Document downloaded from:

http://hdl.handle.net/10251/65033

This paper must be cited as:

Moreno-Conde, A.; Moner Cano, D.; Da Cruz, WD.; Santos, MR.; Maldonado Segura, JA.; Robles Viejo, M.; Kalra, D. (2015). Clinical information modeling processes for semantic interoperability of electronic health records: systematic review and inductive analysis. Journal of the American Medical Informatics Association. 22(4):925-934. doi:10.1093/jamia/ocv008.

The final publication is available at

http://dx.doi.org/10.1093/jamia/ocv008

Copyright Oxford University Press

Additional Information

This is a pre-copyedited, author-produced PDF of an article accepted for publication in Journal of the American Medical Informatics Association following peer review. The version of record is available online at: http://dx.doi.org/10.1093/jamia/ocv008

Clinical Information Modeling Processes for Semantic Interoperability of Electronic Health Records: Systematic Review and Inductive Analysis

Alberto Moreno-Conde[1,2,+], David Moner[3,+], Wellington Dimas da Cruz[4], Marcelo R Santos[5], José Alberto Maldonado[3,6], Montserrat Robles[3], Dipak Kalra[1]

+ These authors have contributed equally to this work

[1] Centre for Health Informatics and Multiprofessional Education, University College London, London, United Kingdom
[2] Hospital Universitario Virgen del Rocío, Sevilla, Spain
[3] Instituto de Aplicaciones de las Tecnologías de la Información y de las Comunicaciones Avanzadas, Universitat Politècnica de València, Valencia, Spain
[4] Universidade FUMEC, Belo Horizonte, Brazil
[5] Centro de Informática em Saúde (CINS), Universidade Federal de Minas Gerais, Belo Horizonte, Brazil
[6] VeraTech for Health SL, Valencia, Spain

Correspondence to: Alberto Moreno Conde, Grupo de Innovación Tecnológica, Hospital Universitario Virgen del Rocío, Edif. Centro de Documentación Clínica Avanzada, Av. Manuel Siurot, s/n. Seville, 41013, Spain. Email: albertomorenoconde@gmail.com

Keywords: Electronic Health Record, Semantic Interoperability, Clinical Information Modeling, Systematic Review, Clinical Information Model

Word count (excluding title page, abstract, references, figures, and tables): 4000

ABSTRACT

OBJECTIVE. This systematic review aims to identify and compare the existing processes and methodologies that have been published in the literature for defining clinical information models that support the semantic interoperability of Electronic Health Record systems.

MATERIAL AND METHODS. Following the PRISMA systematic review methodology, the authors reviewed published papers between 2000 and 2013 about semantic interoperability of Electronic Health Records contained in PubMed, IEEE Xplore and Science Direct. Additionally, an inductive content analysis was done to the selected papers to summarize the steps and methodologies followed in order to build clinical information models.

RESULTS. 378 articles were screened and 36 papers were selected for full review. They were analyzed to extract relevant information for the analysis and characterized according to the steps the authors had followed for clinical information modeling.

DISCUSSION. Most of the reviewed papers lack a detailed description of the modeling methodologies used to create clinical information models. A representative example is the lack of description related to the definition of terminology bindings and the publication of the generated models. However, this systematic review confirms that most of the clinical information modeling activities follow very similar steps for the definition of clinical information models. Having a robust and shared methodology could improve their correctness, reliability and quality.

CONCLUSION. Independently of implementation technologies and standards, it is possible to find common patterns in the development of clinical information models, suggesting the viability of defining a unified good practice methodology to be used by any clinical information modeler.

BACKGROUND AND SIGNIFICANCE

The increased adoption of Electronic Health Record (EHR) systems potentially enables the sharing of patient information across multiple systems to support continuity of care. To this end, standards and technical specifications have been developed; defining how the information contained in EHRs should be structured, semantically described and communicated. Current trends followed by most of those specifications rely on differentiating the representation of data instances from the definition of the clinical information models.

Clinical Information models

In this paper we use the expression clinical information model (CIM) as a generic term that encompasses all technical specifications defining how clinical information is organized and described inside an EHR system, repository or for EHR communication. A CIM defines both the information structure and formal semantics of documented clinical concepts. CIMs are structural and semantic artifacts that facilitate organizing, storing, querying and displaying clinical data, exchanging that data between different information systems, and performing data analytics. Usually, a CIM is defined by constraining the generic data structures of an underlying reference model that provides the basic characteristics and attributes needed to represent data instances. Terminologies such as SNOMED CT, ICD or LOINC also play an important role in the definition of CIMs. The structure of CIMs can be bound (precisely mapped) to clinical terminologies to provide an univocal definition of the model. Furthermore, terminologies are also used to specify value sets, i.e. the set of possible terms that can be assigned as values of the clinical information. Thus, a complete semantically interoperable definition of CIMs can only be

achieved by both using a standard reference model and using terminologies to describe the semantics of the information structures.

Goossen et al. described initiatives that follow a CIM approach, indicating their differences and similarities [1].

