

Comparison and transformation between CDISC ODM and EN13606 EHR standards in connecting EHR data with clinical trial research data

Archana Tapuria¹, Philipp Bruland², Brendan Delaney³, Dipak Kalra⁴
and Vasa Curcin¹

Abstract

Objectives: Integrating Electronic Health Record (EHR) systems into the field of clinical trials still contains several challenges and obstacles. Heterogeneous standards and specifications are used to represent healthcare and clinical trial information. Therefore, this work investigates the mapping and data interoperability between healthcare and research standards: EN13606 used for the EHRs and the Clinical Data Interchange Standards Consortium Operational Data Model (CDISC ODM) used for clinical research.

Methods: Based on the specifications of CDISC ODM 1.3.2 and EN13606, a mapping between the structure and components of both standards has been performed. Archetype Definition Language (ADL) forms built with the EN13606 editor were transformed to ODM XML and reviewed. As a proof of concept, clinical sample data has been transformed into ODM and imported into an electronic data capture system. Reverse transformation from ODM to ADL has also been performed and finally reviewed concerning map-ability.

Results: The mapping between EN13606 and CDISC ODM shows the similarities and differences between the components and overall record structure of the two standards. An EN13606 archetype corresponds with a group of items within CDISC ODM. Transformations of element names, descriptions, different languages, datatypes, cardinality, optionality, units, value range and terminology codes are possible from EN13606 to CDISC ODM and vice versa.

Conclusion: It is feasible to map data elements between EN13606 and CDISC ODM and transformation of forms between ADL and ODM XML format is possible with only minor limitations. EN13606 can accommodate clinical information in a more structured manner with more constraints, whereas CDISC ODM is more suitable and specific for clinical trials and studies. It is feasible to transform EHR data in the EN13606 form to ODM to transfer it into research database. The attempt to use EN13606 to build a study protocol (that was already built with CDISC ODM) also suggests the possibility of using EN13606 standard in place of CDISC ODM if needed to avoid transformations.

Keywords

Operational data model, EN13606, electronic health records, clinical archetypes, interoperability

Submission date: 14 March 2017; Acceptance date: 13 April 2018

Introduction

Electronic health records (EHRs), promising to provide an ideal form of longitudinal patient health record, offer remarkable and enhanced opportunities for clinical research. The reuse of routinely collected clinical data in the form of EHRs for clinical research is being explored as part of the drive to make maximum use of

¹King's College London, UK

²Albert-Schweitzer-Campus, University of Münster, Germany

³Imperial College London, UK

⁴University College London (UCL), UK

Corresponding author:

Archana Tapuria, King's College London, London WC2R 2LS, UK.
Email: archana.tapuria@kcl.ac.uk



the EHR data for clinical trials and studies. The aim is to reduce the effort in extracting the EHR data, reduce duplication and errors in data entry, reduce the costs, increase data quality and facilitate small pragmatic trials. Automatic transfer of data from the EHR to the Clinical Trial Electronic Data Capture (EDC) would save many hours of arduous effort, especially for multi-site data-intensive clinical trials.

There are several challenges in integrating clinical trials with clinical EHR systems.¹ The heterogeneity of the structure and architecture used by the various EHR systems, the incompatibility of the clinical data standards used, the various choices of clinical terminologies and ontologies adopted by different EHR systems and the difference in the workflow and process of clinics and clinical studies make the interoperability and integration of EHRs with the clinical research EDC very challenging.²

In this regard, different approaches of metadata harmonisation were analysed to bring both worlds together: EDC systems are mostly able to handle and communicate its metadata and data in the Clinical Data Interchange Standards Consortium Operational Data Model (CDISC ODM) standard.³ To describe metadata in a comprehensive manner the ISO/IEC 11179 standard for metadata registries is used in several contexts such as the caDSR (Cancer Data Standards Repository) of the NCI (National Cancer Institute).⁴ Studies that are funded by the NCI are encouraged to publish their study metadata within the caDSR and to reuse already specified data elements and forms. Likewise, openEHR archetypes are used to fully specify a (medical) circumstance such as the paragon archetype of 'Blood Pressure', which subsumed the value itself but also additional information about the systolic, diastolic, pulse pressure or location.

Interoperability between ISO 11179 and CDISC ODM has been shown in work by Bruland et al.⁵ and transformation between CDISC ODM and openEHR archetypes has also been done by Bruland and Dugas.⁶ Hume et al.⁷ and Richesson et al.⁸ have suggested the current challenges in using the data standards in clinical research and addressed an ODM's limitations and strengths to support new trends in clinical research informatics. For example, ODM forms support only three levels of depth while HL7 CDA's nested observations can be unlimited in number. This disparity is at least partially a reflection of the difference between protocol-driven clinical research and the event-driven healthcare domain. ODMs also represent controlled terminologies differently. The HL7 CDA standard uses the HL7 Reference Information Model (RIM) to provide an external semantics source and ODM tends to define its own codes without explicitly accounting for semantics.⁹

In addition, diverse projects were tackling the subject of supporting clinical research by the standardisation and

harmonization of data models between healthcare and research worlds. The European FP7 TRANSFoRm project aimed at developing an infrastructure for a Learning Health System in European Primary Care,¹⁰ with concrete use cases in clinical trials, epidemiological studies and diagnostic decision support. In that context, TRANSFoRm developed a Randomised Controlled Trial module that integrated into several European EHRs,¹¹ providing automatic patient eligibility checking, part-filling of electronic Case Report Forms (eCRFs), managing study workflow and storing research data back into the EHR.¹² This was supported by a two-level modelling approach of Detailed Clinical Modelling, which is depicted on the first level as an information model, the Clinical Research Information Model, which defines the workflow and data requirements of the clinical research task, combined with the Clinical Data Integration Model, an ontology of clinical primary care domain^{13,14} that captures the structural and semantic variability of data representations across data sources. At the second level, archetypes are used to constrain the domain concepts and specify the implementation aspects of the data elements within EHR systems or patient registries. The two-level modelling approach, using the concept of archetype for detailed clinical content modelling, has been adopted by EN13606.¹⁵ The archetype defines the data elements that are required by specific application contexts, for example, different clinical studies. While EN13606 uses a hierarchical reference model,¹⁴ TRANSFoRm chose an event-based tabular structure for the reference model of the TRANSFoRm information models.¹⁶ The standards chosen for building the study design information models was CDISC ODM, as it was compatible with this reference model structure and represents the data collected in clinical trials and represents aspects of study design.¹⁷

