Applied Informatics

RESEARCH

Open Access

CrossMark

DDO: a diabetes mellitus diagnosis ontology

Shaker El-Sappagh[1*] and Farman Ali[2]

*Correspondence:
shaker_elsapagh@yahoo.com
[1] Information Systems
Department, Faculty
of Computers and Information,
Minia University, El-Minia, Egypt,
Egypt
Full list of author information
is available at the end of the
article

Abstract

Diabetes mellitus is a major cause of morbidity and mortality in humans. Early diagnosis is the first step toward the management of this condition. However, a diagnosis involves several variables, which makes it difficult to arrive at an accurate and timely diagnosis and to construct accurate personalized treatment plans. An electronic health record system requires an integrated decision support capability, and ontologies are rapidly becoming necessary for the design of efficient, reliable, extendable, reusable, and semantically intelligent knowledge bases. In this study, we take the first step in this direction, by designing an OWL2 diabetes diagnosis ontology (DDO). Protégé 5 software was used for the construction of the ontology. DDO is developed within the framework of the basic formal ontology and the ontology for general medical science to represent entities in the domain of diabetes, and it follows the design principles recommended by the Open Biomedical Ontology Foundry. Currently, DDO contains 6444 concepts, 48 properties, 13,551 annotations, and 27,127 axioms. DDO can serve as a diabetes knowledge base and supports automatic reasoning. It represents a major step toward the development of a new generation of patient-centric decision support tools. DDO is available through BioPortal at: http://www.bioportal.bioontology.org/ontologies/DDO.

Keywords: Diabetes mellitus, Ontology engineering, Clinical decision support system, Knowledge reuse, Semantic similarity, Medical informatics

Background

Diabetes mellitus (DM) encompasses a group of metabolic disorders characterized by a chronic hyperglycemic condition, resulting from defects in insulin secretion, insulin action, or both. It is predicted that it will be the seventh leading cause of death by 2030.[1] According to the International Diabetes Federation (IDF),[2] 387 million people worldwide were living with DM in 2014. This number is expected to increase to over 592 million (~10 % of the world's population) by 2035. Undiagnosed cases of DM are thought to reach up to 179 million (Zarkogianni et al. 2015), and 4.9 million deaths were attributed to DM in 2014. In addition to the financial burden of diabetes, diabetes is associated with an increased risk of morbidity and mortality. The American Diabetes Association (Anyanwagu et al. 2015) categorizes diabetes into two main types: type-I and type-II.

---

[1] http://www.who.int/mediacentre/factsheets/fs312/en/.

[2] http://www.idf.org/.

Type-I accounts for 10 % of diabetics. This type is usually diagnosed in children and young adults. Type-II accounts for the remaining 90 % of diabetics.

Diabetes remains a public health problem, despite the availability of several clinical practice guidelines CPGs, research studies, and medications (Zarkogianni et al. 2015). The adherence to large text-based CPGs, even those that are broadly acceptable, is often low. The excessive and increasing number of guidelines makes it hard for physicians to recall, locate, and appropriately apply them. Diagnosis of diabetes is the first and most important step for its prevention, detection, and management (Zarkogianni et al. 2015). Diagnosis describes the process of identifying the syndrome or disease affecting a patient (Bickley and Szilagyi 2012). Bickley and Szilagyi (2012) asserted that for diagnosis, physicians must first gather the indications of a disease, including the past medical history of a patient, symptoms, family history, and physical examinations. Finally, complementary explorations, such as laboratory tests, can increase or decrease the likelihood of a diagnosis. Regarding diabetes, some laboratory tests, such as HbA1c, FPG, BMI, or OGTT, are generally appropriate for directly screening diabetes or pre-diabetes (Anyanwagu et al. 2015). However, such methods can be limited and may be inadequate for detecting high-risk individuals (Buysschaert et al. 2015). For example, HbA1c measurements may be altered in the setting of anemia or renal failure, and HbA1c may remain in the normal range for some patients with mild dysglycemia.

