OGMS: The Ontology for General Medical Science

Albert Goldfain[1§], Robert Rovetto[2], Lindsay G. Cowell[3], Melanie Courtot[4], Alan Ruttenberg[2], Yu Lin[5], Alexander D. Diehl[2], Mark Jensen[2], Alexander P. Cox[2], Sivaram Arabandi[6], William R. Hogan[7], Mathias Brochhausen[7], Yongqun He[5], Prabhu Shankar[8], Kerstin L. Forsberg[9], Richard H. Scheuermann[10], Werner Ceusters[2], Barry Smith[2*]

[1]Blue Highway, Inc., Syracuse, NY, USA
[2]University at Buffalo, Buffalo, NY, USA
[3]University of Texas Southwestern Medical Center, Dallas, TX, USA
[4]British Columbia Cancer Research Centre, Vancouver, BC, Canada
[5]University of Michigan Medical School, Ann Arbor, MI, USA
[6]ONTOPRO LLC, Houston, TX, USA
[7]University of Arkansas for Medical Sciences, Little Rock, AR, USA
[8]Emory University, Atlanta, GA, USA
[9]AstraZeneca, Mölndal, Sweden
[10]J. Craig Venter Institute, San Diego, CA, USA

*All authors contributed equally to this work
[§]Corresponding author

Email addresses:
AG: albertgoldfain@gmail.com
RR: rrovetto@buffalo.edu
LGC: lindsay.cowell@utsouthwestern.edu
MC: mcourtot@gmail.com
AR: alanruttenberg@gmail.com
YL: linikujp@gmail.com
ADD: addiehl@buffalo.edu
MJ: mpjensen@buffalo.edu
APC: apcox@buffalo.edu
SA: sivaram.arabandi@gmail.com
WRH: hoganwr@gmail.com
MB: mbrochhausen@uams.edu
YH: yonggunh@med.umich.edu
PS: prshank@emory.edu
KLF: kerstin.l.forsberg@gmail.com
RHS: rscheuermann@jcvi.org
WC: ceusters@buffalo.edu
BS: phismith@buffalo.edu

Abstract

Background

The Ontology for General Medical Science (OGMS) is introduced. OGMS serves as a reference ontology for clinical medicine within the framework of the Basic Formal Ontology (BFO) and the Open Biomedical Ontology (OBO) Foundry. The latest release of OGMS in OWL format is freely available at the persistent URL: http://purl.obolibrary.org/obo/ogms.owl.

Construction and Content

The core terms of OGMS are defined, including: 'disease', 'disorder', and 'disease course'. The relations between these terms, drawn from the OBO Relation Ontology are also presented. The OWL implementation of OGMS, which represents these relations as description logic restrictions, is briefly discussed.

Utility and Discussion

OGMS is intended to be used as a framework that can be extended for biomedical ontology development and the construction of clinical-annotation applications that are interoperable with other OBO Foundry ontologies. Various ontologies and applications that utilize OGMS are discussed.

Background

The domain of clinical medicine is a difficult one from an ontology development perspective. Clinical terminology can be inconsistent, vague, and highly dependent on disciplinary context. A formal, explicit, unique, and unambiguous representation of clinical terms can begin to address these difficulties. Currently, there is no shortage of ontologies, terminologies, controlled vocabularies, thesauri, and coding systems intended to support clinical applications. However, the support provided by such resources is often restricted by the fact that it is difficult (if not impossible) to utilize them beyond their initial purpose. Also, such resources contain terminological artifacts – such as use of 'Not Otherwise Classified' – which create difficulties for data integration and formal reasoning and also for consistent management of data over time. For example, the International Classification of Diseases (ICD) is typically used for medical billing and to classify morbidity data during encounters, but it is not built to allow advanced clinical decision support or automated inference and reasoning. Moreover, like terminologies such as SNOMED-CT, it is not maintained in a way that allows it to respond in agile fashion to new developments in translational medicine. We believe OBO Foundry ontologies are in a good position to support the goals of clinical informatics (Smith *et al*, 2007). We also believe that a new formal ontology is needed within the OBO Foundry to explicitly to address the challenges of the clinical domain. The Ontology for General Medical Science (OGMS), presented herein, is intended to address this need.

OGMS is an ontology of the main types of entities involved in a clinical encounter, including those denoted by terms such as 'disease', 'disorder', 'disease course', 'diagnosis', and 'clinical finding'. These entities are referred to in handwritten and dictated notes, charts, electronic health records, clinical research databases, prescription systems, discharge summaries, and reimbursement systems, but these references typically employ terms that are ambiguous and that do not highlight subtle yet important distinctions. Scheuermann, Ceusters, and Smith (2009) laid the groundwork for OGMS by describing a core set of general clinical entities that are classified within the framework of the Basic Formal Ontology (BFO) and the Open Biomedical Ontology (OBO) Foundry. These entities were represented using precisely defined terms. OGMS developers have attempted to adhere closely to OBO Foundry principles[1]. They have paid special attention to the formulation of logically coherent definitions in order to enable more powerful algorithmic processing of clinical data.

Various authors have pointed out the need for a formal ontological treatment of disease (Schulz *et al*, 2010; Bello *et al*, 2011; Mizoguchi *et al*, 2011). Although OGMS is not itself a disease ontology, it can (and has) been used as a framework from which to build ontology modules for a range of different diseases and disease families. OGMS provides a general theory of disease and formal definitions for terms widely used in clinical encounters to describe different aspects of disease. These terms will be further elaborated by specific disease ontologies and by clinical ontologies.