A further example is the EHR4CR (Electronic Health Records for Clinical Research) project. The aim of this Innovative Medicines Initiative funded project was to reduce the cost of conducting clinical trials, through better leveraging routinely collected clinical EHR data. The approach to handling semantic interoperability was based on the realistic assumption that there will remain a co-existence of several standard information models (e.g. EN13606 information model and archetypes, openEHR, Health Level 7 (HL7) RIM and HL7 Fast Healthcare Interoperability Resources (FHIR) specifications, CDISC ODM, etc.) for representing EHR in systems (www.hl7.org). EHR4CR adopted a mediation model and mapping approach to a set of Common Data Elements (CDE) identified as frequently occurring in clinical research protocols.¹⁸ These CDEs were picked from several trials and the coverage in European EHR systems was investigated to foster the reuse of data.¹⁹

The HL7 RIM and EN13606 standards define the semantics of patient care data and clearly demonstrate the need for ‘layers of semantic expressiveness’.²⁰

Table 1 shows some more European health informatics projects focusing on semantic interoperability of EHRs and using different standards for their clinical information models.

In this paper, we are building upon the experiences of the EHR4CR project, EU FP7 TRANSFoRM project, and using as an exemplar the MRC INFORM clinical trial²¹ currently in development. In the following section, we describe the standards that we have used in our work in more detail.

CEN/ISO EN 13606

The CEN/ISO EN13606 is a European norm from the European Committee for Standardization also approved as an international ISO standard.²² The overall goal of the EN13606 standard is to define a rigorous and stable information architecture for communicating part or all of the EHR.²³ EN13606 follows an innovative Dual Model architecture. The former is structured through a reference model that is an object-oriented model used to represent the generic and stable properties of health record information.^{24,25} The latter is based on archetypes.²⁶

The EN13606 reference model is composed of building blocks or classes/entities such as Folder, Composition, Section, Entry, Cluster and Element as shown in Figure 1.²³

An archetype is a structured and constrained combination of entities of a reference model that represents a particular clinical concept, such as a blood pressure measurement or a laboratory analysis result. It provides a semantic meaning to a reference model structure. It is built by constraining the entities in the following different ways.

1. Constraints on the range of attributes of primitive types.
2. Constraints on the existence of attributes, that is, whether a value is mandatory for the attribute in run time data.

3. Constraints on the cardinality of attributes, that is, whether the attribute is multi-valuate or not.
4. Constraints on the occurrences of objects indicating how many times in runtime data an instance of a given class conforming to a particular constraint can occur.

Archetypes are specified using the Archetype Definition Language (ADL). This language provides an abstract syntax, which can be used to express archetypes for any reference model in a standard way. An archetype can include other archetypes and can be used in combination to design the templates for the forms.

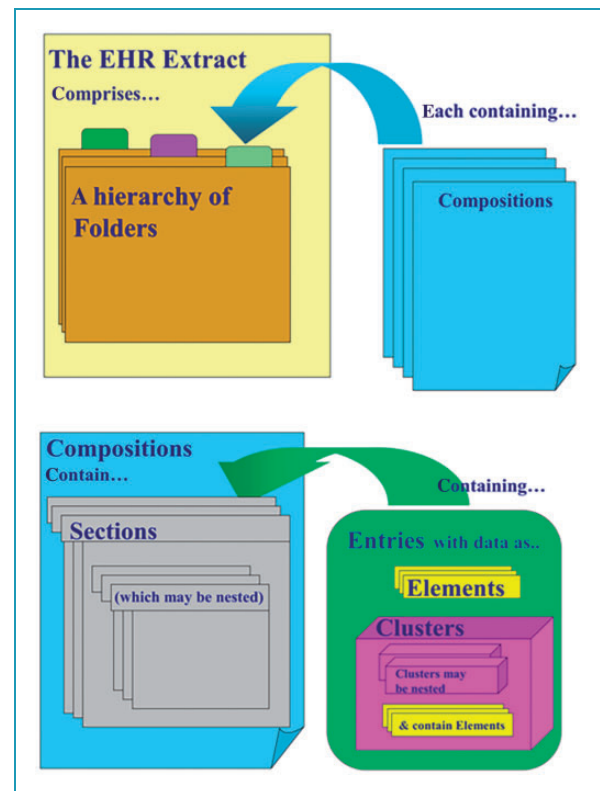


Figure 1. Structure of the reference model of EN13606 and EHR extract hierarchy.

Table 1. European health informatics projects focusing on semantic interoperability of EHRs.

Health informatics related projects	Standards used for clinical information models	Links
EURECA	HL7 v3, BRIDG	http://www.eurorec.org/RD/eureca.cfm
SALUS	HL7 CDA, ISO EN13606	https://www.sec-salus.eu/
Linked2Safety	HL7, OpenEHR	http://www.linked2safety-project.eu/

CDISC Operational Data Model

The Clinical Data Interchange Standards Consortium (CDISC) is an open, multidisciplinary, non-profit standards developing organization that has been working to develop and support global, platform-independent data standards that enable information system interoperability to improve medical research and related areas of healthcare.²⁷ CDISC have established a suite of standards as an end-to-end solution for clinical trials. These include specification of a trial protocol with Protocol Representation Model and its ODM representing a case report form (CRF), specification of study design with Study Design Model, specification of tabulated data with Study Data Tabulation Model and standardised sets of defined data elements from Clinical Data Acquisition Data Standards Harmonisation (CDASH).²⁸

The ODM is a vendor-neutral, platform-independent format for exchanging and archiving clinical study data. ODM is designed to facilitate the regulatory-compliant acquisition, archive and interchange of metadata and data collected in a study, and so is closely aligned with the schedule of activities. It includes all information (clinical data, along with its associated metadata, administrative data, reference data and audit information) necessary to share among different software systems during study setup, operation, analysis and submission and for long-term retention as part of an archive. The ODM is represented in eXtensible Markup language (XML) format and is designed to collect data from many different sources into one document. ODM has become the language of choice for representing case report form content.²⁹ ODM v1.3.2 is the most current version of the standard.