Owing to the asymptomatic nature of the DM disease at the early stage, the patient can be affected by it for 9–12 years prior to being diagnosed (Tripathi and Srivastava 2006; Zarkogianni et al. 2015). At the same time, diabetic retinopathy occurs in three quarters of all people that have diabetes for 15 years, and is the most common cause of blindness. Approximately half of all people with diabetes exhibit some degree of neuropathy (Tripathi and Srivastava 2006). In addition, high urine glucose predisposes patients to urinary tract infections (UTI). As many as 25 % of individuals that are newly diagnosed with diabetes exhibit diabetic retinopathy or microalbuminuria, suggesting that there is an average 7-year gap between the actual onset and the diagnosis of diabetes (Buysschaert et al. 2015). Buysschaert et al. (2015) surveyed the association between diabetes and other complications. As a result, *acute* and *chronic* complications are often present at the time of diagnosis (Tripathi and Srivastava 2006; Anyanwagu et al. 2015; Buysschaert et al. 2015).

The diagnosing of diabetes can be affected by other factors, which can confirm whether diabetes is present regardless of whether laboratory test results are in normal or abnormal ranges (Chen et al. 2012). Some examples of these factors are as follows (Buysschaert et al. 2015): (1) The patient's current complications that are related to diabetes such as *microvascular diseases* (e.g., retinopathy, nephropathy, neuropathy), *macrovascular diseases* (e.g., cerebrovascular, peripheral vascular, coronary artery,), and *infections and other complications (*e.g., psychosocial problems, dental disease, nonalcoholic fatty liver disease (NAFLD), dyslipidemia). For example, Ortiz-Lopez et al. (2012) suggested the usage of NAFLD as a predictor of type-II diabetes. (2) Medications affecting glucose levels such as prednisone, olanzapine, glucocorticoids, chemotherapy agents, antipsychotics, and mood stabilizers. (3) Demographic factors such as gender, age, whether the patient is a smoker, and race. (4) Symptoms such as polydipsia, polyphagia, vision, and polyuria (U.S. Department of Health and Human Services 2016; American

Diabetes Association 2016; Canadian Diabetes Association 2016; National Institute for Health and Care Excellence (NICE) 2016). For example, the American Diabetes Association (ADA) (Anyanwagu et al. 2015) confirmed that the diagnosis of type-II diabetes (T2DM) involves the testing of signs or conditions associated with insulin resistance (e.g., acanthosis nigricans; hypertension; dyslipidemia; polycystic ovarian syndrome; small-for-gestational-age birth weight; or cancer in the liver, pancreas, endometrium, colon/rectum, breast, or bladder). The ADA also confirmed an association between fasting plasma glucose (FPG) results and the presence of retinopathy. The Canadian Diabetes Association (CDA) asserted that the HbA1c laboratory test is affected by many disorders such as hyperbilirubinemia, hypertriglyceridemia, chronic renal failure, chronic liver disease, and hemoglobinopathies. Moreover, some laboratory tests are not compatible with some complications. For example, HbA1c cannot be used for screening patients with homozygous hemoglobinopathies, anemia, EPO therapy, treatments of cancer, hepatitis C, and sever hyperlipidemia.

Many diabetes symptoms are indications of other diseases, such as obesity, which is a major risk factor for DM and cardiovascular diseases (CVD). Tumor markers are used to diagnose pancreatic cancer, but are also an indicator of pancreatic tissue damage, which can be caused by diabetes. In addition, the identified complications that are associated with diabetes help with the creation of a comprehensive care plan for diabetes, containing medication, physical activity, education, and diet programs. There are approximately 100 main complications associated with DM. In 2013, Liu et al. (2013) mined the top 10 complications. Moreover, drugs may induce hyperglycemia through a variety of mechanisms, including alterations in insulin secretion and sensitivity, direct cytotoxic effects on pancreatic cells, and increases in glucose production. These drugs include statins (e.g., fluvastatin, lovastatin, and pravastatin), antibiotics (Fluoroquinolones), atorvastatin (Lipitor), rosuvastatin (Crestor), simvastatin (Zocor), and chemotherapeutic agents (Anyanwagu et al. 2015). However, statins are the most widely prescribed medications for the prevention of cardiovascular events.

