ICBO: International Conference on Biomedical Ontology
July 28-30, 2011 · Buffalo, NY, USA

A Framework Ontology for Computer-Based Patient Record Systems

Chimezie Ogbuji

Case Western University (School of Medicine), Cleveland, OH, USA

Abstract. A lack of uniform content and format standards remains the biggest barrier to the development of the Institute of Medicine's (IOM) Computer-Based Patient Record (CPR). The CPR ontology is a uniform core set of data elements (whose formal semantics are captured in OWL) for use in a Computer-Based Patient Record (CPR). It is meant to be of immediate and practical value to clinical research projects and patient registries that wish to leverage well-defined ontologies for a given domain.

1 Introduction

In 1991, the IOM defined the Computer-Based Patient Record (CPR) as:

> an electronic patient record that resides in a system specifically designed to support users by providing accessibility to complete and accurate data, alerts, reminders, clinical decision support systems, links to medical knowledge, and other aids [4].

This definition emphasizes data that are well-organized, discrete[1], and in a digital form.

One of the major goals of recent USA healthcare policy is to stimulate the *meaningful* use [2] of Healthcare Information Technology (HIT) in order to lay the groundwork for advanced electronic health information systems [7].

In describing the barriers to developing a CPR, the IOM declared that the content of CPRs must be defined such that each contains a uniform, core set of data elements that are named consistently. Some form of vocabulary control must be in place with specific meaning for data elements that describe clinical findings, clinical problems, procedures, and treatments [4]. The use of an ontology such as this one (the *CPR Ontology*) to govern a semantics for these data elements is meant to serve as an infrastructure skeleton to address this particular desiderata.

An ontology specifies a conceptualization of a domain and is often comprised of definitions of a hierarchy of concepts in the domain and restrictions on the relationships between them. An ontology with terms that serve as the basis for a uniform core set of data elements in a CPR will facilitate meaningful use of CPR content. In its current form, this ontology can be a high-level framework for collections of clinical vocabulary systems. All concepts are given a canonical syntax and semantics.

The scope of this ontology is over sub-domains of clinical medicine[2] as they relate to and appear in CPR content and with sufficient enough detail for immediate use in clinical care and research information systems. It can and has been used as a coordinate system for managing or eliciting precise meanings of terminology used by healthcare and research professionals for use in applications involving large scale data management.

To achieve maximal completeness of terminological coverage, this ontology leverages a number of normalization criteria that seek to minimize the amount of implicit differentiation between primitive concepts [11]. Definitions are given as chains of explicit restrictions on relations that differentiate two primitive concepts where one subsumes the other.

[1] The term discrete data is used to refer to data that are comprised of a finite set of values for variables. In the context of medical data, this is often referred to as coded data.

[2] The study of disease by direct examination of the living patient.

The CPR ontology	SNOMED-CT
clinical-finding	Clinical finding
procedure	Procedure
bodily-feature	Observable entity
organism	Organism
pharmacological-substance	Pharmaceutical / biologic product
recorded-clinical-situation	Situation with explicit context
clinical-artifact	Record artifact

Table 1. Mappings to SNOMED-CT

Like many contemporary ontologies that have a related goal, the CPR ontology is based on the Basic Formal Ontology. The following extant desiderata were followed in creating this framework:

1. Consistency with the top levels of the Basic Formal Ontology (BFO): reusing the terms *representational artifacts, occurrents*, and *continuants* [17];
2. A clear separation of digital ("information content") entities and the phenomenon they represent, or describe;
3. The use of a clinically-oriented framework for representing diseases, their causes, manifestations, related diagnostic acts, and pathophysiological phenomena. For these we leverage terms from the BFO, [13, 18], and SNOMED-CT.
4. A principled integration [12] with a reference ontology for (canonical) human anatomy.

2 Related Work

There are a number of issues with existing biomedical ontologies that the author has encountered while trying to harmonize this ontology with them. Contemporary ontology standardization and exchange is often hegemonic, but despite this, ontology harmonization and mapping can support interoperability. This ontology is meant to be used in this way with the contemporary BFO-based ontologies for clinical medicine and patient record content. In this section, we briefly summarize and characterize the overlap between this ontology and existing biomedical ontologies.

There are few principled ontology frameworks or vocabularies that comprehensively meet the needs of a contemporary patient record or (disease) registry. In Bodenreider's review [3] and analysis of high-impact biomedical ontologies, only one of the nine was listed with a scope of clinical medicine and patient record content: SNOMED-CT.