Breadth and Scope

OGMS is a reference ontology intended to be used primarily in the development of clinical application ontologies. Towards this end, OGMS is intentionally small, consisting of approximately 100 terms, and only importing terms from BFO, the Information Artifact Ontology (IAO), and the Ontology for Biomedical Investigations (OBI). In contrast to other reference ontologies that focus on describing the parent-child type-subtype relations of entities within a given domain, e.g., cell types in the Cell Ontology[2] or biological process types in the Gene Ontology[3], OGMS focuses on representing high-level disease and clinical encounter entities and the relations between them. It is built around a view of disease as a certain sort of power or potentiality (roughly the potentiality for signs and symptoms to be manifested), which exists in organisms in virtue of physical disorders in the organism. These powers or potentialities are called 'dispositions', and so OGMS defends the general view according to which diseases are special types of dispositions.

OGMS is designed to serve as the framework for describing how specific classes of physical disorders relate to abnormal dispositions manifesting themselves in pathological processes which are recognized as signs or symptoms in clinical encounters and documented in clinical information

[1] http://www.obofoundry.org/crit.shtml
[2] http://cellontology.org/
[3] http://www.geneontology.org/

systems. Any ontology can import all of OGMS without much overhead. Clinical application ontologies extend OGMS by extending (and thereby refining) the basic taxonomic and relational structure for a particular domain of interest. In so doing, an OGMS-conformant extension can become compliant with BFO without having developers delve into the details of BFO types. An OGMS-conformant extension will thus possess a great degree of interoperability with other such extensions.

The current focus of OGMS development has involved human organisms; however many OGMS terms are compatible with veterinary medicine and even with representations of diseases involving non-human hosts (e.g., plants) (Walls *et al*, 2012). The only requirement is that users have some resources to describe the canonical anatomy and physiology of the organism that is the subject of care.

Construction and content

The development of OGMS is a community effort and is undertaken with clinical input and an eye towards clinical applicability. A group has also been established for extended discussion of theoretical and technical issues involving OGMS[4]. The latest release of OGMS in OWL format is freely available at the persistent URL: http://purl.obolibrary.org/obo/ogms.owl. A Google Code repository[5] is used for source control and issue tracking[6]. OGMS is published to the OBO Library[7], Ontobee[8], and BioPortal[9]. Individual OGMS term URL identifiers are dereferenced to self-sustained HTML and RDF documents automatically by the Ontobee linked data server program.

Natural language definitions are provided for OGMS terms using the Aristotelian form: "*[species]* is a *[genus]* that *[differentia]*". Wherever possible, the OWL representations attempt to match the expressivity and semantics of these natural language definitions, but sometimes these definitions are for the sake of clarity more expressive than their OWL counterparts.

Core Terms

OGMS is organized around a few core entities to create an ontological framework from which multiple different sorts of extensions can be built. Figure 1 shows the core terms of OGMS and the relations that hold between the entities denoted by them. All of the relations are taken from the OBO Relation Ontology (RO) except **has_material_basis**, which is a relation included in BFO2.

[4] http://groups.google.com/group/ogms-discuss
[5] http://code.google.com/p/ogms/
[6] http://code.google.com/p/ogms/issues/list
[7] http://www.obofoundry.org/cgi-bin/detail.cgi?id=OGMS
[8] http://www.ontobee.org/browser/index.php?o=OGMS
[9] http://bioportal.bioontology.org/ontologies/1414

[FIGURE 1 HERE]

Figure 1 OGMS core terms and RO relations between them. Each relation R(x,y) shown here is a type-type relation and is rendered in OWL as "x R SOME y", with universals x and y represented in OWL by classes and relation R represented by an object property. The fact that not all diagnoses are only sometimes about diseases is indicated with a dashed arrow.

The core terms of OGMS are designed to include by those disease-neutral expressions that are most generally used in clinical medicine - expressions that also are frequently conflated in natural language contexts. Consider the way in which terms like "HIV" may denote either a certain virus type, a specific collection of instances of this virus type, the HIV infectious disorder, the HIV infectious disease, the HIV infectious disease course, the diagnosis of HIV infectious disease, and so on. OGMS can be used to factor ambiguous terms such as "HIV" into non-ambiguous terms whose meaning and reference is clear. We believe this is an important step in annotating clinical records and exchanging meaningful electronic health information.

Below, we cover three of the core terms: 'disorder', 'disease', and 'disease course'. These represent a minimal subset of terms from which an OGMS-conformant extension ontology can be built.

Disorder

The current definition of disorder in OGMS is:

disorder: *A disorder is a material entity which is clinically abnormal and part of an extended organism.*

More will be said about 'extended organism' below, for now it can be thought of as a synonym of 'organism'. A disorder is the material basis of a disease. A disease cannot exist without some disorder to serve as its material basis. A disorder is a part of the organism: for example, it is an enlarged facet joint, an inflamed liver, a carcinomatous lung, or an accumulation of apoptotic airway epithelial cells in a case of chronic obstructive pulmonary disease (Smith, Ceusters, Kumar, and Rosse, 2005).

OGMS classifies disorders as BFO material entities in order to remain flexible as to the granularity of any particular disorder and in order to do justice to the fact that disorders may change their granularity as they evolve over time, as when a disturbance at the molecular level of granularity evolves into a tumor at the level of coarse anatomy.

Although disorders are independent continuants in BFO terms, every disorder has a corresponding quality or qualities that make it a disorder (as opposed to an "ordered" canonical anatomical entity). An organism must undergo some sort of etiological process in which a certain ordered configuration becomes a disordered configuration which can then evolve through additional processes, including pathological transformation, pathological derivation, and pathological

invasion (Smith, Ceusters, Kumar, and Rosse, 2005). These are occurrent parts of the process leading from etiology to evolution of a disorder. An example of pathological transformation is the formation of a tumor mass. An example of 'pathological derivation' is the vertical transmission of HIV immediately before and after birth. An example of 'pathological invasion' is the colonization of a host by infectious organisms.

The notion of clinical abnormality mentioned in the definition of 'disorder' above is an undefined primitive term in OGMS. What is clinically abnormal will vary across different clinical domains and thus is not a part of what we are calling 'general medical science'. The term 'clinically abnormal' is intended to set a threshold for what things count as disorders. All parts of a particular organism will have slight deviations from canonical anatomical parts, but most do not reach the threshold of clinical significance.