An ODM file consists of a tree of elements that includes the Subject Data, StudyEventData, FormData, ItemGroupData, ItemData and Annotation elements.

The ODM is composed of two major parts.

- The Metadata part defines what events, forms and questions a study is made up of. The metadata features are StudyEventDef, FormDef, ItemGroupDef and ItemDef.
- The second part is the patient data, which provides a data transport and storage mechanism for the actual clinical data as entered into the eCRFs.

It is commonly used in clinical trials for example, to archive data and metadata of clinical trials.³⁰ ODM is used in different EDC systems as well as in clinical data management systems used by the pharmaceutical industry.³¹ Especially in the context of clinical trials, CDISC standards are more established and many EDC systems already support ODM files.

Objectives

The approaches to standardise the structure of clinical information for EHRs and for clinical trials have historically been led by different standardisation bodies and have resulted in different families of standards and specifications for representing clinical care information and clinical trial information, even though these should be very similar in practice. Both EHR4CR³² and TRANSFoRm³³ have demonstrated the value of using EHR data for research, therefore making it increasingly important for the semantics of these two worlds to come together. Therefore, our research has been to investigate whether the standards used in each of the two domains are capable of being mapped to each other. The paper investigates the data interoperability between the two standards EN13606²² and CDISC ODM²⁷ and thereby also establish the feasibility of converting the EHR data available in the form of EN13606 into an ODM (which is the preferred choice of standard for clinical research) to transfer it into research databases. It also tries to study (with an example) the possibility to use EN13606-standards-based clinical archetypes instead of CDISC ODM to extract the required clinical information from the EHR sources and to make interoperability between the EHR data and the clinical research data even easier. Such a possibility might allow data captured once during clinical care to be reused for research purposes without duplicate data entry or transformation of data structures.

Methods

The specifications of EN13606 and CDISC ODM 1.3.2 were reviewed by clinicians and medical informatics professionals; a mapping model between the structure and components of both standards developed, and the similarities and differences noted.

As a proof of concept the information model to represent the INFORM clinical trial protocol was built with the EN13606 editor tool, to explore the feasibility and challenges in using it for the purpose for which ODM is widely used already. Also, an information model for the 'inclusion criteria' of 'INFORM' clinical trial protocol was created using both CDISC ODM and EN13606 Standards and compared. 'Hypertension' being the main inclusion criteria, an information model for 'Blood Pressure' was built. The ODM designer tool³⁴ was used for the information model created with CDISC ODM. The EN13606 editor called Object Dictionary Client (ODC) was used, which was developed in-house by University College London (UCL), the latest version of which is recently published, and the information models built with it are open source for use.³⁵

ODM is an XML-based standard, which is well described in the CDISC ODM XML schema definition. As EN13606 ADL is also available as XML format,

based on the mapping model we have developed an extensible stylesheet language transformation (XSLT) script to transform ADL into ODM including terminological bindings.

The other way around has been performed with the aid of ODM2ADL-converter,³⁶ which is part of the Portal of Medical Data Models, an information infrastructure to create, share and discuss medical documentation entities.³⁷ The converter produces openEHR archetypes and has been adapted in the way to result in EN13606 ones.

Transformations between EN13606 and CDISC ODM, and vice versa, have been reviewed by two medical informatics professionals.

To test the approach of converting metadata and clinical data from EN13606 to ODM, we have used anonymous clinical data from a comprehensive test study in EN13606, performed the mapping into CDISC ODM and imported the results into our CDISC ODM-compliant x4T-EDC system.³⁸

Results

Mapping between EN13606 and CDISC ODM v1.3.2

Figure 2 shows the mapping model with similarities and differences between the components and overall record structure of the two standards.

ODM ItemGroups contain similar elements/items of a specific clinical domain. Hence, an ItemGroup can be mapped to one archetype of EN13606. An ODM StudyEventDef or FormDef corresponds to a template

or group of archetypes in EN13606. An ODM ItemDef corresponds to an Element in EN13606. Mapping of EN13606 classes like ‘Composition, Section, Entry, Clusters’ to ODM is not possible as the ODM structure does not have classes corresponding to these. Figure 2 shows that the transformation of element names, descriptions, different languages, datatypes, cardinality, optionality, units, value range, and terminology codes is possible between EN13606 and CDISC ODM.

The datatypes provided by CDISC ODM are similar to that of EN13606, but EN13606 has a more extensive list of datatypes (Table 2). Though EN13606 has a more extensive list of datatypes, it is possible to map multiple datatypes from EN13606 to a datatype from CDISC ODM and vice versa. For example, ‘DateTime’, ‘PartialTime’ and ‘PartialDatetime’ from CDISC ODM can be mapped to ‘Date_time’ from EN13606. Similarly, ‘String’ and ‘Character’ from EN13606 can be mapped to ‘String’ from CDISC ODM.

ODM serves the purpose of capturing the clinical data and its representation in a fairly simple manner.³⁹ A single format provides all components needed to describe clinical research data with attribution requirements mandated by regulatory agencies. It reduces the number of unique file formats a clinical application needs to support.⁴⁰ ODM is part of a family of end-to-end standards. It is also a transport standard used for event-based messaging similar to HL7 FHIR, though it is not its primary purpose.