The HL7 v3 modeling approach is based on a standard Reference Information Model (RIM) representing the main business logic of any healthcare environment, from which specific messages and documents can be defined. HL7 v3 messages [2] and HL7 Clinical Document Architecture (HL7 CDA) [3] are standards based on the HL7 RIM. It is possible to define clinical information models for HL7 CDA in the form of HL7 templates that specify how the clinical information is to be contained and organized within each kind of document, for specific clinical communication purposes.

HL7 FHIR [4] is a new generation specification that uses modular components called Resources. These resources (definitions of common reusable patterns of clinical information) can be combined or extended in order to provide particular solutions to health information systems. They are therefore also to some extent CIMs.

Another important modeling approach is based on the dual level methodology [5], based on the definition of a synthesized and generic Reference Model (RM) that is designed to represent the most basic properties and structures of any EHR. Clinical information models are defined in the form of archetypes. Archetypes define how data should be structured in order to be seamlessly stored or transferred between EHR systems. The dual model approach is supported by the EN ISO 13606 standard [6] and the openEHR specifications [7].

Additional modeling approaches have emerged focused on defining generic information models at a conceptual level, without depending on a specific implementation. The Clinical

Information Modeling Initiative (CIMI) [8], Detailed Clinical Models (DCM) [9], and the Clinical Element Model initiative [10] are examples of such generic models.

Figure 1 summarizes the reference models used (i.e. the models that represent data instances) and the CIM technology employed by each of these initiatives.

Clinical information modeling processes

We define a clinical information modeling process (CIMP) as the process of analyzing the domain and requirements, designing, implementing, validating and maintaining CIMs. This process will usually require the cooperation of experts with technical and clinical background, in order to obtain a final implementable definition of CIMs that satisfies the clinical needs, which may be agreed and used at the level of a single care organization, an EHR system vendor user group, a health region or a country. Once CIMs are defined, governance mechanisms can be applied to ensure a correct management and future evolution of the defined models.

The traditional software development process includes requirements definition, a domain analysis, design, implementation and validation [11]. CIMP covers domain analysis, design, implementation and validation of the CIMs, but also includes some special characteristics: CIMs are based on standard specifications and formats, and they can be shared and reused. The participation of both health and technical professionals in this process requires coordination and evaluation mechanisms in order to create relationships of trust towards the developed CIMs. Moreover, having a well-defined CIMP is of extreme importance in order to ensure a comparable quality and homogeneous design of CIMs created by different organizations or professionals [12,13].

OBJECTIVE

This paper aims to identify and compare the existing clinical information modeling processes and methodologies that have been published in the literature. In particular, a systematic review and an inductive content analysis have been performed in order to learn about methodologies and experiences in building CIMs for semantically interoperable EHR systems. The question being addressed in this study is to discover if an emergent consensus (good practice) strategy in building CIM artifacts exists; and to know if it is therefore possible to propose a common or unified CIMP.

MATERIALS AND METHODS

Systematic review

In order to perform the systematic review of the existing literature we have chosen the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology [14]. This methodology proposes a 27 item checklist and a flow diagram in order to guide the authors during the conduct of a systematic review [15].

The eligibility criteria, i.e. the characteristics to be taken into account to perform the search, were:

- Papers with any of the following terms in their title or abstract: “Electronic Health Record”, “Hospital Information System”, “Clinical Information System”, “Health Information System”, “EHR”, “medical record system”, “automated medical systems”, “Electronic Medical Record”.
- Papers with the terms “semantic interoperability” or “clinical information model” in their title or abstract.
- Published between January 2000 and August 2013.

When deciding the search criteria, it was preferred to have a broad scope focused on

semantic interoperability for EHR rather than searching for each of the specific EHR mechanisms that could be applied for clinical information modeling such as “archetype” or “template”. The variability of terms and technologies related to the definition of CIMs is so broad that we needed to avoid the risk of leaving out important references or experiences that used formalisms such as object-oriented models, entity-relationship design, XML Schemas or ontologies. The inclusion of the semantic interoperability filter helped in excluding EHR traditional development experiences that did not have a focus on the reuse of the information structures that were developed.

The sources of information where the search was performed were PubMed [16], IEEE Xplore [17] and ScienceDirect [18]. As an example, Figure 2 shows the search defined in PubMed, according to the previously described search criteria. Search queries in the other sources of information can be found in the supplementary material.

According to the PRISMA methodology, a two-phase procedure was established for the systematic review. In Phase 1 (study screening) a first review was made based on the title and abstract of the papers returned as result of the queries. Two additional exclusion criteria were adopted: (a) the paper does not include information about CIMs, and (b) the paper does not include information about CIMP. In case of doubt due to the limited information available in the titles and abstracts, the papers were accepted for full review. In Phase 2 (full review) the full text of the selected papers was reviewed. The objective of this full review was twofold: to reject those papers that did not fit the purpose of the systematic review and, only from those that were finally accepted, to extract a set of data items and indicators to perform further analysis.