Unacceptable morbidity and mortality rates are still recorded in all countries. The effectiveness of the therapy for a disease depends mainly on the level of accuracy and the timing of its diagnosis. The complexities present practicing medicine make conventional quantitative diagnosis approaches inadequate, and hence new techniques are required. In developing countries, such as Egypt, where a population explosion presents a major concern, maintaining manual screening systems for DM patients throughout the country is not an easy task (Yao and Kumar 2013). A system or search engine is required that is able to both detect asymptomatic DM cases in their early stages by using electronic health records (EHR) and support physicians at the point of care for diagnosing specific cases. Clinical decision support systems (CDSS) play a key role in the search for success in the domain of diabetes (Sanchez et al. 2013; Miller et al. 2015). To be effective, CDSS must be proactive, and provide physicians with the right knowledge, in the right form, at the right time (Schreiber 2000).

As previously asserted, the diagnosis of DM diagnosis requires many features. If CDSS is designed as a standalone system, then physicians must enter these features for each patient. This is not feasible, as it would interrupt the clinical workflow of physicians. The solution for this problem is to integrate CDSS as a component of the EHR system

(Ahmadian et al. 2010). As a result, CDSS can automatically collect a patient's features from their profile, even from a distributed environment. However, some issues can prevent advanced data exchange and integration in an EHR environment, including the frequent usage of diverse terminologies to represent the same concept, a lack of logical and machine-readable relationships between different terms, and the lack of a machine-readable and community supported data exchange format (e.g., the OWL format) for the representation of medical data. These obstacles are an obstacle to computer-assisted automated reasoning.

The automation of data collection requires the standardization of the applicable terminology, in order to solve the interoperability and portability problems (Ahmadian et al. 2010). If we treat the categorical features, such as a patient's current medications, diseases, and symptoms, as regular categorical features, then we will lose the associated semantics. For example, if the patient has hypertension, which can affect the decision of a physician regarding a DM diagnosis, the disease may be not stored in EHR under this name. Hypertension has many synonyms and related sub-diseases. For example, for SNOMED CT (SCT)[3] there are 178 other diseases, such as "eclampsia," which are considered as being related to or subtypes of hypertension. String similarity is not suitable for verifying the clinical relationship between these concepts (Agrawal and Elhanan 2014). When CDSS collects a patient's features, it must detect such semantic relationships between the collected diseases, symptoms, and medications in order to mimic the reasoning of doctors. The best solution for handling this challenge is to bind CDSS with ontologies (Gruber 1995; Chen et al. 2012; Sanchez et al. 2013; Rahimi et al. 2014).

Ontologies play an important role in areas such as knowledge management; data integration, exchange, and semantic interoperability; and decision support and reasoning (Button et al. 2013; Suzuki et al. 2015; Mugzach et al. 2015). Ontology can preserve the semantic relationships between its concepts, and hence improve the intelligence of decision support systems. Many CPG languages, such as GLIF, EON, GEM, and Asbru, have exploited this idea to solve the curly braces problem in the Arden syntax language (Ahmadian et al. 2010). Hsu et al. (2015) enhanced the common observational medical outcomes partnership (OMOP)[4] data model by linking its fields with an ontology for an intracranial aneurysm. As a result, they were able to support semantic queries over EHR data. Building ontologies based on existing standard medical ontologies or terminologies, such as SCT, gene ontology (GO), UMLS,[5] and RxNorm, can further improve the interoperability between CDSS and an EHR environment (Agrawal and Elhanan 2014).

The project described in this paper represents the first step toward a comprehensive research program to develop and evaluate new generation informatics interventions for facilitating problem-solving in diabetes management. The first step of this research is to develop a formal organization system of diabetes-specific diagnosis knowledge by using ontology, and to validate this knowledge base for accuracy, completeness, appropriateness, and clarity. We propose Diabetes Mellitus Diagnosis Ontology (DDO), which will be encoded in the OWL 2 format by using the Protégé 5 tool. The core of DDO is

---

[3] http://www.ihtsdo.org/snomed-ct.

[4] http://www.omop.org/.

