the business case for the present international investment in e-Health and interoperability.

2. The ISO EN 13606

1) The ISO EN 13606 EHR communication standard
This five-part standard ensures that the information governance [10], provenance and clinical context [11] of each health record entry is communicated consistently [12,13], and helps towards building a safer system [14] (Table 2).

2) The ISO EN 13606 reference model
The ISO EN 13606 reference model that underpins the

Table 1. Clinical and health service drivers for integrated Electronic Health Records in e-Health and interoperability

Perspective	Driver
Internal and external environment	Manage increasingly complex clinical care
	Support team-based care
	Connect multiple locations of care delivery
Safety	Deliver evidence-based healthcare
	Improve safety by reducing errors and inequalities
	Improve safety reducing duplication and delay
Patient-centered	Enrich population health management and prevention
	Empower and involve citizens
	Protect patient privacy
Effectiveness	Improve cost effectiveness of health services
	Better inform and exploit bioscience research

Table 2. Five parts of the ISO EN 13606 EHRs communication standard

Part	Standard description
CEN 13606 Part 1	The reference model: a scalable model for representing personal health information
CEN 13606 Part 2	Archetype interchange specification: an archetype model is used to represent archetypes when communicated between archetype repositories within EHRs. Archetypes define (constrain) legal combinations of the reference model (described in Part 1)
CEN 13606 Part 3	Reference archetypes and term lists
CEN 13606 Part 4	Security: a methodology for specifying the privileges necessary to access the EHR data
CEN 13606 Part 5	Exchange models to request and provide an EHR extract, an archetype or an audit log extract

EHR: Electronic Health Record, CEN: European Committee for Standardization.

exchange of EHR information. This model is an information model that contains a set of classes and attributes. The 13606 approach is to represent the reference model as a set of Unified Modeling Language diagrams. The outcome is a hierarchical model reflecting the hierarchical nature of real heath records. The 13606-1 reference model is composed of a number of classes which build on each other to provide the representation of an EHR Extract. These classes include EHR extract class, recorded components, and other classes such as audit, record linkages, access policy, and message.

The EHR Extract class identifies who the extract is about and what system the extract has been extracted from. The EHR data is (optionally) comprised of directory of folders and made up of compositions (records authored together and committed to the EHR), demographics, and access control policies. And the recorded component is a super class of other classes. These record component classes build from a simple element to more and more complex structures. These classes include element, item, entry, section, composition, and folder, as described in Table 3. Using such a hierarchy, an EHR Extract can be built starting with elements and then successively grouping into items, entries, sections, compositions, and folders.

However, the way in which clinical information is represented within the EHR also needs to be formalised. This includes, for example, definitions for individual data elements and how they should be combined, which data elements should be mandatory, what kinds of data value are appropriate (a term, a quantity, etc.) units of measurement, value ranges, valid terms. These clinical model specifications are commonly known as archetypes.

II. Case Description

1. Clinical Archetypes
Clinical archetypes provide a systematic approach to repre-

senting the definition of any EHR data structure. The archetype approach is itself somewhat generic, and could represent data structures for any profession, specialty or service.

Archetypes have been adopted by European Committee for Standardization (CEN) as a European Standard (EN 13606 Part 2) and as an international ISO standard (ISO 13606 Part 2) following more than a decade of research in Europe and Australia and some further development by the openEHR Foundation. The standards set out the requirements for clinical domain modelling and formalisms that meet the requirements can be regarded as conforming. This allows more modern paradigms, such as that provided by the Semantic Web to be used [15].

Clinical archetypes represent a formal statement of agreed consensus on best practice when recording clinical data structures. They are specifications of the knowledge data and their inter-relationships that play an important role in determining how clinical information is represented and organized inside EHRs and when they are communicated between systems. They are content specifications expressed in terms of constraints on a specific reference model. Archetypes will often also influence the way in which clinical data are managed within individual EHR systems, how users enter data and how data are presented.