The key idea here is that some disorders may or may not yield a disposition towards a pathological process. A population of infectious organisms may be contained in an organism without becoming an infectious disorder. The threshold can be set in a probabilistic way: when a particular material entity in an organism is of a type that reliably and repeatably gives rise to a particular disease, then the material entity in question is clinically significant, and thus a disorder.

Ontologies using OGMS may represent subclasses of disorder by using the method of cross-products (Hill *et al,* 2002; Smith *et al*, 2005). For disordered anatomical parts this will mean using terms from the Foundational Model of Anatomy (Rosse and Mejino, 2007) to represent subtypes for example of 'bone disorder' or 'liver disorder'. Not every case of a cross-product definition will involve the disorder being *located in* the anatomical entity. Cross-product definitions may refer, for example, to disorders affecting the communication of fluids between the body's anatomical compartments. Also, part of a disorder may be located in an anatomical entity without being a disorder *of that entity* (for example a metastatic lung tumor may spread to the liver without being a liver disorder). Other disorders may involve foreign substances in an organism as parts of disorders, for example in a case of infectious disorders, covered by the Infectious Disease Ontology (IDO), are examples of a disorder with parts (i.e. the infectious agents) that are either foreign to the host organism, or have become overly abundant due to an immunodeficiency or immunosuppression on the part of the host organism, and have moved to a foreign anatomical location in which they can do some harm to the host organism. Infectious disorders can exist in organism sites that are holes and cavities. The extended organism includes these immaterial entities as parts:

extended organism*: An object aggregate consisting of an organism and all material entities located within the organism, overlapping the organism, or occupying sites formed in part by the organism.*

An extended organism can be thought of geometrically as the interior of the convex hull (external boundary) of the organism. Every organism has a corresponding extended organism. If an organism without any holes or cavi-

ties were to exist, the terms ‘organism’ and ‘extended organism’ would for that organism be synonomous.

Disorders in the OGMS sense are not processes or dependent continuants. As such, many things which contain the label ‘disorder’, such as ‘sleep disorder’ or ‘obsessive compulsive disorder’ are not disorders in the OGMS sense (unless, of course, such a label were used to describe a potential material basis in the brain for what might be called obsessive compulsive disease).

Disease

Disease is defined in OGMS as follows:

disease*: A disposition (i) to undergo pathological processes that (ii) exists in an organism because of one or more disorders in that organism.*

A disease, like any other disposition, can be borne without being realized. This is especially salient for diseases that can have long periods of dormancy. A colony of infectious organisms may establish an infection well before pathological processes begin to occur. A disease is borne at the same time as the corresponding disorder develops, i.e. when the clinical significance threshold is crossed and the organism bears a tendency towards such processes.

As was mentioned above, any part of a particular disorder is either part of the organism or contained in the extended organism, and is the material basis for a disease. However, diseases are borne by whole organisms. Clinical medicine sometimes refers to such things as a ‘diseased liver’, suggesting that the organism part is the bearer of the disease. However, a liver that is removed from an individual’s body cannot bear a liver disease because the rest of the body’s functioning is required for the realization of a pathological process. Livers must be embodied to realize both their functions in times of health and their dysfunctions during liver disease courses. So in OGMS, the bearer of a disease is always an entire organism.

Although instances of disease instances are typically ascribed to organisms, the realization of a disease often requires a complex interaction between the organism bearing the disease, other organisms, and environmental factors that may be either internal or external to the organism. A particular instance of the flu, for example, is realized in a host only when the influenza virus realizes its disposition to colonize and replicate in the host organism and when the internal environment of this host realizes a disposition to serve as an infection site. These secondary dispositions can be modelled as pairwise complementary dispositions whose mutual realization realizes the disease (Goldfain, Smith, and Cowell, 2010).

Disease Course

The definition of ‘disease course’ is:

disease course*: the totality of all processes through which a given disease instance is realized.*

This 'totality' includes pathological processes that are recognized through signs and symptoms; but it also includes potentially asymptomatic early stages of a disease and pathological processes occurring during convalescence. Disease courses, particularly chronic disease courses, are in many cases divided into stages. The process boundaries of these stages may be described as being fiat, bona fide, vague or even left unspecified. This is important clinically as the recommended treatment options for many diseases change depending on the stage of the disease course the patient is experiencing. The etiological processes leading up to the manifestation of a disease also may have a complex parthood structure that clinicians refer to. Infectious disease courses, for example, are preceded by a vector transmission process.

Currently, OGMS subtypes disease course according to the different ways a disease course may unfold. For example, 'chronic disease course' is a subtype of 'disease course'. When clinical medicine recognizes a repeatable progression of a disease course, a term is often introduced to refer to it. For example, HIV researchers have recently identified a long-term non-progressive type of disease course in HIV positive individuals (Walker, 2007).

Controversial Terms

We believe that the development of OGMS should be guided by the usage of general terms in clinical practice. Whenever possible, the users of OGMS should be shielded from the debates about ontological commitments and the implementation details of the representation formalism in which OGMS is rendered. Early development of OGMS has suggested that a few terms used in clinical medicine cannot be given a formal definition that is easily reconciled with general usage and the realist assumptions of BFO. For example, the term 'symptom' can be used in natural language to refer to a subjective mental experience (e.g., the feeling of vertigo), an observable quality (e.g., the redness of a rash), or a process (e.g., a bout of coughing). To insert a single term into the BFO framework that accommodates all of these would violate the disjointness of continuants and occurrents, which is a presupposition of the BFO architecture. In such cases, rather than abandoning the realist framework of BFO, we believe that clinical informatics researchers can benefit by thinking about the ontological commitments which they make when they refer to entities of different sorts. Simply by providing a robust framework for a discussion of these difficult terms, we believe the development of OGMS has exposed some of the inconsistencies in usage and potential formalizations of the problematic terms.