Neither the CPR ontology by itself nor many of the current BFO-based ontologies such as the Ontology of Biomedical Investigation (OBI), the Ontology for General Medical Science (OGMS), Biotop, or the Ontology of Information Artifacts (IAO) – by themselves (without extensions) – meet the IOM desiderata for use in patient record and registry systems in the manner that SNOMED-CT does. However, in sharing a common foundation, this ontology was designed to be semantically interoperable with them. Many existing biomedical ontologies don't have a framework for representing the process of diagnostic inquiry and relationships with the information content that result from them. They also do not have a framework for the representation of a clinical diagnoses and the role of treatment in classifying diseases in a manner similar to that proposed in [19].

This ontology predates[3] OBI (but not the MGED Ontology on which it is based), IAO, and OGMS. It has evolved as a result of interaction with the authors of many of them and is meant to facilitate harmonization for the benefit of users of both. A more comprehensive assessment of where these ontologies fail to provide coverage sufficient for use in detailed representation of the core data elements of a CPR is beyond the scope of this work.

The CPR ontology	The FMA
anatomical-structure	Material anatomical entity
immaterial-anatomical-continuant	Immaterial physical anatomical entity

Table 2. Mappings to the FMA

[3] http://www.w3.org/wiki/HCLS/POMROntology (December 2006)

3 The Structure of the CPR Ontology

3.1 Running Example

Iuliano et al. [6] report the case of a young female patient with a previously clinically diagnosed Myocardial infarction (MI), a medical history of hypertension, no history of diabetes mellitus, hypercholesterolemia or premature coronary artery disease in her family. Ten years earlier, she had an inferior MI treated with systemic thrombolysis and other prescribed pharmacological substances.

During her hospitalization at that time, the echocardiogram revealed akinesis of the posterior-basal wall with an estimated ejection fraction of 50%. Laboratory tests including serum glucose and other requested components were performed. During the more recent hospitalization, a physical examination showed blood pressure of 135/80 mmHg, heart rate of 64 BPM, BMI of 27 kg/m2, and waist circumference of 85 cm.

In subsequent sections, we will demonstrate how many of the clinical concepts involved in this case can be represented using the CPR ontology.

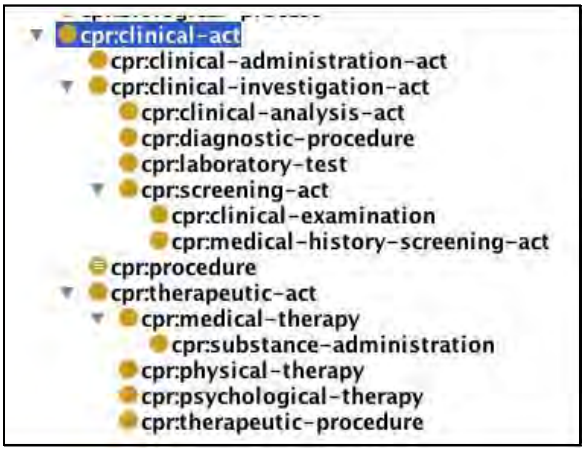

Figure 1. Diagram of clinical act taxonomy

3.2 Clinical acts

The **clinical-act** class corresponds to the root of a hierarchy of the actions that comprise the healthcare workflow, This hierarchy is placed directly under the *span:ProcessualEntity* class in the BFO ontology as shown in figure 1.

The recorded presence and absence of a medical history is collected as a result of a **medical-history-screening-act**. These are a **screening-act** that results in anamnesis [4]. Other clinical acts that correspond to concepts in the case report are: **diagnostic-procedure** (echocardiogram), **laboratory-test**, **therapeutic-procedure** (thrombolysis), and **substance-administration** (prescription of the various drug regiments).

A **clinical-investigation-act** is distinguished from a clinical act by the restriction that they have an active participant playing an investigative role. Similarly, a therapeutic act is distinguished from a clinical act by the restriction that they have a continuant playing a therapeutic role.

An echocardiogram diagnostic process can be represented in this way (using the American Heritage Stedman's Medical Dictionary definition):

```
diagnostic-procedure THAT
  hasMethod SOME UltrasoundRecording AND
  investigates SOME
    ( hasConsequence SOME
      ( pathological-disposition THAT
        ro:located_in SOME Heart ) ) AND
  hasOutput SOME Echocardiogram
```

The **investigates** relation holds between a diagnostic procedure and either an *etiological agent* or some indication of a **therapeutic-act**. The former is defined in 3.6 along with a *pathological disposition*. An Echocardiogram is a representational artifact, which is defined in the next section. In our example, the thrombolysis was indicated (**hasIndication**) by the result of one or more diagnostic procedures.

3.3 Representational Artifact

In [17], a *representational artifact* is defined as an idea, image, record, or description that refers to (is of or about), or is intended to refer to, some entity or entities external to the representation.