Inductive analysis

In addition to the systematic review, a methodology called Inductive Content Analysis

was applied [19] to discover the CIMP steps described in the selected papers.. This methodology is recommended to avoid creating preconceived categories when the existing literature is limited or heterogeneous. According to this methodology, a set of tags that qualify the CIM definition processes described in the papers were iteratively refined to represent an abstraction of CIMP steps. The information about the modeling processes was organized into categories, in order to provide a high level and summarized description of those steps.

RESULTS

As a result of the search 374 papers were found, 18 of which were duplicated. Additionally, the authors identified four additional references that met the search criteria and were relevant to the review, but not indexed by the search engines [1,20–22]. In total, 360 paper titles and abstracts were screened by the authors, and 53 of them were accepted for a full-text review, where it was discovered that only 36 papers contained relevant data for the objectives of this research. The summary of this review process is presented in Figure 3.

The main reasons for exclusion were that the analyzed papers did not contain information about modeling or clinical information models. In three cases the full text of the articles was not available.

Table 1 shows the annual distribution of the selected papers. Note that the search in 2013 only included the period between January and August.

Table 1. Annual distribution of papers

Year	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**
No. of papers	1	0	0	3	4	6	4	4	8	6

Analysis of the indicators collected from the selected papers

Table 2 details a summary of the information collected in the paper review. The complete list of publications and information collected and can be found in the supplementary material.

Table 2. Indicators of the analyzed papers

Indicator	Values	References	Total
Type of CIM	HL7 templates	[23], [24], [25], [26], [27], [28], [29], [30], [31]	9 (25.0%)
	EN ISO 13606 or openEHR archetypes	[32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [20], [21]	16 (44.4%)
	Other	[46], [47], [48], [49], [50], [51], [52], [53], [1], [22], [54]	11 (30.6%)
Reference Model	HL7 v3 / HL7 CDA	[30], [25], [28], [24], [26], [29], [36], [23], [31]	9 (25.0%)
	openEHR	[39], [42], [45], [33], [32], [44], [41], [38], [40]	9 (25.0%)
	EN ISO 13606	[20], [21], [35], [43], [34], [37]	6 (16.7%)
	Other	[1], [22], [52], [46], [47], [48], [49], [50], [53], [27], [51], [54]	12 (36.1%)
CIMP is described	Yes	[23], [32], [33], [34], [35], [46], [26], [47], [48], [49], [29], [41], [43], [45], [30], [20], [21], [31], [54]	19 (52.8%)
	No	[24], [36], [37], [38], [25], [27], [48], [40], [28], [50], [51], [52], [53], [42], [44], [1], [22]	17 (47.2%)
Are existing CIMs reused?	Yes	[1], [20], [21], [22], [35], [52], [30], [39], [25], [53], [42], [43], [45], [34], [33], [41], [38], [37], [26], [27], [40], [51], [36], [23], [31], [54]	26 (72.2%)
	No	[46], [47], [28], [49], [29], [32], [24],	8

		[44]	(22.2%)
	Not specified	[48], [50]	2 (5.6%)
Are resulting CIMs shared?	Yes	[30], [35], [36], [37], [38], [41], [45], [53]	8 (22.2%)
	Planned	[1], [20], [21], [22], [39], [40]	6 (16.7%)
	Not specified	[23], [32], [33], [34], [46], [24], [25], [27], [26], [42], [47], [48], [28], [49], [50], [51], [52], [29], [43], [44], [31], [54]	22 (61.2%)
Terminologies used	SNOMED CT	[52], [25], [50], [53], [33], [24], [41], [38]	8 (22.2%)
	Other	[35], [30], [46], [47], [48], [28], [49], [29], [43], [34], [32], [51], [23], [31]	14 (38.9%)
	Not specified	[1], [20], [21], [22], [39], [42], [45], [44], [37], [26], [27], [40], [36], [54]	14 (36.1%)
Tools used	Archetype editor	[32], [33], [35], [36], [37], [38], [39], [40], [53], [41], [42], [43], [44], [45], [20]	15 (41.6%)
	UML/Visual modeler editor	[46], [27], [26], [29], [30]	5 (13.9%)
	Protégé	[47], [51]	2 (5.6%)
	Other	[25], [28]	2 (5.6%)
	Not specified	[23], [34], [24], [48], [49], [50], [52], [21], [1], [22], [31], [54]	12 (33.3%)
Application domains	Theoretic application	[1], [21], [22], [37], [27]	5 (13.9%)
	Regional/national projects	[35], [25], [27], [31], [54]	5 (13.9%)
	One healthcare department	[20], [52], [30], [32], [48], [28], [29], [42], [43], [45], [49], [33], [24], [44], [41], [38], [36], [53]	18 (50%)
	Multiple healthcare	[34], [51], [23], [40], [39], [46] [47],	8

	departments	[50]	(22.2%)
Implementation in real environment	Yes	[1], [20], [35], [52], [30], [46], [39], [47], [25], [48], [28], [49], [50], [53], [42], [43], [45], [34], [33], [32], [24], [44], [41], [38], [37], [26], [27], [40], [31], [54]	30 (83.3%)
	No	[13], [21], [29], [51]	4 (13.9%)
	Not specified	[36], [23]	2 (5.6%)
Participation of health professionals	Yes	[1], [20], [21], [35], [52], [30], [39], [47], [28], [49], [50], [53], [29], [42], [43], [45], [34], [32], [24], [44], [41], [38], [37], [51], [31], [54]	26 (73.2%)
	Not specified	[22], [46], [25], [48], [33], [26], [27], [40], [36], [23]	10 (27.8%)