Organization

OGMS is organized in a single hierarchy beneath BFO. In its use of BFO, OGMS inherits the commitments of ontological realism. Terms in OGMS denote universals and these terms are used to form singular referring expres-

sions (such as 'John's subglottic stenosis') to refer to particular instances of these universals. As of this writing, the released version of OGMS is based on BFO version 1.1, however a migration to BFO2 has already begun in the development version of OGMS, including temporalized relations and process patterns. BFO makes a top level distinction between continuants and occurrents. Figure 2 illustrates part of the asserted hierarchy of continuants and occurrents in OGMS.

Additionally, OGMS includes some undefined primitive terms which are given an elucidation rather than a formal definition.

OGMS utilizes part of the Information Artifact Ontology (IAO), to make a clear distinction between information content entities, such as clinical findings, and the referents they are about.

[FIGURE 2 HERE]

Figure 2 Part of the OGMS Hierarchy

Utility and Discussion

OGMS is being actively used by various application ontologies. Typical uses range from importing a few general terms to using OGMS as an interface point to BFO to organizing the entire taxonomy around the OGMS core terms. Some active application ontologies which utilize OGMS are: the Cardiovascular Disease Ontology[10], Oral Health and Disease Ontology[11], Health Data Ontology Trunk[12], Vital Sign Ontology[13], and Ontology of Medically Relevant Entities[14], Mental Functioning Ontology[15], Ontology for Pain Mental Health and Quality of Life, Sleep Domain Ontology[16], Ontology for Newborn Screening and Translational Research[17], Infectious Disease Ontology[18], Neurological Disease Ontology[19], Ontology of Adverse Events[20], and Adverse Event Reporting Ontology[21]. Below we describe in more detail some of the ways these ontologies use OGMS.

Mental Functioning Ontology (MFO)
The basis of MFO is laid down by the introduction of what are called 'cognitive representations' – and other entities such as beliefs, emotions and desires – which in line with OGMS have a physical basis (in the brain), in the relevant

[10] https://code.google.com/p/cvdo/
[11] https://code.google.com/p/ohd-ontology/
[12] http://code.google.com/p/hdot/
[13] https://code.google.com/p/vital-sign-ontology/
[14] https://code.google.com/p/ogms/wiki/OMRE
[15] http://code.google.com/p/mental-functioning-ontology/
[16] http://purl.bioontology.org/ontology/SDO
[17] http://purl.bioontology.org/ontology/ONSTR
[18] http://infectiousdiseaseontology.org/page/Main_Page
[19] https://code.google.com/p/neurological-disease-ontology/
[20] http://www.oae-ontology.org/
[21] http://code.google.com/p/adverse-event-reporting-ontology/

components of which there occur processes of certain sorts such as: activations of neurons, formation of synapses between cells, and flow of electrons. The corresponding physical components in the patient organism – components which are involved in both mental disease and normal cognitive functioning – are called 'mental functioning related anatomical structures' (Ceusters and Smith, 2010).

Ontology for Pain, Mental Health and Quality of Life (OPMQoL)
The goal of the OPMQoL-project funded by the National Institutes of Health (NIH) is to obtain better insight into the complexity of pain disorders, specifically concerning the assessment of different pain types in the orofacial region, as well as into pain-related disablement and its association with mental health and quality of life (Smith *et al,* 2011). Pain is classified in OPMQoL as an ogms:pathological process. OPMQoL builds also further on OMF because there can be no pain in absence of mental functioning.

Sleep Domain Ontology (SDO)
SDO is a domain-specific application ontology and provides a common framework for merging physiological and clinical data in the area of Sleep Medicine. It was developed as part of the NCRR-funded PhysioMIMI project aimed at developing a federated data integration environment supporting collaborative clinical and translational research using physiological, clinical and genomic data across institutions and databases. It builds on upper-level and reference ontologies such as BFO, FMA, CPR and OGMS. It consists of nearly 1400 classes that extend OGMS terms such as 'clinical finding', 'disorder', 'pathological bodily process', and describes polysomnography data (EEG, ECG, EMG, EOG, SpO2, etc), clinical information, demographics and measurement units. Design patterns in the SDO are leveraged by the PhysioMIMI query interface to generate abstract queries that are translated to database specific queries using database-to-ontology mappings. Released in 2010, it was the first application ontology built on the OGMS framework and is available for download from the BioPortal (Arabandi *et al*, 2010).

Ontology for Newborn Screening and Translational Research (ONSTR)
ONSTR is an application ontology intended to serve as a core component of the NBSDC (Newborn Screening, Follow-up and Translational Research Data Integration Collaborative). Ultimately, ONSTR will be capable of integrating and aggregating many types of data from disparate sources collected during newborn screening, as well as from short-term and long-term follow-up (STFU and LTFU) of patients diagnosed after positive newborn screen. Current ONSTR development efforts focus on representing the ONSTR subdomain for covering follow-up of patients with Inherited Metabolic Disorders (IMD), up to 50 of which are included in the extended state-mandated NBS program in USA.

ONSTR follows OBO Foundry best practices by using BFO as top-level ontology and building class hierarchies under relevant classes imported from OBO Foundry ontologies. OGMS, as an OBO foundry candidate ontology that covers entities involved in clinical encounters and provides higher-level terms used across medical disciplines (e.g. disease, disorder, clinical finding), is

used as a reliable source of importable classes representing the medical domain.