[5] https://www.nlm.nih.gov/research/umls/.

developed within the framework of the basic formal ontology (BFO)[6] (Arp et al. 2015) and the ontology for general medical sciences (OGMS)[7] (Richard et al. 2009). The ontology concepts are collected from existing standard ontologies, including disease ontology (DOID) (Schrim and Mitraka 2015), symptom ontology (SYMP), RxNorm ontology, SCT terminology, UMLS, units of measurements (UO),[8] and other OBO ontologies (BioPortal repository 2016). The remainder of this paper is organized as follows. "Related work" section describes related work. "Methods" section discusses the methodology for creating DDO. "Results and discussion" section presents our results and a discussion regarding them. Finally, "Conclusion" section presents the conclusion and a discussion of future work.

Related work

The diagnosis of DM has been studied using various techniques (Zarkogianni et al. 2015). One popular method involves using risk score calculators such as the American Diabetes Association calculator,[9] QDiabetes,[10] CANRISK,[11] and Diabetes Australia.[12] Regarding diabetes complications, risk engines include the global diabetes model (GDM) (Brown et al. 2000). However, such calculators are not sufficient for diabetes diagnosis, as they are based on a small number of features, and the features are collected in a crisp format for numerical features and a string format for categorical features. Shankaracharya et al. (2010) identified a list of the most critical features required to diagnose diabetes. Various modeling approaches, along with different combinations of data acquired from heterogeneous sources, can be used to provide a clinically meaningful output (Zarkogianni et al. 2015). A CDSS is required at the point of care (Zarkogianni et al. 2015). The development of such systems depends on the availability of systematic, structured, and computable knowledge bases. Zarkogianni et al. (2015) reviewed recent trends in CDSS regarding diabetes management. Although different approaches and technologies have been proposed since 1960, there remain open gaps that need to be bridged. CDSS must be integrated as a component in an EHR environment (Zhang et al. 2016). However, the full integration of CDSS in modern clinical environments has not yet been fully achieved (Sanchez et al. 2013). One of the most suitable solutions for enhancing CDSS semantics and the interoperability, integrity, and reusability of CDSS knowledge is to utilize standard ontologies (Schrim and Mitraka 2015). Standard ontologies are ontologies built using existing standard medical terminologies or ontologies such as SCT, UMLS, LOINC, GO, DOID, RxNorm, BFO, and OGMS.

Ontology-based CDSS for clinical diagnosis has been extensively investigated in many medical studies (Zhang et al. 2014). For the modeling of diseases, DOID is a general ontology, which collects the concepts associated with many diseases (Schrim and

---

[6] http://www.ifomis.uni-saarland.de/bfo/.

[7] https://www.bioportal.bioontology.org/ontologies/OGMS.

[8] http://www.bioportal.bioontology.org/ontologies/UO.

[9] http://www.diabetes.org/are-you-at-risk/diabetes-risk-test/.

[10] http://www.qdiabetes.org/.

[11] http://www.healthycanadians.gc.ca/diseases-conditions-maladies-affections/disease-maladie/diabetes-diabete/can-risk/index-eng.php.

[12] https://www.diabetesaustralia.com.au/risk-calculator.

Mitraka 2015). Suzuki et al. (2015) asserted that specific ontologies must be developed for specific diseases, because every disease includes knowledge that is specific to that disease. In diabetes, ontologies have been used for every aspect of diabetes management (Chen et al. 2012; Rahimi et al. 2014). For example, Chen et al. (2012) used an ontology to build a recommendation system for drug selection. Rahimi et al. (2014) developed a type 2 diabetes mellitus (T2DM) ontology (DMO), to diagnose and manage patients with diabetes. They proposed an algorithm to query the ePBRN data repository in order to diagnose T2DM. Chalortham et al. (2009) proposed a type-II diabetes ontology. Lin and Sakamoto (2009) defined the ontology of glucose metabolism disorder (OGMD). This is applied to the ontology of geographical regions (OGR) and the ontology of genetic susceptibility factor (OGSF), which describes the genetic susceptibility factors related to diabetes mellitus. This ontology is largely related to diabetes-related complications. The BioMedBridges Diabetes Ontology (http://www.bioportal.bioontology.org/ontologies/DIAB) has been proposed based on SCT taxonomy. This ontology is mainly concerned with diabetes-related phenotypes. Although there are a growing number of ontologies for DM, there has been no attempt to establish a common framework and to collect, organize, and share formal knowledge regarding a complete picture of DM diagnosis.