50% of the selected papers were focused on one specialized care department, while the others were focused on multiple departments, national/regional projects or described a theoretical approach. The papers cover a large variety of clinical domains, including nursing, oncology, neonatology, genetics or infectious diseases. Most of the papers (83.3%) described a real deployment. 73.2% of the papers also mentioned the participation of health professionals during the development process.

The preferred type of technical artefacts used to implement CIMs were archetypes (44.4%) followed by HL7 templates (25.0%). With regard to the reference models used for the definition of CIMs, openEHR (25.0 HL7 v3 (25.0% including messages and CDA), and EN ISO 13606 (16.7%) were the most mentioned. Other works made use of proprietary reference models, expressed in UML, XML or as ontologies.

All the references included in this systematic review apply a CIMP for defining CIMs, but only 52.8% of them described it with some degree of detail.

In most of the studied papers, modeling of CIMs was centered on the structural definition (e.g. a hierarchy of fields and grouping headings) without detailing how these structures were bound to terminologies (i.e. without mapping the field names to terms, nor specifying terminology value lists for fields with textual values). 36.1% of analyzed papers did not include any mention to the use of terminologies. In the others, SNOMED CT was the most widely adopted terminology (22.2%). Only four of the papers [21,24,35,41] provided a detailed description about how they conducted the terminology binding process. The same lack of information can be found about the metadata associated with the CIMs created (provenance, authorship, endorsements, related bibliography, etc.), which was rarely mentioned.

Sharing publicly the defined CIMs at the end of the CIMP is a mechanism to provide credibility and acceptance of developed artifacts, and to facilitate their reuse. Only 38.9% of papers mentioned sharing the defined CIMs publicly. 72.2% of papers mentioned reusing existing CIMs as part of their development process.

A recurring demand nowadays in healthcare is to use and produce specific tools and processes to solve problems related to electronic recording of clinical data [35]. The use of appropriate design tools helps users to manage the complexity of a detailed specification and helps to ensure the syntactical correctness of the resulting model. Tool use should therefore contribute to the quality of the CIMs. In this context, we found that 67.7% of publications mentioned the use of specific tools for the creation of CIMs. Archetype editors had the leading adoption (41.6%), followed by UML or similar visual design tools (13.9%). The other papers mentioned the use of tools such as spreadsheets, mind maps, XML editors or Protégé.

Inductive analysis of clinical information modeling processes steps

After the tagging and categorization of the information extracted about the CIMP

described in the selected papers, the following non-mutually exclusive categories of related information were found. Table 3 summarizes the papers including information related to each category.

- *Scope definition leading to selection of the domain and selecting relevant experts.* Whether the scope of a CIM is local or it is designed for wider use, it will be needed to identify the domain to be covered and the expected uses of the CIMs to be developed [39]. Based on the identification of the care setting, healthcare activities, and clinical requirements, it is possible to create a work group of relevant experts in that clinical domain, responsible for the design of the CIMs [30].
- *Analysis of the information covered in the specific domain.* In order to create complete CIMs definitions, it is required to obtain an understanding of clinical scenarios, workflows and users, to determine the data items that will be supported by CIMs [25–27,44,46][R41, R18]. It is necessary to identify how the existing systems have been implemented and documented [48,51]. Reviewing guidelines, literature and validated clinical scales [30,32,53] allows the design team to ensure that information covered by the CIMs will meet the requirements of clinical practice. To collect this information, interviews and workshops with clinical experts may be performed [35,36].
- *Design of clinical information models.* After identifying the necessary data items, these are merged and harmonized into CIMs avoiding possible overlapping [20,45,47]. Each CIM will detail the possible set of attributes associated with it in a structured way [1,21,41]. Each data item associated with a clinical concept can be detailed in the form of a value set or CIM node [24,38,50,52]. It is also important to identify domain terminologies that are applicable to the studied domain, in order to map them to the CIMs

[28,35]. The definition of CIMs can be focused either on just determining the essential data sets as common minimum communication requirements [32] or on satisfying the application of CIMs for multiple purposes, ensuring a basic compatibility across domains.