An archetype provides a standardized way of specifying EHR clinical data hierarchies and the kinds of data values that may be stored within each kind of entry. It defines (or constrains) relationships between the data values within an EHR data structure, expressed as algorithms, formulae or rules. It may logically include other archetypes, and may be a specialization of another archetype. In order for it to be managed and used appropriately, its metadata needs to define its core concept, purpose and use, evidence basis, authorship, versioning and maintenance information [16].

Archetypes offer a tractable way of binding generic EHR models to compositional terminology. They provide target

Table 3. Level of record component classes

Level	Description
Element	A record component containing a single data value.
Item	A single element, a list of elements, a cluster or a list of clusters. Item therefore allows the representation of a wide range of data structures.
Entry	Items recorded for a single recording in the EHR (e.g., a single observation).
Section	Entries grouped together.
Composition	Set of record components authored during users' clinical sessions and stored in the EHR (e.g., a progress note).
Folder	Group of records. Folders can include other folders, compositions or used to organize (selected subset of) the EHR extract.

EHR: Electronic Health Record.

knowledge representations for use by guideline and care pathway systems, and so support knowledge level interoperability: systems may interoperate not only at the data level, but also at the level of intended clinical meaning. EHR components identify the archetypes used when the data were created, and/or to which they map, which aids future interpretation, analysis, querying.

Figure 1 shows an example of an archetype for adverse reaction to medication, authored using an archetype editor developed by the University of Linköping in Sweden. In the main central panel of this figure is a hierarchical data structure that represents the components of the documentation of an adverse reaction. For each node in the hierarchy the icon used indicates the data type of the patient-specific value: textual, coded, date or time, quantity, etc. The right hand panes show (upper pane) the number of occurrences that are permitted within instances of EHR data and, for coded entries (lower pane), the terminology values that may be used. Other panes (not shown) permit further constraints to be applied, such as numeric ranges and measurement units. This archetype therefore defines the "shape" within an EHR for representing adverse reactions, and thereby offers some predictability to any application or system component that needs to query EHR data to obtain adverse reaction information.

A new initiative from ISO is to develop quality requirements for Detailed Clinical Models, a generic clinical representation approach that could embrace archetypes, Health Level Seven International (HL7) templates, and other equivalently expressive modeling formalisms. The transformation of archetypes from ISO EN 13606 to openEHR ones and vice versa has been addressed by Martinez-Costa et al. [17]. Their results show that such an exchange is possible [17].

The openEHR Foundation has grown a worldwide community that is defining archetypes. To adopt archetypes for local healthcare environment, local activities have engaged in their domains. At first, Japanese team launched in 2007, and now many countries have their domestic openEHR community. Most of them are domestic or geographical community, but Spanish/Portuguese team covers Latin America, too. Each community acts as promoters of the openEHR community in their domain, and cooperates with other Standards Developing Organisations to utilise clinical standards. The reasons to localise archetypes are categorised by three main reasons: 1) domestic healthcare environment, 2) language translation, and 3) clinical specialist demands.

Concerning domestic healthcare environment, each country has its own healthcare system, such as insurance, general practitioner (GP) program, and emergency systems. It is necessary for their domain to specialise archetypes for their environment. As for language translation, archetypes are designed for a multi-lingual environment. However, the terms are not matched one by one in translations. At last, for clinical specialist demands, the archetypes, which are shared using a Clinical Knowledge Manager repository, have been designed for GP use. Clinical specialists, e.g., hematologists, need more detailed archetypes for their use.