Infectious Disease Ontology (IDO)
IDO is a suite of interoperable ontologies (Cowell and Smith, 2009). The suite consists of a core ontology (IDO Core) covering terms and relations generally relevant to the infectious disease domain, and a set of disease- or pathogen-specific ontologies developed as extensions from the core. The core IDO imports terms such as “disease”, “disorder”, “disease course”, and “treatment” from OGMS, and provides infectious disease-specific terms such as “pathogen”, “vector”, “herd immunity”, “fomite”, “virulence”, “focal infection”, “carrier”, “seroprevalence”, “epidemic”, and “antibiogram”. IDO illustrates how the disease-disorder-disease course structure in OGMS is specialized.

ido:infectious disease is_a ogms:disease
ido:infectious disorder is_a ogms:disorder
ido:infectious disease course is_a ogms:disease course

IDO extension ontologies represent this same framework in even more specialized form. For example, in the Staphylococcus aureus extension (IDO-STAPH) we have:

ido-staph:staphylococcal infectious disease is_a ido:infectious_disease
ido-staph:staphylococcal infectious disorder is_a ido:infectious_disorder
ido-staph:staphylococcal infectious disease course is_a ido:infectious disease course

IDO extensions must have enough terms to model infectious diseases from various perspectives. Staphylococcal infectious diseases can be differentiated by the host anatomical location in which the infection (consisting of an infectious organism population) becomes an infectious disorder (i.e., has colonized and become clinically significant). For this purpose, the Foundational Model of Anatomy (FMA) ontology can be used. Also, the genes and gene products present in particular Staphylococcus aureus isolates have an impact on the type of disease (and disease course) the host organism acquires. For example, using the OGMS framework, along with other OBO Foundry ontologies, it is possible to model the finding that Staphylococcus aureus strains carrying the gene for Panton-Valentine leukocidin are associated with infectious disorders of the lung that are the material basis for highly lethal, necrotizing pneumonia in young, otherwise immunocompetent patients (Gillet *et al*, 2002). These findings speak to the nature of the organism (has_part SOME PVL), the location of the infectious disorder in which it is a constituent (has_location SOME fma:lung), the disease (is_a staphylococcal pneumonia) and the disease course (lethal, necrotizing). The OGMS framework is rich enough to represent many of the important aspects of such findings.

Ontology of Medically Relevant Social Entities (OMRSE)
OMRSE was designed initially to represent entities such as the roles that people have in the healthcare system, gender roles, marriage contracts (to enable capture of data about marital status and history), organizations includ-

ing healthcare organizations, and organizational roles (e.g. outpatient facility). Its relationship to OGMS is that healthcare roles such as physician, patient, and so on are realized in an ogms:healthcare process. These roles are sanctioned by various omrse:organizations.

Neurological Disease Ontology (ND)

ND is an extension of OGMS that provides a set of classes to represent neurological diseases along with their associated signs and symptoms, assessments, diagnoses, and interventions encountered in the course of clinical practice and research [see Jensen et al., 2013, "The Neurological Disease Ontology", Journal of Biomedical Semantics, this issue]. ND is developed with an asserted hierarchy based on the primary mechanism of disease, such as degeneration versus demyelination, and not one of symptomology or location of disorder. Use of OGMS has clarified certain ambiguities that commonly arise, such as recognizing dementia as a syndrome rather than a disease, or that secondary progressive multiple sclerosis is not a separate disease type, but rather a unique part of a disease course. Terms are defined in ND utilizing both natural language and axiomatized definitions that describe and formalize the relations between classes within the ontology itself as well as to external ontologies such as the Gene Ontology, Cell Ontology, Protein Ontology, and Chemical Entities of Biological Interest. In addition, external references are made where possible to related terms in other ontologies, vocabularies and terminologies that attempt to classify neurological diseases, for example NIF-Dysfunction. Initial work on ND is focused on the areas of dementia and Alzheimer's disease, multiple sclerosis, and stroke and cerebrovascular disease. Extensions for additional neurological diseases are planned.

Ontology for Adverse Events (OAE)

OAE builds on (but also deviates from) work conducted in the European ReMINE project in which a strict objective view on adverse events is maintained (He *et al*, 2011; Ceusters *et al*, 2011). The ReMINE ontology for adverse events and OAE are examples of how distinct extension ontologies in the same domain can be compatible with OGMS, yet differ in ontological commitments. The whole OAE ontology is based on the ogms:'pathological bodily process'. OAE defines an adverse event (OAE_0000001) as "an ogms:'pathological bodily process' that occurs after a medical intervention. To be consistent with the current adverse event definition in medical field, an adverse event in OAE does not have to be causally induced by a medical intervention. OAE currently defines a 'causal adverse event' as an adverse event that starts by a medical intervention and is causally induced by the medical intervention. The ReMINE ontology, in contrast, acknowledges only as adverse event what in OAE is called 'causal adverse event'. Many intermediate processes may occur in between a medical intervention and an adverse event outcome. It is a major task for OAE to represent and analyze possible causal relations between an adverse event and a medical intervention.

Adverse Event Reporting Ontology (AERO)

AERO allows for unambiguous data representation and interpretation by encoding guidelines classifying adverse events. AERO is used for automated classification of the patients based on a set of signs and symptoms they present, and the associated clinical findings assessed by their physician in compliance with a selected guideline. It also allows for other guidelines representation and therefore integration with other reporting systems or legacy data. The main goal of AERO is to support large-scale, automated analysis of adverse events following immunization reports from national systems such as the Vaccine Adverse Event Reporting System (VAERS) used in the United States and the Canadian Adverse Events Following Immunization Surveillance System (CAEFISS) in Canada. AERO focuses on accurate representation of clinical guidelines, specifically the Brighton case definitions in the domain of adverse events following immunization. It builds under the ogms:'clinical finding' class which is of information entities that are about medically relevant entities - material entities, qualities, processes, dispositions that are typically localized in an anatomical system or region. AERO is a driving effort for the Ontology of Medically Relevant Entities (OMRE), an OGMS extension, to which it submits all signs and symptoms definitions, as those are not specific to AERO but rather intended to be used by other efforts. Clinical findings described in AERO relate to the medically relevant entities and to the body systems using subproperties of *iao:is about*, a general relation between information and things in the world. Medically relevant entities are to be considered a generalization of symptoms or conditions and are directly related to the patient or part of the patient. For example the entity *omre:low blood pressure* is *localized in* the *cardiovascular system*. Taken together, AERO and OMRE allow classification of adverse events following immunization reports along different axes – by Brighton diagnosis or by system involved for example.