It is the basis for a distinction between information content [5] and the phenomena they represent, or describe. Its use in the CPR ontology is meant to resolve certain ontological inconsistencies [16] that exist in common vocabulary standards such as HL7 RIM that

[4] a patient's account of a medical history

conflate the two. It denotes CPR content or a component thereof that stand in the **representationOf** relation to some entity.

The primary representational artifact relevant to the domain is the **clinical-artifact** class. It is explicitly differentiated from a representational artifact by the restriction that its instances are composed by a person playing a relevant role in the care process and have a person as the subject of their description. The **subjectOfDescription** relation holds between a clinical artifact and its subject. It is similar to the *SUBJECT RELATIONSHIP CONTEXT* attribute in SNOMED-CT that is used to specify the subject of the *clinical finding* or *procedure* being recorded. Clinical artifacts are also often consumed by (**hasInput**) or produced from (**hasOutput**) a **clinical-act** such as an echocardiogram diagnostic process.

3.4 Clinical Findings

Expanding on the definition from [13], the CPR ontology defines the **clinical-finding** primitive as a **clinical-artifact** that is composed by a clinician, is the **outputOf** a **clinical-act**, and represents a bodily feature. Examples of findings from the case report are the medical history of hypertension (**anamnesis**) and akinesis of the posterior-basal revealed from the previous echocardiogram.

3.5 Surgical Deeds

This CPR ontology distinguishes a **procedure** from a **clinical-act** by the following explicit existential role restrictions: 1) it **actsOn** [8] an organismal continuant [12] or object, as in the case of device insertions; 2) it stands in an *approach site* relation to an immaterial anatomical entity (a cavity, whole, etc.); 3) It stands in a **hasMethod** relation[5] with the deed [8] or method that achieved it. SNOMED-CT specifies the semantics of the *SURGICAL APPROACH* attribute as the directional, relational, or spatial access to the site of a surgical procedure.

A **diagnostic-procedure** (such as an echocardiogram) is a primitive procedure that is also a **clinical-investigation-act**.

3.6 Disease Manifestation, Etiology, and Pathophysiology

The concept of a disease as a pathological disposition and its relation to a disease course is introduced in [13]. A disposition is an attribute of an organism in virtue of which it will initiate specific processes when certain conditions are satisfied. Whitbeck's theory [18] of *etiology* and *disease entities* is incorporated in the CPR ontology **pathological-disposition** term.

It is a primitive that is explicitly distinguished from a disposition via the following existential role restrictions: 1) it is located in a patient 2) it stands in the **isConsequenceOf** relation to one or more instances of **etiological-agent**, either directly or via transitive closure of this causal relationship; 3) it is realized[6] as a pathological process that has, as a component, consistent *physiological* or *anatomical alterations* (**morphological-alteration**). The **isConsequenceOf** relation denotes a causal relationship between entities (occurrent and continuant alike).

3.7 Diagnosis as Clinical Analysis and Hypothesis

Whitbeck provided [19] an ontological basis for clinical diagnoses, how they fit into the terminology of clinical medicine, and their relationship to diseases and etiological agents. The CPR ontology defines a **clinical-diagnosis** as a clinical artifact that is the output of a **clinical-investigation-act**. It stands in the **hypothesizedProblem** relation to a disease that is hypothesized by a clinician to inhere in the patient. Intuitively, this is the recording of the conclusion to an interpretive (scientific) process that investigates problems that require management.

[5] SNOMED-CT defines a METHOD attribute of procedures that relates it to the action being performed to accomplish the procedure.

[6] Currently, a causal, agentive relation is used, however, it is not clear (given [13, 18]) if the relationship between a disease and anatomical alterations is causal or mereological.

3.8 Relationship with Measurement Unit Ontology

The CPR ontology includes a **measurementOf** relation that serves as a bridge between it and the Measurement Unit Ontology (MUO)[7] . It holds between a clinical finding and a **muo:QualityValue**. The MUO defines a *quality value* as the (reified) value of a quality that is related to exactly one unit of measure (via the **muo:measuredIn** relation) and a corresponding scalar quantity (via the **muo:numericalValue** relation).

So, the semantics of the **measurementOf** relation can be captured in the following role chain:

$$subjectOfDescription \circ bearerOf \circ measuresQuality^{-} \circ measuredIn^{-}$$

In this representation, *bearerOf* corresponds to the implicit BFO relationship between an independent continuant and the dependent continuant (quality, function, etc.) that inheres in it and **muo:measuresQuality** is the MUO relation that holds between a unit and the quality it measures.