In Japan, the community has designed more than 50 archetypes for 6 model disease amongst national intractable disease repository program. For domestic use, health insurance and demographics needed to adjust Japanese conventional style. The Clinical Knowledge Manager provides qualified
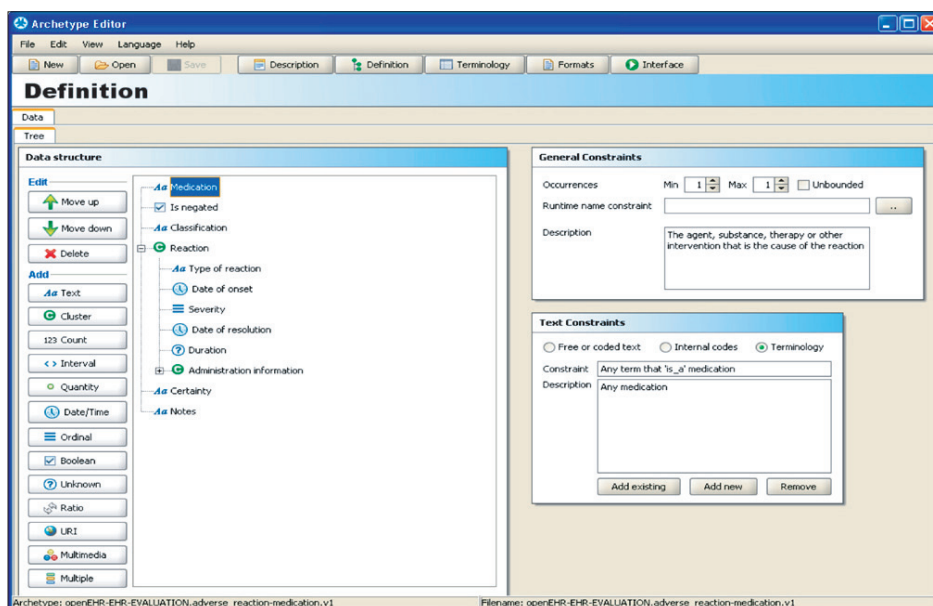

Figure 1. Example screen showing the main portion of an archetype for adverse reaction (to medication).

archetypes for this project, but twenty archetypes were specialised to record disease specific information. One of the most frequent specialisations is to determine severity for each disease and criterion. Each local activity is now considering to make a repository to share it artefacts. To utilise archetype to local condition, specialisation is necessary, but it is more important to discuss how to normalise such information for other domain. Worldwide collaboration would incubate qualified clinical domain concepts.

The current challenge is to identify how evidence of best practice and multi-professional clinical consensus should best be combined to define archetypes at the optimal level of granularity and specificity for wide adoption. Patients and social care communities will increasingly be involved in sharing records and so need to be included when archetypes are being defined. Definitive archetypes will need to be quality labelled and disseminated. SemanticHealthNet, an EU Network of Excellence, is exploring many of these design and adoption challenges [18].

2. Semantic Interoperability Challenges

Over the past few years it has become clear that some the greatest areas of challenge in semantic interoperability for the EHR lie firstly in the definition and sharing of suitable data structure definitions and secondly in the binding of them to terminology. By binding an archetype to a part of a terminology system (for example, to specify that the value for a node called "location of fracture" must be a term from a hierarchy of bones in the skeletal system) it should be possible to foster consistency and reliability (correctness) in how EHR data are represented, communicated, and interpreted.

However, this is only partially true in practice. There are a number of difficulties that mean the binding of an archetype to a constrained set of terms can be problematic. Record structures and terminology systems have been developed in relative isolation, with very little or no cooperation on their mutual requirements or scope, resulting in overlapping coverage and often a clumsy fit.

Successful information models are focused on meeting specific use cases and requirements. However, the contemporary health information reference models have been defined for a very broad range of use cases and scopes, in order to be applicable to needs right across healthcare, such as the HL7 Reference Information Model. Whilst on one level this can be seen as a measure of success, it inevitably results in such models having many optional properties. This in turn creates the risk that the model will be used inconsistently by different vendors within multiple systems because there are multiple ways in which clinical data may be represented.

Context & recursion is sometimes poorly modelled, risking ambiguity or nonsense (for example, being able to define an Entry about a patient's mother but to include within it some data about a different relative). Thus, semantic interoperability is a basic challenge to be met by the new distributed and communicating health information systems (HIS). Lopez and Blobel [19] has provided an appropriate development framework for semantic interoperable HIS; the usability of which has been exemplified in a public health scenario.