- *Definition of implementable clinical information model specifications.* In order to make the defined CIMs compatible with existing EHR information standards an implementable technical specification is needed. The process of implementing the modeled CIMs into technical artifacts starts with the search and review of existing CIMs [23,37,42]. Those CIMs that suit the scope of the project will be reused or adapted [33]. This will increase the interoperability between systems with different local needs but using similar CIMs. For those clinical concepts that are not covered by existing CIMs, new ones will be created.
- *Validation.* Multiple techniques have been adopted to validate the defined models, including peer review and the creation of prototype screens [29,34,49]. Further evaluation using routinely collected clinical data from multiple patients will provide stronger validation for the defined CIMs [43].
- *Publishing and maintenance.* Those CIMs that are created should be transferred into a public repository in order to be accessible by any other user [40]. CIMs published in the repository should include a method for receiving feedback from those projects and organizations that adopt them [22].
- *Governance.* This final category is not properly a step of the CIMP, but closely related to it [40]. The organization responsible for the development and maintenance of CIMs will be in charge of establishing an effective governance of them. This governance will

determine the process for quality review and publication of CIMs, and the relationship with other projects and organizations working in the same domain covered by those CIMs [1]. This could include certification of CIMs by the developer organization or other external bodies [22].

Table 3. Categories found after the inductive analysis of CIMP steps

Category	Published papers
1. Scope definition and creation of a work team	[52], [30], [46], [39]
2. Analysis of the information covered in the specific domain	[1], [20], [21], [52], [30], [46], [39], [36], [25], [47], [48], [28], [49], [50], [53], [29], [42], [43], [45], [34], [33], [32], [35], [24], [51], [44], [31], [54]
3. Design of clinical information models	[1], [20], [21], [52], [39], [47], [28], [49], [50], [53], [43], [45], [35], [24], [41], [38], [31], [54]
4. Definition of implementable clinical information models specifications	[1], [20], [21], [30], [36], [25], [47], [48], [49], [50], [29], [42], [43], [45], [34], [33], [32], [35], [24], [51], [41], [38], [37], [23], [26], [27], [54]
5. Validation	[1], [20], [30], [39], [49], [53], [29], [43], [34], [51], [41], [54]
6. Publishing and maintenance	[39], [53], [43], [31]

7. Governance	[1], [22]

DISCUSSION

This systematic review analyzed the reported clinical information modeling processes that have been adopted to support the semantic interoperability of EHR systems. Our reflection on the results of the publication searches confirmed that the decision not to include more specific search criteria was appropriate. Using a generic search without including specific terms for the types of CIM proved to be successful, since it allowed the inclusion of an extensive range of experiences of CIMs development, using different technologies and standards.

Discussion on the extracted indicators

The extracted indicators from the selected papers raise several interesting discussion points.

- *Limited information about the CIMP used to create clinical information models.* All the selected papers rely on the use of CIMs as a kernel piece of their information systems. However, the methodology followed to create them was not usually described in detail and sometimes not even mentioned. This lack of information might reduce the level of third party trust in the quality of the developed CIMs. Given that currently a standard CIMP does not exist in the literature, we had expected that more authors would have included a detailed description of their own modeling and validation steps. It is particularly relevant the limited information about terminology bindings that is provided in the studied papers. 36% of the reviewed papers did not even mention the

terminological aspect, and most of the others only referred to it as a future work. A CIM cannot be semantically interoperable if it lacks terminological references that describe its contents. The definition of a particular information structure can be affected by the expressivity of the selected terminologies that accompany it and, vice-versa, the design of a particular information structure affects how the value sets to be used in it should be created [10]. Moreover, a loose definition and use of terminological value sets also affects the final quality and interoperability of the clinical data that is produced [55].

- *Resultant CIMs are not shared.* It was observed that most of the analyzed experiences didn't provide any mechanism to access the resultant CIMs. Although it is not mandatory to share them with external groups, it would be a good practice to share these models openly unless there are copyright restrictions. This can improve the quality of the defined models through feedback [12,56] and supports the harmonization of multiple groups developing CIMs in parallel in the same domain, and thus, the semantic interoperability of EHR information.
- *Modeling tools.* CIM definition is a multidisciplinary task where health professionals and technicians collaborate. To this end, it is important to have the appropriate tools that ease the definition and review processes. This study suggests that most of the modeling efforts use generic tools to carry out this work, such as UML technologies, mind maps, spreadsheets or XML tools. Only those which rely on the archetype approach make use of specific tooling. In any case, several of the reviewed papers warn about the immaturity of modeling support tools [33,41,45]. We can conclude from these results that there is a need for better modeling tools. However, it has to be taken into account that the mentioned papers are from 2007, 2009 and 2011. It should be expected that

improvements have been made in this topic over the last few years.

- *Mapping to implementable specifications*. Transforming generic CIM definitions into implementable specifications (i.e. archetypes or templates) is not a direct process since it requires accommodating the information attributes of the CIM in a specific RM structure [34]. This implies that a shared CIM could be implemented in different technical artifacts or standards that were not completely equivalent.

Discussion on the inductive analysis results

The methodological approach to create CIMs has been discovered to be similar in all the studied papers where information was available. Figure 4 summarizes the steps obtained from the inductive analysis of the content related to CIMP and the relationships between them. The process starts with the selection of the scope and the work team, followed by a domain analysis (including the research of references or existing CIMs that could be reused), the design and definition of the structure and semantics of new CIMs (or the modification of existing ones), the validation by health professionals and, finally, the publication of the resulting CIMs.