EN13606 reference model along with the archetypes can hold more detailed clinical data in a hierarchical structure with more specific constraints compared to the ODM. ODM forms support three levels of depth,

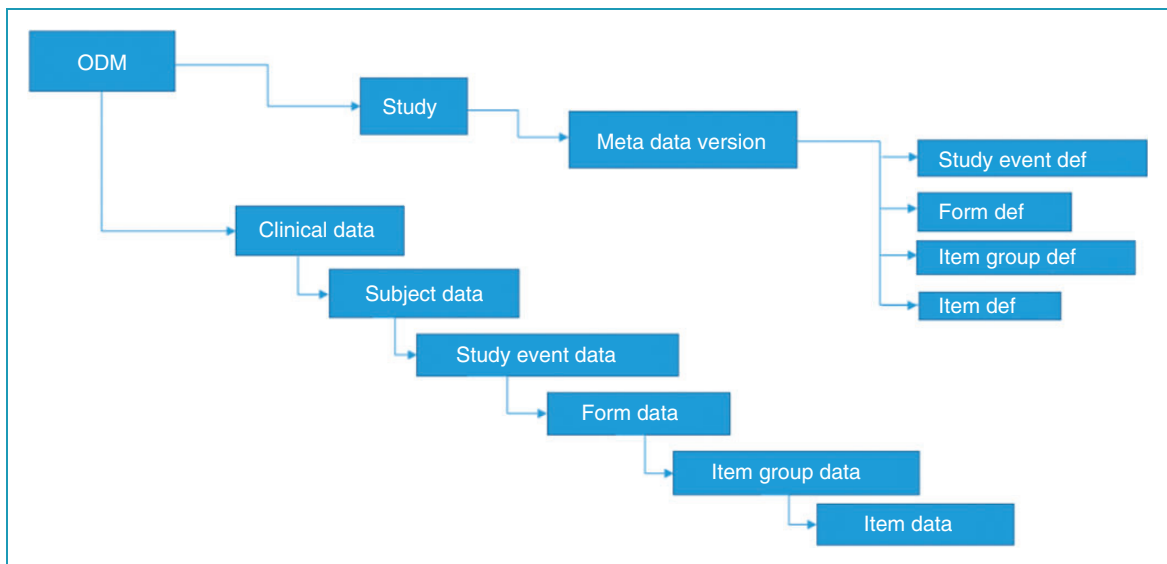


Figure 2. Structure of CDISC ODM metadata and data.

Table 2. Mapping between datatypes of CDISC ODM and EN13606.

CDISC ODM	EN13606	Description
Date	Date	
Time	Time	
PartialTime		
PartialDatetime		
Datetime	Date_time	
String	String	
	Character	
Boolean	Boolean	
Double	Double	
HexBinary	Array<BYTE>	
Base64Binary	Array <BYTE>	
HexFloat	Real	
Base64Float	Real	
	Integer	
IntervalDatetime	Interval <T> (interval of instances of any ordered type)	'T' represents Time
	List<T> (ordered, duplicates allowed)	
	Set<T> (unordered, no duplicates)	
	Hash <T, K> (keyed list of items of any type)	

while EN13606 archetype supports practically unlimited number. This reflects the difference between protocol-driven clinical research data and event-driven healthcare records. ODM makes use of the Alias element to capture semantic information, though it is a poor and unstructured way to capture semantic information. EN13606 tries to achieve semantic interoperability by standardising the structure and representation of the clinical data using archetypes.

The validation of the conversion showed that the converted ADL files could be opened with the EN13606 editor and were reusable in the ADL format. Semantic information obtained from ODMs Alias element was transferred to the term bindings element of ADL to ensure semantic interoperability. Similarly, it was possible to convert the ADL forms created with the ODC tool to ODM XML format

(with the medical data models tools) and these were reviewed using the ODM designer tool.

Information model for an example concept domain with CDISC ODM and EN13606

In the INFORM trial, the main inclusion criteria being 'hypertension', a 'blood pressure' information model needs to be built. First, the blood pressure archetype (Figure 4) was built with the ODC, an EN13606-based editor. Later an information model (Figure 5) for the same was built with the ODM designer tool. A nested tree structure of the data elements was built with the 13606 editor, whereas a 'list' structure was built with the ODM designer tool.

The ODM structure showed limitations in representing the information in the required hierarchy. Especially it was not possible to cluster the elements such as 'systolic blood pressure' and 'diastolic blood pressure', whereas in EN13606 structure, it is possible to practically form unlimited clusters within clusters to help represent the clinical data in the required hierarchy.

Using EN13606 editor to capture the INFORM trial protocol

Figure 6 shows a screenshot of the attempt to build INFORM clinical trial protocol with the EN13606 editor. It was possible to represent the study protocol along with the details of the schedules and visits of the participants. The main limitation observed is that there can be a number of ways in which the protocol, especially the study schedules and visits, could be represented (especially in terms of hierarchy).

Evaluating the mapping between EN13606 and CDISC ODM

As a proof of concept in the real world, we used anonymous clinical data from a comprehensive test study in EN13606, performed the mapping into CDISC ODM and imported the results out into CDISC ODM compliant x4T-EDC system.³⁸ The transformation of the EN13606 source study into ODM was successful. A resulting example of the import into our x4T-EDC study database is shown in Figure 7.

Discussion

Unlike CDISC ODM, the EN13606 standard is not specifically designed for clinical studies. The primary purpose of EN13606 is to define the structure of the EHRs, while CDISC ODM aims to capture data and metadata of clinical research data. CDISC ODM is also