There are also a number of problems with clinical terminologies. Modern terminologies attempt to provide a comprehensive coverage of healthcare, although many terms cannot be precisely defined, or agreed upon; if many of the terms in a terminology are not used consistently what use is it having them there? Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) in particular defines over 1.2 million relationships between its 400,000+ unique concepts and the complexity of these relationships may exceed what can be safely implemented and reliably used. We do not yet have enough operational user experience to show otherwise. Terminologies also often include multiple representations for the same clinical concept, adding to the semantic interoperability challenge. The binding of SNOMED CT to HL7 artefacts is also known to be problematic, and the publication of a draft standard known as TermInfo [20] has highlighted the potential complexity of any resolution to this.

It must be remembered that human interpretation of clinical notes and correspondence on paper has for decades been sufficient to meet most shared care needs without the use of formalized record structures or terminology. However, we know for example that prescribing systems can help prevent error [21,22], and that they require access to comprehensive allergy, diagnosis and past medication data from the EHR in a processable form in order to offer safety alerts to the prescriber. Other scenarios include the use of electronic guidelines and care pathway systems, and clinical audit systems.

Many of the safety-critical scenarios requiring the computational support of health IT are knowledge management failings or communication gaps. Particular points in the clinical process which are often not currently documented in computable forms and which are not always done well (i.e., in which care steps might be delayed or omitted, or dangers introduced), and for which sufficient knowledge now exists to improve safety, include new medication prescriptions, reminders, evidence based care, care transfers, and care coordination, as listed in Table 4.

The development of clinical content to meet these needs, for example as archetypes binding to SNOMED CT, will require the additional development of ontology resources in order

Table 4. Safety issues requiring the computational support of health IT

Issue	Description
New medication prescriptions	The safety of prescriptions is often compromised by a lack of comprehensive information on concurrent medication (including purchased drugs) and details of known allergies, in particular since this information might be split across multiple care organizations and health records.
Reminders and prompts	Reminders and prompts for overdue or overlooked healthcare actions and interventions
Evidence-based care	The use of clinical guidelines and other forms of evidence to determine the optimal management strategy and care pathway for a given patient
Care transfers	Referrals and within-team workflow, such as the degree of urgency and the expectations of the referring clinician from another team member
Care coordination	Ensuring that a high-level view can be taken of distributed (multi-team) care to protect against duplication, delay and incompatible interventions

to map out the domain and its semantic sub-components, to ensure that coverage is complete without overlap, and also to semantically index each archetype for discovery and to help identify semantic equivalence. Although informatics research still has much to offer to these challenges, a focus on the most useful and tractable clinical areas would be best undertaken through large scale pilot projects that can partner academia with multiple vendors and multi-site care teams across Europe, and validate the results across larger patient groups. These pilots must address clinically driven use cases that meet genuine gaps in safe or high quality care delivery.

Clinical engagement on a wide scale is now essential, to help grow libraries of consensus and evidence-based clinically useful archetypes. Vendor and e-Health program engagement is needed to establish the means to validate such archetypes in achieving the safe and meaningful exchange of EHR data between systems.

III. Discussion

There is widespread recognition globally that a formalized and scalable means of defining and sharing clinical data structures is needed to achieve the value of investment in e-Health. Vendor approaches based on archetypes are gaining acceptance as the best supported methodology for defining these structures, reflected in its international standardization.