Conclusions

OGMS occupies a niche in an ecosystem which includes medical terminologies, electronic health record metadata, taxonomies, and ontologies of varying scope, longevity, expressivity, and quality. Among these are: SNOMED-CT, UMLS, MeSH, HL7-RIM, Biomedical Research Integrated Domain Group model, (Human) Disease Ontology (DO), Symptom Ontology, BioTop, openEHR Archetypes, Clinical Patient Record Ontology, and the International Classification of Disease (ICD). Despite the expansive coverage and widespread use of some of these resources, we believe OGMS uniquely provides a compact, formal, application-agnostic, and general ontology for clinical medicine that can serve as a gateway for other systems to interoperate with the OBO Foundry. Steps are already being taken, for example in the case of DO, to align the upper level terms with OGMS (Schriml *et al*, 2012).

The development of OGMS has always relied on human clinical input and discussion rather than automated ontology construction or automated mappings to other resources. This is consistent with its scope and the fact that it is not attempting to be a terminology for all of clinical medicine, but rather as an

organizing framework for the construction of such resources. The disorder-disease-disease course distinction presented herein is at the core of this framework and allows for a better alignment between clinical data of various sorts and the underlying reality the data refer to. As such, some have called for using OGMS as resource in clinical data standards (Forsberg, 2012).

Future development on OGMS will include a migration to BFO2 and an increased effort to harmonize the 'diagnosis' and 'clinical finding' branches of the ontology with the Ontology for Biomedical Investigations.

Competing interests

None of the authors report any competing interests.

Authors' contributions

All co-authors contributed equally to this paper.

Acknowledgements

The work described is funded in part by grant 1R01DE021917-01A1 from the National Institute of Dental and Craniofacial Research (NIDCR), and award UL1TR000039 from the National Center for Advancing Translational Science (NCATS), and award AI077706 from the National Institute of Allergy and Infectious Diseases (NIAID). LGC's contributions were additionally funded by a Burroughs Wellcome Fund (BWF) Career Award at the Scientific Interface. The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of the NIDCR, NCATS, NIAID, BWF, or the National Institutes of Health.

References

S. Arabandi, C. Ogbuji, S. Redline, R. Chervin, J. Boero, R. Benca (2010), "Developing a sleep domain ontology (Abstract)", *2010 AMIA Clinical Research Informatics Summit*.

S. Bello, A.P. Davis, T.C. Wiegers, M. E. Dolan, C. Smith, J. Richardson, J. Blake, C. Mattingly, and J. T. Eppig (2011), "Waiting for a Robust Disease Ontology: A Merger of OMIM and MeSH as a Practical Interim Solution", *Proceedings of ICBO2011*: 231-233.

R. R. Brinkman, M. Courtot, D. Derom, J. M. Fostel, Y. He, P. Lord, J. Malone, H. Parkinson, B. Peters, P. Rocca-Serra, A. Ruttenberg, S. Sansone, L. N. Soldatova, C. J. Stoeckert, J. A. Turner, J. Zheng, and the OBI Consortium (2010), "Modelling biomedical experimental processes with OBI", *Journal of Biomedical Semantics*, 1(Suppl 1):S7.

W. Ceusters and B. Smith (2010), “Foundations for a realist ontology of mental disease”, *Journal of Biomedical Semantics*, 1(10).

W. Ceusters, M. Capolupo, G. De Moor, J. Devlies, and B. Smith (2011), “An Evolutionary Approach to Realism-Based Adverse Event Representations”, *Methods of Information in Medicine,* 50(1): 62-73.

L. G. Cowell and B. Smith (2009), “Infectious Disease Ontology”, in Vitali Sintchenko (ed) *Infectious Disease Informatics*: 373-395.

K. Forsberg (2012), “Mind Maps Just Begging for RDF Triples”, http://kerfors.blogspot.se/2012/09/mind-maps-just-begging-for-rdf-triples.html. Retrieved 06/21/2013.

Y. Gillet, B. Issartel, P. Vanhems, J. Fournet, G. Lina, M. Bes, F. Vandenesch, Y. Piémont, N. Brousse, D. Floret and J Etienne (2002), “Association between *Staphylococcus aureus* strains carrying gene for Panton-Valentine leukocidin and highly lethal necrotising pneumonia in young immunocompetent patients”, *Lancet* 359: 753-759.

A. Goldfain, B. Smith, S. Arabandi, M. Brochhausen, and W. R. Hogan (2011) “Vital Sign Ontology”, *Proceedings of the Workshop on Bio-Ontologies*, ISMB, Vienna: 71-74.

A. Goldfain, B. Smith, L. G. Cowell (2010) “Dispositions and the Infectious Disease Ontology”, *Proceedings of FOIS 2010*, IOS Press: 400-413.

Y. He, Z. Xiang, S. Sarntivijai, L. Toldo, and W. Ceusters (2011), “AEO: a realism-based biomedical ontology for the representation of adverse events”, *Adverse Event Representation Workshop, ICBO 2011*.

D. P. Hill, J. A. Blake, J. E. Richardson, and M. Ringwald (2002), “Extension and integration of the gene ontology (GO): combining GO vocabularies with external vocabularies”, *Genome Research* 12: 1982-1991.

R. Mizoguchi, K. Kozakil, H. Kou, Y. Yamagata, T. Imai, K. Waki, and K. Ohe (2011), “River Flow Model of Diseases”, *Proceedings of ICBO2011*: 63-70.

C. Rosse and J. L. V. Mejino (2007), “The Foundational Model of Anatomy Ontology”, in A. Burger, D. Davidson, and R. Badock (eds) *Anatomy Ontologies for Bioinformatics: Principles and Practice Vol 6*: 59-117.

R. H. Scheuermann, W. Ceusters, and B. Smith (2009) “Toward an Ontological Treatment of Disease and Diagnosis”, *Proceedings of the 2009 AMIA Summit on Translational Bioinformatics*: 116-120.