Although these steps were defined based on the partial information available in the published literature, the level of similarities found suggests that it would be possible to define a unified process to guide CIM definition, including the description of best practices to increase the quality of the CIMs.

Finally, the identified CIMP steps are encompassed by a general governance process. This governance is in charge of identifying when a new CIM needs to be created of if existing ones should be reviewed. The governance of CIMs is a separate topic that has also received attention by researchers [57,58].

Limitations and risk of bias of this systematic review

The authors recognize that the inclusion of the "semantic interoperability" criterion could have limited the papers found in the search, since the use of this term was limited in the early 2000s. Anyhow, this criterion allowed collecting early experiences of CIM-based approaches from promoters of the semantic interoperability concept at that time. Nearly 20 papers published until 2005 were found including that term.

In order to limit the risk of bias of this systematic review, all papers were screened by at least two of the authors of this paper, who had to agree on their suitability for the full-text review phase. In the full review phase, the papers were interchanged between the authors. Thus, every paper was either screened or fully reviewed independently by different authors. In the inductive analysis the authors also achieved a consensus on the steps and classifications of the selected papers, based on the information contained in them.

Regarding the obtained results and conclusions, a limitation of the performed review is that most papers did not describe in detail the CIMP that was followed in order to define the CIMs. In many cases the modeling process was just mentioned or scattered across the text. This necessitated a careful and detailed reading of each of the papers in order to find out the steps followed by the original authors.

CONCLUSION

The use of CIMs has gained recognition as one of the essential aspects of the creation of standardized and interoperable EHR systems. Different standards and technical approaches exist (e.g. EN ISO 13606 and openEHR using archetypes, or HL7 v3 using templates), but the idea of separating the definition of the CIMs from the actual representation and persistence of the data values is shared among all of them. Moreover, the work of international modeling initiatives

such as CIMI indicates an increased interest in creating reusable CIMs. Thus, it is important that the CIMP used to create those models follows clear and well defined steps.

This research characterized published experiences related to the creation of semantically interoperable EHR systems between 2000 and 2013, in order to obtain a better understanding of the steps followed by all of them during the creation of CIMs. It was found that most of the experiences share a similar approach. This suggests that it should be possible to create a common or unified methodology for clinical information modeling in the future. This conclusion is however limited due to the lack of detail describing the used CIMP in the selected papers. It is important to advocate further collaboration between the main organizations and professionals involved in CIM development, to reach a consensus in the definition of a unified best practice CIMP.

A commonly agreed CIMP will promote and emphasize the importance of analyzing the information covered in a particular domain, the collaboration between different clinical and technical professionals and the search for consensus in the definition of CIMs. It will also minimize the diversity of ways in which a CIM can be designed and will make terminology bindings more consistent. This is directly related to the improvement of the quality of CIMs [59,60].

REFERENCES

1 Goossen W, Goossen-Baremans A, van der Zel M. Detailed Clinical Models: A Review. *Healthc Inform Res* 2010;**16**:201. doi:10.4258/hir.2010.16.4.201

2 Beeler GW. HL7 Version 3—An object-oriented methodology for collaborative standards development. *Int J Med Inf* 1998;**48**:151–61. doi:10.1016/S1386-5056(97)00121-4

3 Dolin RH, Alschuler L, Boyer S, *et al.* HL7 Clinical Document Architecture, Release 2. *J Am Med Inform Assoc* 2006;**13**:30–9. doi:10.1197/jamia.M1888

4 Fast Health Interoperability Resources (FHIR). http://www.hl7.org/fhir/ (accessed 18 Jul2014).

5 Beale T. Archetypes: Constraint-based Domain Models for Futureproof Information Systems. *OOPSLA 2002 Workshop Behav Semant* 2002.

6 ISO 13606:2008 - Health informatics - Electronic health record communication. 2008.

7 openEHR. http://www.openehr.org/ (accessed 18 Jul2014).

8 Clinical Information Modeling Initiative (CIMI). http://www.opencimi.org/ (accessed 18 Jul2014).

9 Goossen WT. Using detailed clinical models to bridge the gap between clinicians and HIT. *Stud Health Technol Inf* 2008;**141**:3–10.

10 Oniki TA, Coyle JF, Parker CG, *et al.* Lessons learned in detailed clinical modeling at Intermountain Healthcare. *J Am Med Inform Assoc* 2014;:amiajnl – 2014–002875. doi:10.1136/amiajnl-2014-002875

11 Jacobson I, Booch G, Rumbaugh J, *et al. The unified software development process*. Addison-Wesley Reading 1999.

12 Ahn S, Huff SM, Kim Y, *et al.* Quality metrics for detailed clinical models. *Int J Med Inf* 2013;**82**:408–17.

13 Kalra D. Editorial principles for the development of standards for the structure and content of health records. 2012.https://www.rcplondon.ac.uk/sites/default/files/documents/editorial-principles-for-the-development-of-record-standards.pdf (accessed 16 Jul2014).