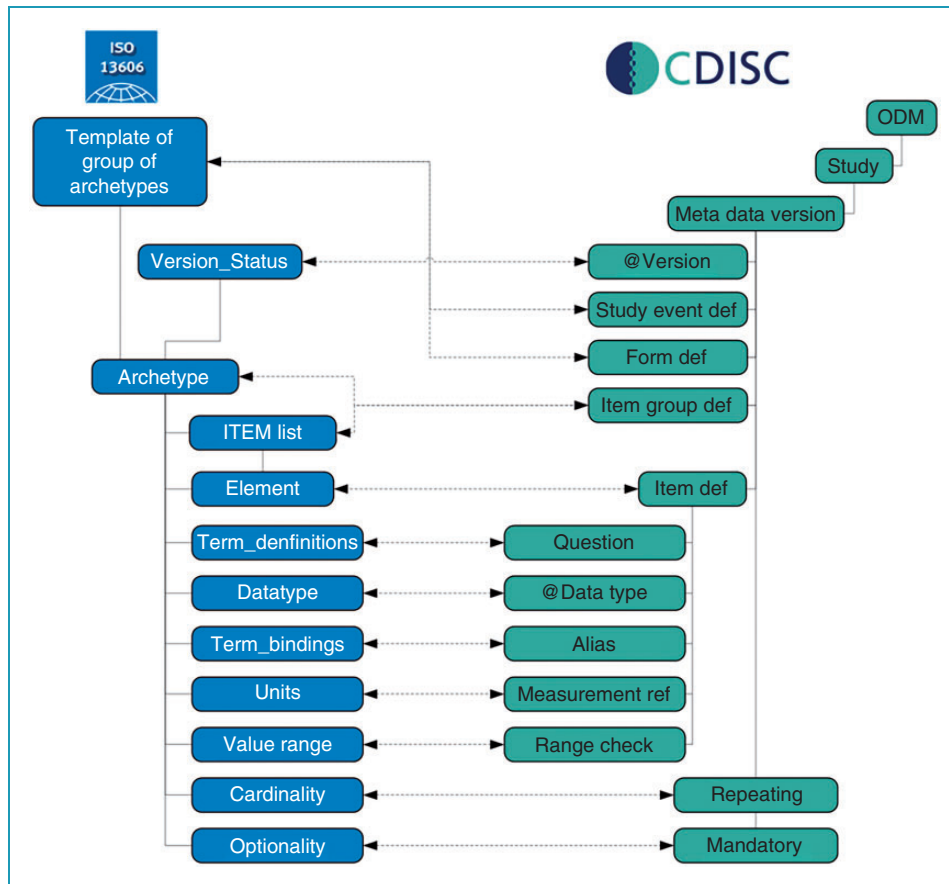


Figure 3. Mapping model between EN13606 and CDISC ODM v1.3.2.

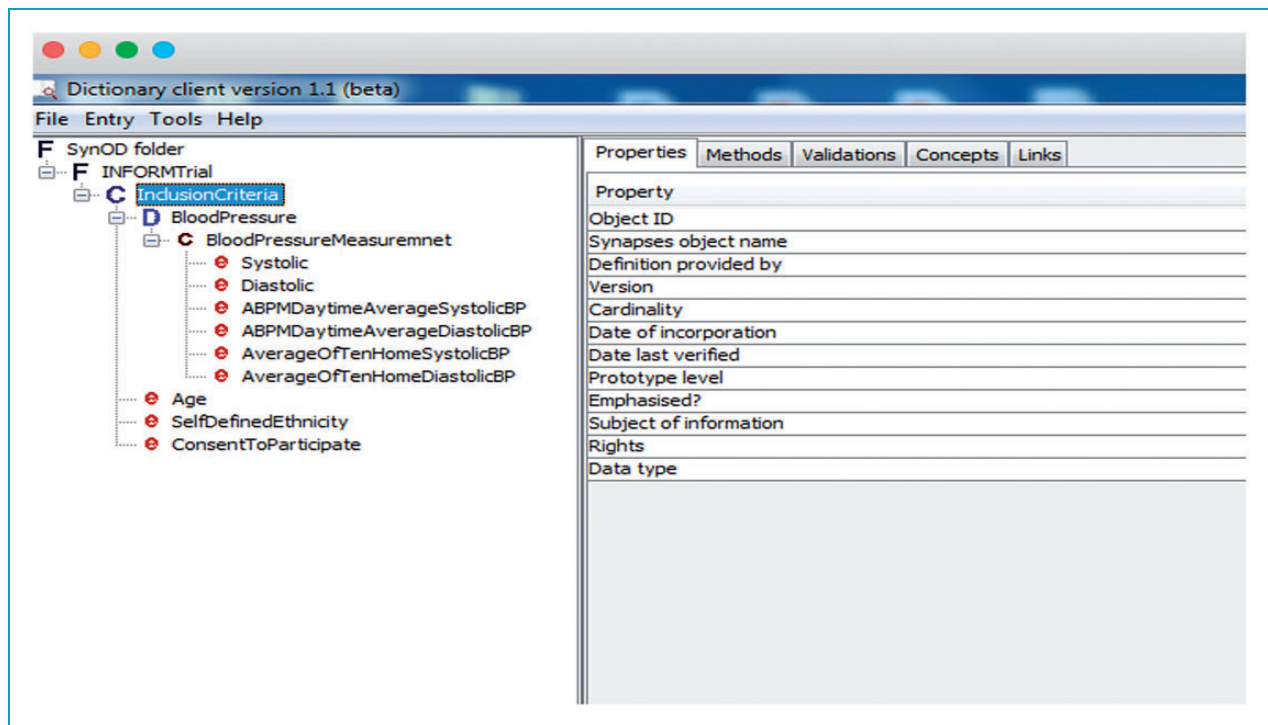


Figure 4. Screenshot of information model for blood pressure with ODC EN13606 editor tool.

Item ID	Name	DataType	Length	SignificantDigits	SASField
BP01	Systolic	integer		3	
BP02	Diastolic	integer		3	
BP03	Mean Arterial Pressure	integer		3	
BP04	Pulse Pressure	integer			
BP05	Comment	text			
BP06	Cuff Size	string			
BP07	Location	string			
BP08	Method	string			
BP09	Mean Arterial Pressure Formula	string			
BP010	Device	string			
BP011	Position	text			
BP012	Exertion	text			
BP013	24 hour average	float			
BP014	Date	datetime			

Figure 5. Screenshot of information model for blood pressure with ODM designer tool.