L. M. Schriml, C. Arze, S. Nadendla, Y. W. Chang, M. Mazaitis, V. Felix, G. Feng, and W. A. Kibbe (2012) “Disease Ontology: a backbone for disease semantic integration”, *Nucleic Acids Research*, 40: D940-D946.

S. Schulz, K. Spackman, A. James, C. Cocos, and M. Boeker (2010) “Scalable representations of diseases in biomedical ontologies”, *Journal of Biomedical Semantics*, 2(Suppl 2): S6.

B. Smith, W. Ceusters, A. Kumar, C. Rosse (2005), “On Carcinomas and Other Pathological Entities”, *Comparative and Functional Genomics* 6(7/8): 379-387.

B. Smith, W. Ceusters, B. Klagges, J. Köhler, A. Kumar, J. Lomax, C. Mungall, F. Neuhaus, A. L. Rector, and C. Rosse (2005), “Relations in Biomedical Ontologies”, *Genome Biology* 6(R46).

B. Smith, M. Ashburner, C. Rosse, J. Bard, W. Bug, W. Ceusters, L. J. Goldberg, K. Eilbeck, A. Ireland, C. J. Mungall, The OBI Consortium, Neocles Leontis, P. Rocca-Serra, A. Ruttenberg, S. Sansone, R. H. Scheuermann, N. Shah, P. L. Whetzel, and S. Lewis (2007), “The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration”, *Nature Biotechnology* 25: 1251-1255.

B. Smith, W. Ceusters, L. J. Goldberg, and R. Ohrbach (2011), “Towards an Ontology of Pain”, in Mitsu Okada (ed.) *Proceedings of the Conference on Logic and Ontology*:23-32.

B. D. Walker (2007), “Elite Control of HIV Infection: Implications for Vaccines and Treatments”, *Topics in HIV Medicine*, 15(4): 134-136.

R. Walls, B. Smith, J. Elser, A. Goldfain, D. W. Stevenson, P. Jaiswal (2012), “A plant disease extension of the Infectious Disease Ontology”, *Proceedings of ICBO2012*: .

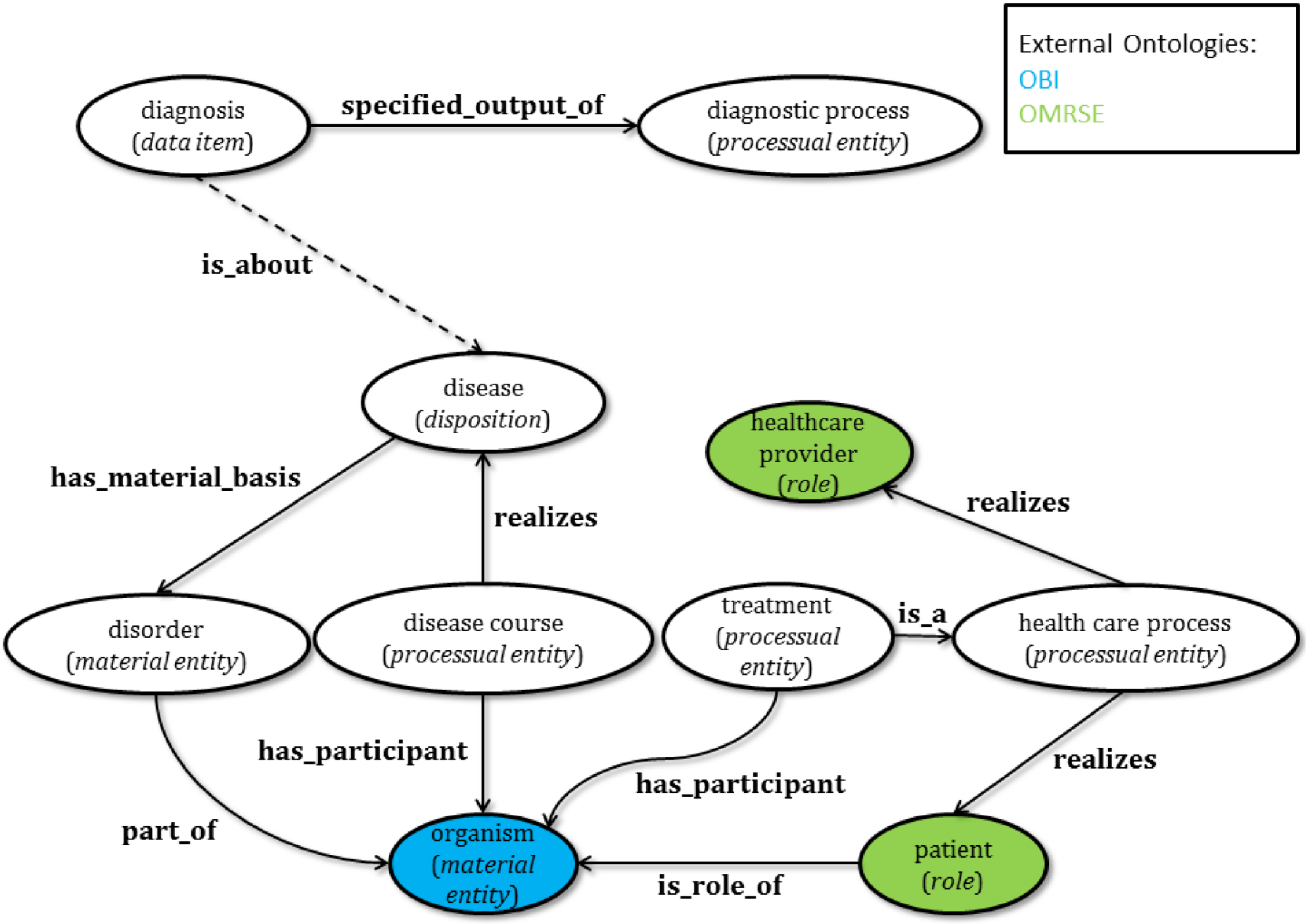
External Ontologies:
OBI
OMRSE
diagnosis
(data item)
specified_output_of
diagnostic process
(processual entity)
is_about
disease
(disposition)
healthcare
provider
(role)
has_material_basis
realizes
realizes
disorder
(material entity)
disease course
(processual entity)
treatment
(processual
entity)
is_a
health care process
(processual entity)
has_participant
has_participant
realizes
part_of
organism
(material
entity)
patient
(role)
is_role_of