Large and comprehensive sets of clinical data models are now needed that cover whole domains in a systematic and inclusive way, catering for the inevitable diversity of use cases and users but helping to foster consensus and best practice. For these to be endorsed they need to be quality assured, and to be published and maintained by reliable, certified, non-partisan sources. This process will include organizational

governance, artefact governance and quality assured processes, and will build on the approach and criteria presented in this deliverable.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Acknowledgments

References

1. Kalra D. ISO/DIS 18308 Requirements for an Electronic Health Record reference architecture. Geneva, Switzerland: International Organization for Standardization; 2010.
2. Kalra D, Lloyd D. ISO 13606 Electronic Health Record communication. Part 1: Reference Model. Geneva, Switzerland: International Organization for Standardization; 2008.
3. Kalra D. ISO 13606 Electronic Health Record communication. Part 3: Reference archetypes and term lists. Geneva, Switzerland: International Organization for Standardization; 2009.
4. Kalra D. ISO/TS 13606 Electronic Health Record Communication. Part 4: Security. Geneva, Switzerland: International Organization for Standardization; 2009.
5. Kalra D. ISO/EN 13606 Electronic Health Record communication. Part 5: Interface specification. Geneva,

Switzerland: International Organization for Standardization; 2010.

6. Kalra D, Beale T, Lloyd D, Heard S. ISO 13606 Electronic Health Record communication. Part 2: Archetype interchange specification. Geneva, Switzerland: International Organization for Standardization; 2008.

7. openEHR [Internet]. London: openEHR Foundation; c2013 [cited at 2013 Dec 1]. Available from: http://www.openehr.org.

8. SemanticHealthNet [Internet]. [unknown]: SemanticHealthNet; c2013 [cited at 2013 Dec 1]. Available from: http://www.semantichealthnet.eu.

9. Kalra D. Clinical foundations and information architecture for the implementation of a federated health record service [dissertation]. London: University College London; 2002.

10. Smith R. What clinical information do doctors need? BMJ 1996;313(7064):1062-8.

11. Patterson D, Ingram D, Kalra D. Information for clinical governance. In: Lugon M, Secker-Walker J, editors. Clinical governance: making it happen. London: The Royal Society of Medicine Press; 1999.

12. Waegemann CP. Medical Record Institute's survey of electronic health record trends and usage. In: Proceedings of Toward an Electronic Health Record Europe '99; 1999 Nov 14-17; London, UK. p. 147-58.

13. Audit Commission for Local Authorities and the National Health Service in England and Wales. A spoonful of sugar: medicines management in NHS hospitals. London: Audit Commission; 2001.

14. Kohn LT, Corrigan JM, Donaldson MS. To err is human: building a safer health system. Washington (DC): National Academies Press; 2000.

15. Hammond WE, Pollard DL, Straube MJ. Managing healthcare: a view of tomorrow. Stud Health Technol Inform 1998;52 Pt 1:26-30.

16. Rothschild AS, Dietrich L, Ball MJ, Wurtz H, Farish-Hunt H, Cortes-Comerer N. Leveraging systems thinking to design patient-centered clinical documentation systems. Int J Med Inform 2005;74(5):395-8.

17. Martinez-Costa C, Menarguez-Tortosa M, Fernandez-Breis JT. An approach for the semantic interoperability of ISO EN 13606 and OpenEHR archetypes. J Biomed Inform 2010;43(5):736-46.

18. Rector AL, Nowlan WA, Kay S. Foundations for an electronic medical record. Methods Inf Med 1991;30(3):179-86.

19. Lopez DM, Blobel BG. A development framework for semantically interoperable health information systems. Int J Med Inform 2009;78(2):83-103.

20. Health Level Seven International. Using SNOMED CT in HL7 Version 3: Implementation Guide, Release 1.5 [Internet]. Ann Arbor (MI): Health Level Seven International; c2009 [cited 2013 Dec 1]. Available from: http://wiki.hl7.org/index.php?title=Using_SNOMED_CT_in_HL7_Version_3;_Implementation_Guide,_Release_1.5.

21. Nightingale PG, Adu D, Richards NT, Peters M. Implementation of rules based computerised bedside prescribing and administration: intervention study. BMJ 2000;320(7237):750-3.

22. Bates DW. Using information technology to reduce rates of medication errors in hospitals. BMJ 2000;320(7237):788-91.