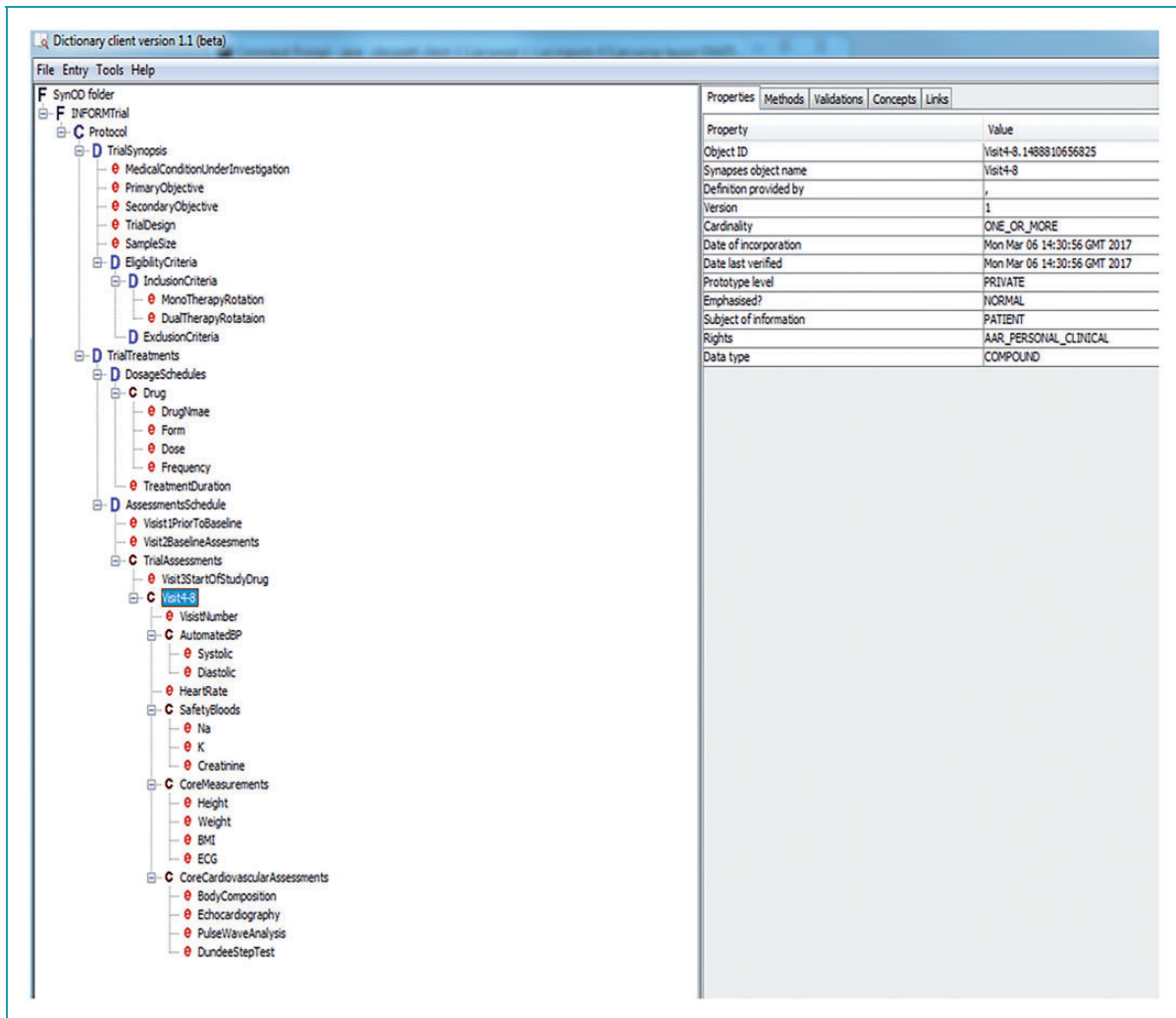


Figure 6. Screenshot of the archetype model with constraints built for the INFORM protocol with EN13606 editor. F: folder; C: composition; D: entry; c: cluster; e: element.

PID: 001 [Uniklinikum Münster]

Start / Daten / Patienten-ID: 001 / Formular: Baseline Visit Form

Common

Site number * 23

Subject ID * 001

Visit Date * 27.02.2010

Demographics

Date of Birth * 07.05.1957

Sex * Male Female

Race Caucasian

Smoking History

Check when the subject is a smoker * Ja Nein

Number of cigarettes per day Less Than 10 cigarettes per day 10 to 20 cigarettes per day Greater Than 20 cigarettes per day

Drinking History

Number of alcoholic drinks per day * Less Than 1 drink per day 1 to 2 drinks per day Greater Than 2 drinks per day

Physical Exam

Height [in] * 193

Weight [lbs] * 90

BMI *

Systolic blood pressure [mm Hg] 120

Diastolic blood pressure [mm Hg] * 80

Does the subject feel dizzy when standing up from a sitting position * Nein Ja

XRay

Server or File location of X-Ray photograph *

Complaints due to smoking

Figure 7. Screenshot of the x4T-EDC system showing the imported CDISC ODM metadata and clinical data of a sample patient.

a transport or exchange standard similar to HL7 FHIR, though it is not its primary purpose. While ODM provides a vehicle to communicate the study results back to the regulatory body, it lacks a rich-enough information model to capture the innate contextual information of the clinical study data.⁴⁰ The HL7 FHIR framework, which has been swiftly adopted by the healthcare community, looks to be the likely candidate for overcoming this challenge.⁴¹ Leroux and Lefort⁴¹ and Doods et al.⁴² have presented an approach to integrate the CDISC ODM standard with the FHIR resources to enrich longitudinal clinical study data extracted from ODM.

Nevertheless, we have shown that it is feasible to map data elements and forms based on CDISC ODM format to EN13606 EHR standards and vice versa. It is possible to represent the study definitions of a clinical trial protocol using EN13606. The transformation of EN13606 into ODM strengthens the data transfer between clinical routine world and research world. And it is easier to transfer data and have the semantic

and provenance of data clear when data is to come from a more hierarchically and structured format (EN13606) to a less hierarchic format (ODM) in order to move data from an EHR system to an EDC system. With regard to the Learning Health System, it is essential that knowledge gained from clinical research is returned into the healthcare domain.

EHR systems contain a huge amount of clinical data that is potentially eligible to be reused for secondary purposes.¹⁹ However, where the intention is to use healthcare data for secondary purposes such as clinical research, it is indispensable to consider the provenance, purposes of collection and the quality of routine healthcare data.^{43,44} In this regard, medical experts need to be involved in the process of identifying the required data in the appropriate clinical context. In order to support the data identification process, semantic annotation of data elements within primary healthcare systems (e.g. EHR) is a promising approach to easily discover the meaning and context of medical data.