14 Moher D, Liberati A, Tetzlaff J, *et al.* Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;**339**:b2535–b2535. doi:10.1136/bmj.b2535

15 Transparent reporting of systematic reviews and meta-analyses (PRISMA). http://www.prisma-statement.org/ (accessed 18 Jul2014).

16 US National Library of Medicine. http://www.pubmed.gov (accessed 18 Jul2014).

17 IEEE Xplore Digital Library. http://ieeexplore.ieee.org/ (accessed 18 Jul2014).

18 ScienceDirect. http://www.sciencedirect.com/ (accessed 18 Jul2014).

19 Elo S, Kyngäs H. The qualitative content analysis process. *J Adv Nurs* 2008;**62**:107–15. doi:10.1111/j.1365-2648.2007.04569.x

20 Rinner C, Kohler M, Hübner-Bloder G, *et al.* Creating ISO/EN 13606 archetypes based on clinical information needs. In: *Proceedings of EFMI Special Topic Conference" e-Health Across Borders Without Boundaries*. 2011. 14–5.

21 Muñoz Carrero A, Romero Gutiérrez A, Marco Cuenca G, *et al. Manual práctico de interoperabilidad semántica para entornos sanitarios basada en arquetipos*. Unidad de

investigación en Telemedicina y e-Salud. Instituto de Salud Carlos III - Ministerio de Economía y Competitividad. 2013.

22 Kalra D. Editorial principles for the development of standards for the structure and content of health records. 2012.https://www.rcplondon.ac.uk/sites/default/files/documents/editorial-principles-for-the-development-of-record-standards.pdf

23 Yuksel M, Dogac A. Interoperability of Medical Device Information and the Clinical Applications: An HL7 RMIM based on the ISO/IEEE 11073 DIM. *IEEE Trans Inf Technol Biomed* 2011;**15**:557–66. doi:10.1109/TITB.2011.2151868

24 Nagy M, Hanzlicek P, Precková P, *et al.* Semantic interoperability in Czech healthcare environment supported by HL7 version 3. *Methods Inf Med* 2010;**49**:186.

25 Lopez DM, Blobel BG. A development framework for semantically interoperable health information systems. *Int J Med Inf* 2009;**78**:83–103.

26 Lopez DM, Blobel B. Enhanced semantic interoperability by profiling health informatics standards. *Methods Inf Med* 2009;**48**:170–7.

27 Lopez DM, Blobel B. Enhanced semantic interpretability by healthcare standards profiling. *Stud Health Technol Inform* 2008;**136**:735.

28 Knaup P, Garde S, Haux R. Systematic planning of patient records for cooperative care and multicenter research. *Int J Med Inf* 2007;**76**:109–17.

29 Goossen WT, Ozbolt JG, Coenen A, *et al.* Development of a provisional domain model for the nursing process for use within the Health Level 7 reference information model. *J Am Med Inform Assoc* 2004;**11**:186–94.

30 Anderson HV, Weintraub WS, Radford MJ, *et al.* Standardized Cardiovascular Data for Clinical Research, Registries, and Patient CareA Report From the Data Standards Workgroup of the National Cardiovascular Research Infrastructure Project. *J Am Coll Cardiol* 2013;**61**:1835–46. doi:10.1016/j.jacc.2012.12.047

31 Jian W-S, Hsu C-Y, Hao T-H, *et al.* Building a portable data and information interoperability infrastructure—framework for a standard Taiwan Electronic Medical Record Template. *Comput Methods Programs Biomed* 2007;**88**:102–11. doi:10.1016/j.cmpb.2007.07.014

32 Spigolon DN, Moro CMC. Essential data set's archetypes for nursing care of endometriosis patients. *Rev Gaúcha Enferm* 2012;**33**:22–32. doi:10.1590/S1983-14472012000400003

33 Späth MB, Grimson J. Applying the archetype approach to the database of a biobank information management system. *Int J Med Inf* 2011;**80**:205–26.

34 Smith K, Kalra D. Electronic health records in complementary and alternative medicine. *Int J Med Inf* 2008;**77**:576–88.

35 Santos MR, Bax MP, Kalra D. Dealing with the Archetypes Development Process for a Regional EHR System. *Appl Clin Inform* 2012;**3**:258–75. doi:10.4338/ACI-2011-12-RA-0074

36 Moner D, Moreno A, Maldonado JA, *et al.* Using archetypes for defining CDA templates. *Stud Health Technol Inform* 2012;**180**:53–7.

37 Moner D, Maldonado JA, Boscá D, *et al.* CEN EN13606 normalisation framework implementation experiences. In: *Seamless Care, Safe Care: The Challenges of Interoperability and Patient Safety in Health Care: Proceedings of the EFMI Special Topic Conference, June 2-4, 2010, Reykjavik, Iceland*. IOS Press 2010. 136.