An alternative approach is to create a new role under **muo:qualityValue** in the role hierarchy that holds between a continuant and the value of a quality associated with the role. So, going back to our medical case, the diastolic and systolic blood pressure values recorded as a result of the physical examination can be captured using a *diastolic* and *systolic* role each of which is subsumed by **muo:qualityValue** and understood to hold between a person and the values of the diastolic and systolic blood pressure qualities, respectively and at the time of measurement.

4 Discussion

In formulating and curating this ontology, there were a number of issues that straddle the line between metaphysics and the philosophy of medicine but are mentioned here.

4.1 Qualities of occurrents

In Sayed's comparison [14] of BFO and DOLCE, the following question was raised:

> Why can't a heartbeat rate be a quality of its heartbeating [sic] event, given it has no meaning outside of this event?

The consequence of BFO's requirement that qualities can only inhere in continuants is that not all bodily features can have qualities, and thus a concept corresponding to the quality of a heart-beating event or process cannot be straightforwardly represented. A similar example is from the medical case.

Ejection fraction is a measure of the fraction of blood pumped out of the right and left ventricles with each heart beat. Intuitively, it seems inappropriate to consider it a measure of the quality inherent in the heart or in the ventricles but rather as a measurement of the flow of blood through the latter.

4.2 Transitive Causal Chains

Another issue is related to transitive, causal chains between two continuants or where the agent is a process. The CPR ontology re-uses the ro:agent in from the Relations Ontology. However, that relation is designed to explicitly rule out the inheritance of agency along causal chains [15] and to ensure that only continuants are agents. This restriction on ro:agent in and the fact that a disease is modeled as a continuant presents a challenge and is the reason why the **isConsequenceOf** is needed.

In describing an etiological agent, Whitbeck gives two criteria that govern the choice of the causal factor to be regarded as the etiologic agent. The first is proximity and the second is a preference for a factor which exists in the environment prior to contact with the patient's body and which may then act upon it [18]. She describes the rationale for this second criterion in the way in which we classify diseases in which it is the toxins elaborated during metabolism and growth of the pathogen which cause the damage.

A detailed account of this shortcoming and of the role of causality in an ontological account of clinical medicine is beyond the scope of this paper. However, the interested reader should see [18].

4.3 Signs, Symptoms, and Epistemology

As mentioned earlier, there are issues associated with the semantics of the medical

[7] http://forge.morfeo-project.org/wiki en/index.php/Units_of_measurement_ontology

terms *symptom* and *sign*. In the CPR ontology, they are modeled as representational artifacts and not special kinds of bodily features. It is generally a bad idea to mix ontology and epistemology [10]. The former is about things in reality and the latter is regarding how cognitive subjects come to know the truth about phenomena. Some bodily features can exist before they are perceived by a patient. The patient may subsequently hypothesize that the bodily features are symptoms that indicate a disease. However, they are not considered symptoms initially.

So, there is nothing about a cough (for example) that makes it a symptom besides the way it is first perceived and then reported. Even if a symptom or finding is recorded by a nurse as a result of eliciting anamnesis from the patient, it is the patient that originally presents with it. This is not an ontological distinction and is mostly also the case with signs as well. However, what makes a bodily feature a sign is a little more objective [13]. A sign is typically observed in a **clinical-examination** that is performed by a clinician. Still, the primary distinction between signs, symptoms and other bodily features is with regards to who observed them.

Since this distinction can be captured via the **composedBy** relation between a **clinical-artifact** and a **person** the CPR ontology has two defined classes: **record-of-symptom** and **record-of-sign**. It also has a **self-examination** class that is an action performed by a patient on his or herself to determine the existence of a medical problem.

So, the logical distinction between a symptom and other bodily features is that they are represented by a clinical artifact with provenance indicating authorship by the patient.

5 Conclusion

This paper introduces a minimal, high-level framework that addresses the CPR requirements of a core, common, data dictionary and facilitates the meaningful use of discrete CPR content via biomedical ontologies. It contributes to meaningful use in the following ways: ontology is used to facilitate data that is discrete, precise in its meaning for the purpose of communication and use in different contexts such as reporting quality measures and clinical care and research information system infrastructure, and as a framework for capturing a large number of clinical conditions.

In using an independent, running example of a medical case, we are able to demonstrate how the CPR ontology acts as a coordinate system that is sufficient to capture the key axes of meaning. As a result, this ontology is able to post-coordinate terms from within a large region of the domain of interest: patient record content. It supports terms that capture the specific meaning of clinical findings, problems, procedures, and treatments.