Semantic interoperability plays a major role in the understanding and exchangeability of healthcare data. Beside the EN13606s structured information model in which data elements are specified, this standard allows the annotation of concepts with semantic codes of diverse code systems, whereas ODM has no native opportunity to place concept codes for elements within its hierarchy. The developers of CDISC ODM suggest using the free-text based ‘Alias’-element for this purpose. One of the largest repositories of ODM files, the Medical Data Models portal⁴⁵ attaches semantic concept codes with the aid of the Alias element to allow rediscovery and further analyses.^{6,46,47} However, this requires an agreement on how concept codes are exactly specified in the free-text attributes. More advanced solutions would be among others the definition of XML schema extensions within ODM. For instance, the NCI has published the CDISC CDASH elements in ODM with an ODM extension to assign their NCI thesaurus codes. Further investigations on the semantic level have been performed by Leroux and Lefort, providing ontological bindings to ODM elements.⁴⁸

Clinical archetypes are also a means of describing the database against which queries are run. How to cause the database to be written against those archetypes, or to impose an archetype structure on what is there already, is an interesting research area. It is a challenge for software engineers to choose the right database for data stored with clinical archetypes based on a particular standard. The EN13606 archetypes use ADL as the preferred format. The CDISC ODM has data stored as XML files and the commonly used databases with EN13606 are MySQL, PostgreSQL,⁴⁹ Oracle and SQL server.

Conclusion

It is feasible to transform EN13606-based archetypes in ADL format into CDISC ODM and vice versa. The transformation of element names, descriptions, different languages, datatype, cardinality, optionality, units, value range and terminology codes from ADL to ODM is possible. EN13606 can accommodate a broader range of detailed clinical information and in a more structured and hierarchical manner with more constraints compared to ODM. Thus, in transformation of data from EN13606 into ODM, the richness of metadata in terms of hierarchy and context gets lost. Nevertheless, ODM is mostly used in the context of clinical trials in which the hierarchy of data is generally not requested. Practically, this mapping model could be used to transform EHR data available in EN13606 form to ODM (the preferred standard for clinical research) to integrate with research databases. With

the aim of more efficient and meaningful interoperability between EHRs and clinical research, the paper also suggests the possibility to use EN13606 standards in place of CDISC ODM to build the study protocols and to extract EHR data into the research database, which could avoid duplication of data and transformations.

Acknowledgements: We thank the Institute of Medical Informatics, University of Münster, Germany, for allowing us to use the Medical Data Models portal.

Contributorship: Not applicable.

Declaration of Conflicting Interests: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical approval: Not applicable.

Funding: Not applicable.

Guarantor: Archana Tapuria.

Peer review: This manuscript was reviewed by three individuals who have chosen to remain anonymous.

References

1. Weng C, Appelbaum P, Hripcsak G, et al. Using EHRs to integrate research with patient care: promises and challenges. *J Am Med Inform Assoc* 2012; 19(5): 684–687.
2. Jensen PB, Jensen LJ and Brunak S. Mining electronic health records: towards better research applications and clinical care. *Nat Rev Genet* 2012; 13(6): 395–405.
3. CDISC ODM Certified Products. <https://www.cdisc.org/resources/odm> (accessed 10 April 2018).
4. Warzel DB, Andonaydis C, McCurry B, et al. Common data element (CDE) management and deployment in clinical trials. *AMIA Annu Symp Proc* 2003: 1048.
5. Bruland P, Breil B, Fritz F, et al. Interoperability in clinical research: from metadata registries to semantically annotated CDISC ODM. *Stud Health Technol Inform* 2012; 180: 564–568.
6. Bruland P and Dugas M. *Transformations between CDISC ODM and openEHR Archetypes*. InMIE, 2014, Istanbul, Turkey: IOS Press, p. 1225.
7. Hume S, Aerts J, Sarnikar S, et al. Current applications and future directions for the CDISC Operational Data Model standard: A methodological review. *J Biomed Informat* 2016; 60: 352–362.
8. Richesson RL and Nadkarni P. Data standards for clinical research data collection forms: current status and challenges. *J Am Med Inform Asso* 2011; 18(3): 341–346. doi: 10.1136/amiajnl-2011-000107.
9. Kush R, Alschuler L, Ruggeri R, et al. Implementing Single Source: the STARBRITE proof-of-concept study. *J Am Med Inform Assoc* 2007; 14(5): 662–673.
10. Delaney BC, Curcin V, Andreasson A, et al. Translational medicine and patient safety in Europe: TRANSFoRm—architecture for the learning health system in Europe. *BioMed Res Int* 2015; 2015: 1–8.