Chen et al. (2012) introduced an ontology for diabetes medication, and an ontology for the symptoms of patients. These ontologies utilize the Semantic Web Rule Language (SWRL) and Java Expert System Shell (JESS) to determine potential prescriptions for patients. Sherimon et al. (2014) proposed a dynamic adaptive questionnaire ontology for gathering the medical histories of diabetic patients. Hayuhardhika et al. (2014) developed an ontology of the diabetes disease and used a weighted tree similarity algorithm for diagnosis. However, the majority of these studies have not built complete ontologies. In addition, none of these ontologies has been built using a systematic method, and they are not based on globally agreed standard top-level ontologies such as BFO. Moreover, such ontologies are not suitable for an accurate diabetes diagnosis, because none of them integrate the symptoms, laboratory tests, drugs, and diseases that affect a physician's decision. The most critical issue is that the majority of these studies have built a demo ontology and used it as a knowledge base for implementing CDSS by using SWRL and JESS. This technique has not been adopted for many other diseases. For example, Suzuki et al. (2015) have proposed a separate ontology for periodontitis, Button et al. (2013) for the rehabilitation of knee conditions, Mugzach et al. (2015) for autism spectrum disorder, and The Infectious Disease Ontology (IDO)[13] has been proposed for infectious diseases.

We have searched existing literature, including ScienceDirect, IEEE Xplore, Google Scholar, Springer Link, and PubMed for available studies regarding diabetes management ontologies and the topic of CDSS. In addition, ontology libraries have been searched, including BioPortal (BioPortal repository 2016) and the Open Biomedical Ontology (OBO) Foundry (OBO 2016). We have found that there have been many studies regarding disease-specific ontologies and ontology-based CDSS, but there has been no research regarding the building of complete DM diagnosis ontologies that contain all the necessary features and are developed in a systematic way by using widely applied

---

[13] http://www.infectiousdiseaseontology.org.

standard medical and top-level ontologies. All of the existing studies are insufficient. On the other hand, when we searched for other diseases, such as Alzheimer's, we were able to find self-contained ontologies that can be utilized in diverse CDSS applications (Malhotra et al. 2014).

With this motivation, in this paper we aim to build a comprehensive, correct, consistent, and standard OWL 2 diabetes diagnosis ontology. This global ontology follows in full the rules set by the OBO Foundry consortium (Smith et al. 2007). This ontology can serve many purposes, including decision support and interoperability with EHR. We keep in mind two pre-requisites. First, that the ontology must achieve maximum interoperability, and second that it will be amenable to future expansion that adds aspects, such as treatments, that are not included in the initial version.

DDO is an ongoing project, which aims to represent every facet of the diabetes disease accurately, in as much detail as possible. These details include, among others, their clinical presentation, diagnosis, treatment, physical manifestation, course of development, and laboratory tests. DDO is presently in the first stages of development. The fact that the ontology is perfectly created based on well-designed top-level ontologies, i.e., BFO 2 and OGMS, and existing standard ontologies, such as disease ontology and SCT, facilitates reusability and continuous enhancement. Therefore, in the future we will enrich the system with aspects of medication, including appropriate diabetic food, diet, drugs, drug–drug interactions, drug–food interactions, exercises, drug–disease interactions, education programs, and temporal aspects. Moreover, in the future, we will utilize these ontologies to build a diabetes management CDSS, based on semantic and fuzzy reasoning. Diabetes CPG can be added into the ontology in the form of rules by using SWRL, which can be queried using SQWRL or SPARQL. The uncertain nature of diabetes diagnosis can be modeled using fuzzy semantic rules, such as f-SWRL, semantically enhanced regular fuzzy systems, or fuzzy ontology. All of these future targets depend critically on the current stage of building a complete, well-designed, and standard ontology. In 2015, Hempo et al. (2015) attempted to build such an ontology, but they developed a shallow and local ontology that is not standardized.

By the word standardized, we specifically mean that the ontology obtains its concepts from standard ontologies and is built using standard upper-level ontologies. The concentration on standardization is motivated by the fact that the same knowledge can be modeled differently by using different ontology engineering methodologies. Moreover, many engineering errors are present in existing ontologies. For example, Ahmed (2011) proposed an ontology for the management of diabetes type-II, which included the facts that "a *diabetes type 2 complication* is_a *diabetes type 2"* and "*Hyperinsulinemia* is_a *diabetes type 2.*" These axioms are not correct.