This ontology has been used as part of the *SemanticDB*[8] project in the Cleveland Clinic's Heart and Vascular Institute, where the author was a lead software architect. It was also used as part of the development of recent research methods for managing disparate relational, polysomnography data via web-based management [1, 9].

A number of issues are still unresolved and raised as input to the frameworks leveraged by the CPR ontology. The ontology is part of a project hosted on Google Code repository[9]. Any feedback or comments are welcome and can be directed to the associated Google Group[10].

Acknowledgements

The author would like to thank the following people for their substantive input into this ontology: Alan Rector M.D. and Ph.D., Sivaram Arabandi M.D. and M.S., Eugene Blackstone M.D., Barry Smith Ph.D., Michel Dumontier Ph.D., Songmao Zhang Ph.D., GQ Zhang Ph.D., Satya Sahoo Ph.D., Helen Chen Ph.D., and Jos Deroo.

References

1. S. Arabandi, Ogbuji C., Redline S., Chervin R., Boero J., Benca R., and Zhang G.Q. Developing a Sleep Domain Ontology (abstract). American Medical Informatics Association, 2010.
2. D. Blumenthal. Launching HITECH. New England Journal of Medicine, 2010.

[8] http://www.w3.org/2001/sw/sweo/public/UseCases/ClevelandClinic/

[9] http://code.google.com/p/cpr-ontology

[10] http://groups.gpogle.com/group/cpr-ontology

3. O. Bodenreider. Biomedical ontologies in action: role in knowledge management, data integration and decision support. Yearb Med Inform, 67:79, 2008.
4. R.S. Dick, E.B. Steen, and D.E. Detmer. The computer-based patient record: an essential technology for health care. Natl Academy Pr, 1997.
5. M. Dumontier and R. Hoehndorf. Realism for scientific ontologies. In Proceeding of the 2010 conference on Formal Ontology in Information Systems: Proceedings of the Sixth International Conference (FOIS 2010), pages 387–399. IOS Press, 2010.
6. L. Iuliano, F. Micheletta, A. Napoli, and C. Catalano. Myocardial infarction with normal coronary arteries: a case report and review of the literature. Journal of Medical Case Reports, 3:24, 2009.
7. B. Kadry, I.C. Sanderson, and A. Macario. Challenges that limit meaningful use of health information technology. Current Opinion in Anesthesiology, 23(2):184, 2010.
8. A. Mori, A. Gangemi, G. Steve, F. Consorti, and E. Galeazzi. An ontological analysis of surgical deeds. Artificial Intelligence in Medicine, pages 361–372, 1997.
9. C. Ogbuji, S. Arabandi, S. Zhang, and G.Q. Zhang. Segmenting and merging domain-specific ontology modules for clinical informatics. pages 414–427, 2010.
10. B.S. Olivier Bodenreider and A. Burgun. The ontology-epistemology divide: A case study in medical terminology. In Formal ontology in information systems: proceedings of the Third International Conference (FOIS-2004), page 185. Ios Pr Inc, 2004.
11. A.L. Rector. Modularisation of domain ontologies implemented in description logics and related formalisms including OWL. In Proceedings of the 2nd international conference on Knowledge capture, pages 121–128. ACM, 2003.
12. C. Rosse, A. Kumar, J.L.V. Mejino Jr, D.L. Cook, L.T. Detwiler, and B. Smith. A strategy for improving and integrating biomedical ontologies. In AMIA Annual Symposium Proceedings, volume 2005, page 639. American Medical Informatics Association, 2005.
13. R.H. Scheuermann, W. Ceusters, and B. Smith. Toward an ontological treatment of disease and diagnosis. Proceedings of the 2009 AMIA Summit on Translational Bioinformatics, pages 116–120, 2009.
14. A.P. Seyed. BFO/DOLCE Primitive Relation Comparison. 2009.
15. B. Smith, W. Ceusters, B. Klagges, J. Kohler, A. Kumar, J. Lomax, C. Mungall, F. Neuhaus, A.L. Rector, and C. Rosse. Relations in biomedical ontologies. Genome biology, 6(5):R46, 2005.
16. B. Smith and W. Ceusters. HL7 RIM: An incoherent standard. Studies in health technology and informatics, 124:133–138, 2006.
17. B. Smith, W. Kusnierczyk, D. Schober, and W. Ceusters. Towards a reference terminology for ontology research and development in the biomedical domain. In Proceedings of KR-MED, volume 2006, pages 57–65. Citeseer, 2006.
18. C. Whitbeck. Causation in medicine: The disease entity model. Philosophy of Science, 44(4):619–637, 1977.
19. C. Whitbeck. What is diagnosis? Some critical reflections. Theoretical Medicine and Bioethics, 2(3):319–329, 1981.