11. Lim SN, Keung C, Zhao L, et al. TRANSFoRm: implementing a learning healthcare system in Europe through embedding clinical research into clinical practice. In: *System Sciences (HICSS)*, 48th Hawaii International Conference, 5 January 2015, pp. 3176–3185. Koloa, Kauai, Hawaii, US: IEEE.
12. Mastellos N, Blizniuk G, Czopnik D, et al. Feasibility and acceptability of TRANSFoRm to improve clinical trial recruitment in primary care. *Fam Pract* 2016; 33: 186–191.
13. Lim Choi Keung SN, Zhao L, Rossiter J, et al. *Detailed clinical modelling approach to data extraction from heterogeneous data sources for clinical research*. 2014 AMIA Clinical Research Informatics Summit, San Francisco, April 2014.
14. Ethier J-F, Curcin V, Barton A, et al. Clinical data integration model core interoperability ontology for research using primary care data. *Methods Inf Med* 2015; 54: 16–23.
15. Muñoz P, Trigo J, Martínez I, et al. The ISO/EN 13606 Standard for the Interoperable Exchange of Electronic Health Records. *J Healthc Eng* 2011; 2(1): 1–24.
16. Kay S and Marley T (eds). *EHCR communications: Part 1 Electronic healthcare record architecture*. ENV 13606. Brussels: CEN, 1999.
17. Saganowski S, Misiąszek A, Bródka P, et al. TRANSFoRm eHealth solution for quality of life monitoring. *AMIA Jt Summits Transl Sci Proc* 2016; 2016: 231–239.
18. Daniel C, Ouagne D, Sadou E, et al. Cross border semantic interoperability for clinical research: the EHR4CR semantic resources and services. *AMIA Jt Summits Transl Sci Proc* 2016; 2016: 51–59.
19. Bruland P, McGilchrist M, Zapletal E, et al. Common data elements for secondary use of electronic health record data for clinical trial execution and serious adverse event reporting. *BMC Med Res Method* 2016; 16: 159. doi:10.1186/s12874-016-0259-3.
20. De Moor G, Sundgren M, Kalra D, et al. Using electronic health records for clinical research: the case of the EHR4CR project. *J Biomed Informat* 2015; 53: 162–173.
21. AIM Hy Study. <http://www.aimhy.org.uk> (accessed 10 April 2018).
22. CDISC Operational Data Model (ODM). <https://www.cdisc.org/standards/foundational/odm> (accessed 1 April 2018).
23. Health informatics – Electronic health record communication – Part 1: Reference model (ISO 13606-1:2008). <https://www.iso.org/standard/40784.html> (accessed 1 April 2018).
24. ISO. BS ISO EN 13606 part 1 (No. EN 13606) (eds D Kalra & D Lloyd) 2007; 1: 1–103.
25. Austin T, Lim YS, Nguyen D, et al. Design of an electronic healthcare record server based on Part 1 of ISO EN 13606. *J Healthc Eng* 2011; 2: 143–160. <http://multiscience.metapress.com/content/121507>.
26. ISO. BS ISO EN 13606 part 3 (No. EN 13606) (eds D Kalra & D Lloyd) 2008; 3: 1–47.
27. ISO EN 13606 Standard. <http://www.en13606.org/the-ceniso-en13606-standard> (accessed 1 April 2018).
28. CDISC Clinical Data Acquisition Standards Harmonization (CDASH). <https://www.cdisc.org/standards/foundational/cdash> (accessed 5 April 2018).
29. Hume S, Aerts J, Sarnikar S, et al. Current applications and future directions for the CDISC Operational Data Model standard: A methodological review. *J Biomed Inform* 2016; 60: 352–362.
30. Kuchinke W, Aerts J, Semler SC, et al. CDISC standard-based electronic archiving of clinical trials. *Methods Inf Med* 2009; 48(5): 408–413.
31. El Emam K, Jonker E, Sampson M, et al. (Eysenbach G, ed.) The use of electronic data capture tools in clinical trials: web-survey of 259 Canadian trials. *J Med Internet Res* 2009; 11(1): e8. doi:10.2196/jmir.1120.
32. Electronic Health Records for Clinical Research. <http://www.ehr4cr.eu> (accessed 05 April 2018).
33. i-HD. <http://www.i-hd.eu/index.cfm/resources/ec-projects-results/transform/> (accessed 05 April 2018).
34. XML4Pharma ODM Study Designer. http://www.xml4pharma.com/CDISC_Products/ODMDesigner_2016_ODM_details.html (accessed 01 April 2018).
35. Austin T, Sun S, Lea N, et al. Patterns: a simple but expressive data modelling formalism. *Int J Knowledge Eng Data Mining* 2016; 4(1): 74–92.
36. Bruland P and Dugas M. Transformations between CDISC ODM and openEHR Archetypes. *Stud Health Technol Informat* 2014; 205: 1225.
37. Dugas M, Neuhaus P, Meidt A, et al. Portal of medical data models: information infrastructure for medical research and healthcare. *Database (Oxford)* 2016; 2016: bav121. doi:10.1093/database/bav121.
38. Bruland P, Forster C, Breil B, et al. Does single-source create an added value? Evaluating the impact of introducing x4T into the clinical routine on workflow modifications, data quality and cost–benefit. *Int J Med Inform* 2014; 83(12): 915–928.
39. Huser V, Sastry C, Breymaier M, et al. Standardizing data exchange for clinical research protocols and case report forms: An assessment of the suitability of the Clinical Data Interchange Standards Consortium (CDISC) Operational Data Model (ODM). *J Biomed Informat* 2015; 57: 88–99.
40. CDISC. http://portal.cdisc.org/CDISC%20User%20Networks/Europe/Italian%20Language/Presentations/ItalianCDISC_UGMeeting01/9_ODM_PVerplancke.pdf (accessed 01 April 2018).
41. Leroux H, Metke-Jimenez A and Lawley MJ. Towards achieving semantic interoperability of clinical study data with FHIR. *J Biomed Semantics* 2017; 8(1): 41. doi: 10.1186/s13326-017-0148-7.
42. Doods J, Neuhaus P and Dugas M. Converting ODM Metadata to FHIR Questionnaire Resources. *Stud Health Technol Informat* 2016; 228: 456.
43. van der Lei J. Use and abuse of computer-stored medical records. *Methods Inf Med* 1991; 30: 79–80.
44. Weiskopf NG and Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. *J Am Med Inform Assoc* 2013; 20: 144–151.

45. Breil B, Kenneweg J, Fritz F, et al. Multilingual medical data models in ODM Format: a novel form-based approach to semantic interoperability between routine healthcare and clinical research. *Appl Clin Inform* 2012; 3(3): 276–289.
46. Hegselmann S, Gessner S, Neuhaus P, et al. Automatic conversion of metadata from the study of health in Pomerania to ODM. *Stud Health Technol Inform* 2017; 236: 88–96.
47. Storck M, Krumm R and Dugas M. ODM Summary: a tool for automatic structured comparison of multiple medical forms based on semantic annotation with the Unified Medical Language System. *PLoS One* 2016; 11(10): e0164569.
48. Leroux H and Lefort L. Semantic enrichment of longitudinal clinical study data using the CDISC standards and the semantic statistics vocabularies. *J Biomed Semantics* 2015; 6: 16.
49. Austin T, Sun S, Lim Y, et al. An electronic healthcare record server implemented in PostgreSQL. *J Healthc Eng* 2015; 6(3): 325–344.