38 Marcos M, Maldonado JA, Martínez-Salvador B, *et al.* Interoperability of clinical decision-support systems and electronic health records using archetypes: a case study in clinical trial eligibility. *J Biomed Inform* 2013;**46**:676–89.

39 Leslie H. International developments in openEHR archetypes and templates. *Health Inf Manag J* 2008;**37**:38.

40 Kohl CD, Garde S, Knaup P. Facilitating secondary use of medical data by using openEHR archetypes. *Stud Health Technol Inform* 2009;**160**:1117–21.

41 Garde S, Hovenga E, Buck J, *et al.* Expressing clinical data sets with openEHR archetypes: A solid basis for ubiquitous computing. *Int J Med Inf* 2007;**76, Supplement 3**:S334–41. doi:10.1016/j.ijmedinf.2007.02.004

42 Garcia D, Moro CM, Cicogna PE, *et al.* Method to Integrate Clinical Guidelines into the Electronic Health Record (EHR) by Applying the Archetypes Approach. *Stud Health Technol Inform* 2012;**192**:871–5.

43 Duftschmid G, Rinner C, Kohler M, *et al.* The EHR-ARCHE project: Satisfying clinical information needs in a Shared Electronic Health Record System based on IHE XDS and Archetypes. *Int J Med Inf* 2013;**82**:1195–207.

44 Dias RD, Cook TW, Freire SM. Modeling healthcare authorization and claim submissions using the openEHR dual-model approach. *BMC Med Inform Decis Mak* 2011;**11**:60. doi:10.1186/1472-6947-11-60

45 Buck J, Garde S, Kohl CD, *et al.* Towards a comprehensive electronic patient record to support an innovative individual care concept for premature infants using the openEHR approach. *Int J Med Inf* 2009;**78**:521–31. doi:10.1016/j.ijmedinf.2009.03.001

46 Puentes J, Roux M, Montagner J, *et al.* Development framework for a patient-centered record. *Comput Methods Programs Biomed* 2012;**108**:1036–51.

47 Liu D, Wang X, Pan F, *et al.* Harmonization of health data at national level: a pilot study in China. *Int J Med Inf* 2010;**79**:450–8.

48 Liu D, Wang X, Pan F, *et al.* Web-based infectious disease reporting using XML forms. *Int J Med Inf* 2008;**77**:630–40.

49 Kim Y, Park H-A. Development and Validation of Detailed Clinical Models for Nursing Problems in Perinatal care. *Appl Clin Inform* 2011;**2**:225–39. doi:10.4338/ACI-2011-01-RA-0007

50 Khan WA, Hussain M, Afzal M, *et al.* Personalized-detailed clinical model for data interoperability among clinical standards. *Telemed E-Health* 2013;**19**:632–42.

51 Jing X, Kay S, Marley T, *et al.* Incorporating personalized gene sequence variants, molecular genetics knowledge, and health knowledge into an EHR prototype based on the Continuity of Care Record standard. *J Biomed Inform* 2012;**45**:82–92. doi:10.1016/j.jbi.2011.09.001

52 Hsu W, Taira RK, El-Saden S, *et al.* Context-based electronic health record: toward patient specific healthcare. *Inf Technol Biomed IEEE Trans On* 2012;**16**:228–34.

53 Hoy D, Hardiker NR, McNicoll IT, *et al.* Collaborative development of clinical templates as a national resource. *Int J Med Inf* 2009;**78**:S3–8.

54 Buyl R, Nyssen M. Structured electronic physiotherapy records. *Int J Med Inf* 2009;**78**:473–81. doi:10.1016/j.ijmedinf.2009.02.007

55 D'Amore JD, Mandel JC, Kreda DA, *et al.* Are Meaningful Use Stage 2 certified EHRs ready for interoperability? Findings from the SMART C-CDA Collaborative. *J Am Med Inform Assoc* 2014;:amiajnl – 2014–002883. doi:10.1136/amiajnl-2014-002883

56 Kalra D, Tapuria A, Austin T, *et al.* Quality requirements for EHR archetypes. In: *MIE*. 2012. 48–52.

57 Garde S, Hovenga EJ, Gränz J, *et al.* Managing archetypes for sustainable and semantically interoperable electronic health records. *Electron J Health Inform* 2007;**2**:e9.

58 Madsen M, Leslie H, Hovenga EJS, *et al.* Sustainable clinical knowledge management: an archetype development life cycle. *Stud Health Technol Inform* 2010;**151**:115–32.

59 Kohl CD, Garde S, Knaup P. Facilitating the openEHR approach-organizational structures for defining high-quality archetypes. *Stud Health Technol Inform* 2008;**136**:437.

60 Stroetmann VN, Kalra D, Lewalle P, *et al.* Semantic interoperability for better health and safer healthcare. European Commission, Directorate-General Information Society and Media 2009. http://dx.doi.org/10.2759/38514

Figure legends

Figure 1. Summary of Reference Models and their Clinical Information Model definition artifacts

Figure 2. Search performed in PubMed

Figure 3. Summary of the systematic review process

Figure 4. Summary of the clinical information modeling process steps