Methods

In this section, we describe the construction of DDO. It is based on standard ontologies, to ensure a shared understanding between people and an interoperability between systems (Mohammed and Benlamri 2014). No existing ontology establishes relationships between all of the terms and vocabulary relating to diabetes complications, symptoms, and medications. We can rely on existing terminologies such as SCT. However, such terminologies are huge and involve many limitations regarding ontological design

(López-García et al. 2014). On the contrary, the building of ontologies based on small and precise subsets of these standard ontologies is well documented for many applications (López-García et al. 2014).

For this reason, there is a need for an effective way to create a diabetes ontology to fit the decision support purpose. To enhance interoperability, consistency, sharing, portability, and reusability, our ontology must be extracted from existing standard ontologies and terminologies (Malhotra et al. 2014). An ontology alignment technique can be employed. This entails the idea of combining multiple ontologies to create a single one, where defining the relationships between the concepts of the ontologies forms the new ontology. There are two main approaches to alignment: ontology matching and ontology linking. Ontology matching techniques are used to relate ontologies on the same domain or on partially overlapping domains (Doan et al. 2003), while ontology linking allows elements from distinct ontologies to be coupled using links (Homola and Serafini 2010). Because linked ontologies are disjoint, the ontology linking technique is appropriate for building DDO.

Existing ontologies related to diseases, symptoms, drugs, and other aspects are evaluated for the suitability of their content coverage and depth of knowledge, as well as the potential to support a successful inference. See Fig. 1. Ontology repositories, such as the NCBO BioPortal (BioPortal repository 2016), Protégé Wiki (http://www.protegewiki.stanford.edu/), Swoogle (http://www.swoogle.umbc.edu/), Ontology Lookup Service from the European Bioinformatics Institute (http://www.ebi.ac.uk/ontology-lookup/), and OBO Foundry (OBO 2016), are searched to identify potentially relevant ontologies. To preserve the consistency, interoperability, reusability, and formality of the resulting knowledge, the DDO ontology extends the OGMS ontology and uses BFO 2 as an upper-level overarching ontology (Spear 2015). OGMS is an ontology for the representation of diseases, signs, symptoms, clinical processes, diagnosis, treatment, and outcomes.

In addition, DDO is compliant with the principles of the OBO Foundry (OBO 2016). BFO provides the top-level concepts, which are independent of any specific domain. Using BFO as a foundation, OGMS expands the range of representational units of BFO, to include the concepts of general medical science. This consists of approximately 100 terms that describe the fundamental aspects of medicine such as 'disorder', 'diagnosis', 'disease', 'disease course', 'symptom', and 'syndrome'. DDO seeks to produce definitions in a similar manner within the domain of DM.

Fig. 1 The DDO design methodology
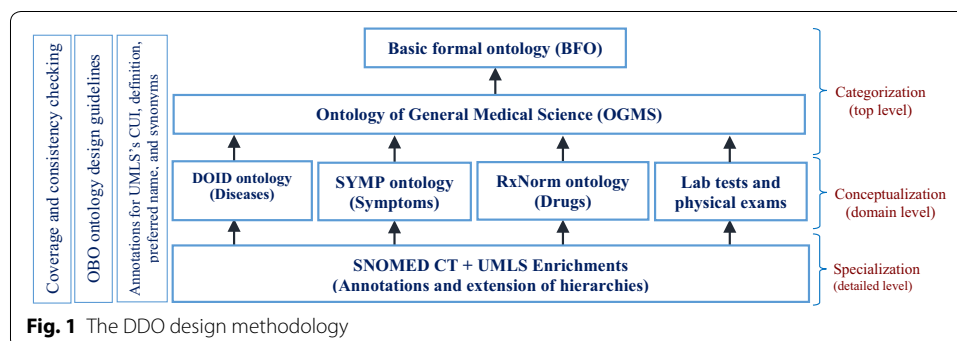
DDO is developed following a top-down and bottom-up approach to catalog relevant entities, relationships, and attributes. We employ a *top-down approach* for creating high-level or generic classes in DDO by analyzing the types of diabetes complications, drugs, symptoms, laboratory tests, and so on that are presented in clinical literature and determining how to classify them within the ontology. Like GO, which contains three main sub-ontologies (biological process, cellular component, and molecular function), DDO focuses on eight main areas (i.e., *disease*, *symptom*, *disorder*, *chemical substance*, *drug*, *demographics*, *physical examination*, and *laboratory tests*). The *bottom-up approach* involves consulting primary research articles, review articles, the texts of CPGs, books, and medical professionals to inform the development of DDO (U.S. Department of Health and Human Services 2016; American Diabetes Association 2016; Canadian Diabetes Association 2016; National Institute for Health and Care Excellence (NICE) 2016). Domain experts provide constructive feedback and guide decision-making regarding controversial material. This approach provides and refines the majority of classes and definitions in DDO. The bottom-up approach leads to the inclusion of new terms in DDO, as well as more detailed classifications of particular concepts. Both approaches are necessary for the completion of the project.