Figure 1

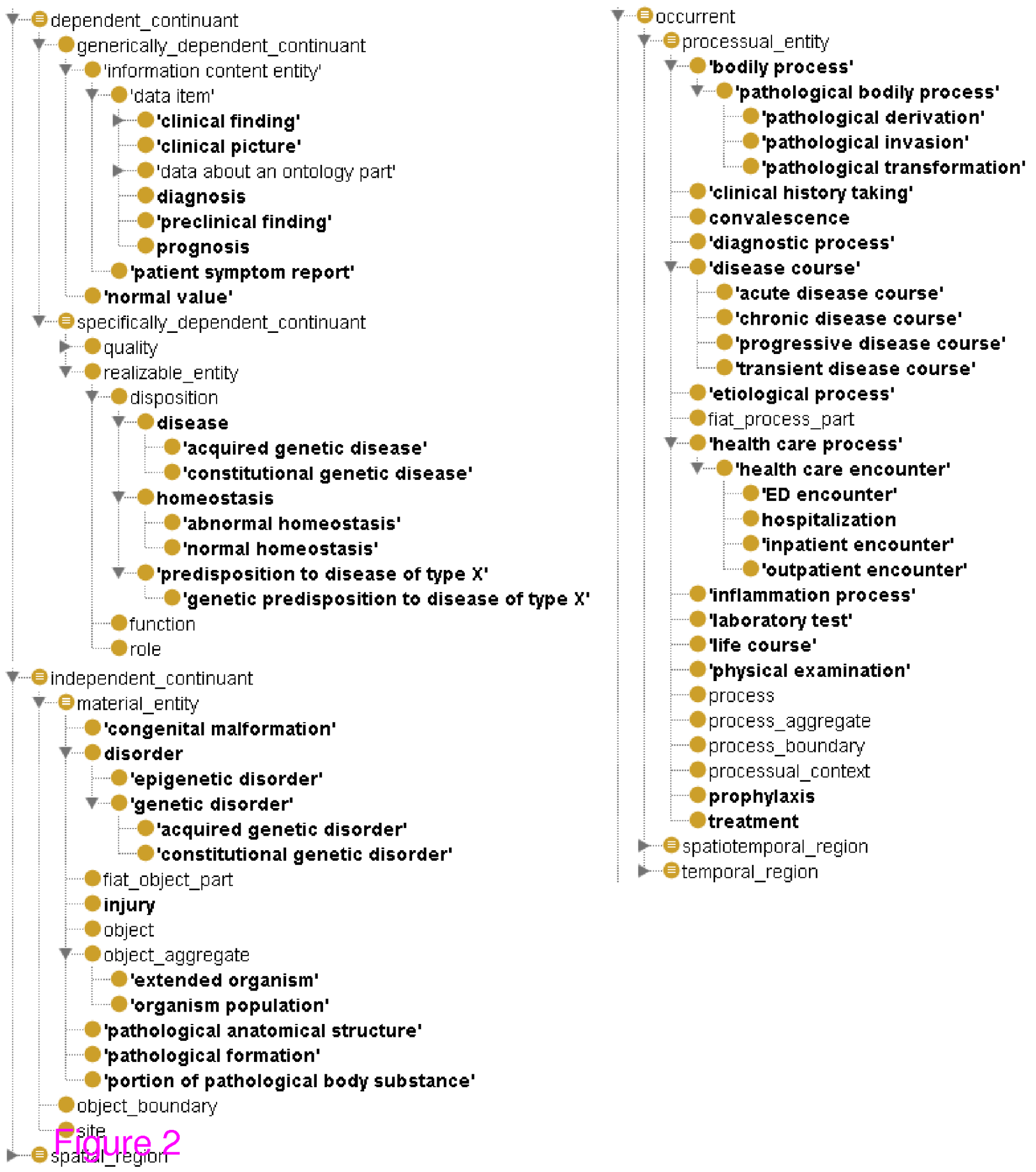
dependent_continuant
generically_dependent_continuant
'information content entity'
'data item'
'clinical finding'
'clinical picture'
'data about an ontology part'
diagnosis
'preclinical finding'
prognosis
'patient symptom report'
'normal value'
specifically_dependent_continuant
quality
realizable_entity
disposition
disease
'acquired genetic disease'
'constitutional genetic disease'
homeostasis
'abnormal homeostasis'
'normal homeostasis'
'predisposition to disease of type X'
'genetic predisposition to disease of type X'
function
role
independent_continuant
material_entity
'congenital malformation'
disorder
'epigenetic disorder'
'genetic disorder'
'acquired genetic disorder'
'constitutional genetic disorder'
fiat_object_part
injury
object
object_aggregate
'extended organism'
'organism population'
'pathological anatomical structure'
'pathological formation'
'portion of pathological body substance'
object_boundary
site
spatial_region
occurrent
processual_entity
'bodily process'
'pathological bodily process'
'pathological derivation'
'pathological invasion'
'pathological transformation'
'clinical history taking'
convalescence
'diagnostic process'
'disease course'
'acute disease course'
'chronic disease course'
'progressive disease course'
'transient disease course'
'etiological process'
fiat_process_part
'health care process'
'health care encounter'
'ED encounter'
hospitalization
'inpatient encounter'
'outpatient encounter'
'inflammation process'
'laboratory test'
'life course'
'physical examination'
process
process_aggregate
process_boundary
processual_context
prophylaxis
treatment
spatiotemporal_region
temporal_region

Figure 2