By adhering to the OGMS and BFO frameworks, DDO is consistent and developed collaboratively with other ontologies that use OGMS. Within our diabetes domain, we reuse the most recent, prominent, and standard ontologies, including DOID (Schrim and Mitraka 2015), ontology of glucose metabolism disorder (OGMD),[14] RxNorm 2015 release,[15] and symptom ontology (SYMP),[16] to create a hierarchical set of concepts stored in an OWL 2 ontology. The SCT 2015 release is utilized to enrich all of these ontologies, as described in (Saitwal et al. 2012). In these ontologies, terms are the core concepts and include abstract groups, sets, collections, or types of objects. The words concept, class, and term are used interchangeably. Note that individuals (or instances) are the basic ground level objects of an ontology. However, an ontology need not include any individuals to be more general (La-Ongsri and Roddick 2015). The methodology of building an ontology based on reusing existing ontologies has been employed in many studies such as for PeriO (Suzuki et al. 2015) for periodontitis, TRAK (Button et al. 2013) for rehabilitation of knee conditions, and by Mugzach et al. (2015) for an autism spectrum disorder ontology.

SCT standard terminology represents the most comprehensive medical ontology. It contains around 400,000 active concepts, including their descriptions and relationships. It is owned and distributed around the world by the International Health Terminology Standards Development Organization (IHTSDO).

DOID standard ontology began in 2003 as part of the NUgene project at the Northwestern University. It contains over 8600 known human diseases and 14,600 terms extracted from many ontologies, including SCT 2010, ICD 9, and MeSH. Presently, it has been adopted by the OBO Foundry. DOID is freely available under the Creative Commons Attribution 3.0 Unported License (http://www.creativecommons.org/

---

[14] https://www.bioportal.bioontology.org/ontologies/OGMD.

[15] https://www.nlm.nih.gov/research/umls/rxnorm/.

[16] http://www.symptomontologywiki.igs.umaryland.edu/mediawiki/index.php/Main_Page.

licenses/by/3.0/). The SYMP standard ontology was developed in 2005 at the Institute for Genome Sciences (IGS) at the University of Maryland. Today, it contains more than 900 symptoms. RxNorm is a thesaurus, taxonomy, ontology, nomenclature, and coding system that provides normalized names and concept codes (called RxCUI) for clinical drugs. Its coverage rate is 99.995 % (Saitwal et al. 2012). This ontology was represented in the OWL 2 language as a single hierarchical structure, using the Protégé editor (http://www.protege.stanford.edu). We extract subsets of the participating ontologies for specific diabetes complications, symptoms, and medications. For example, a subset of DOID has been used to create a cancer ontology named *DO_cancer_slim* with only 393 cancer concepts (Wu et al. 2015). SCT concepts can extend some DOID concepts that are related to diabetes complications.

DDO engineering phases

Our methodology for creating DDO consists of five main phases. These phases are executed sequentially, as illustrated in Fig. 2.

Scope and purpose determination

In this phase, the purpose and scope of DDO are determined. DDO has only one specific purpose, which is disease diagnosis. Moreover, it has a specific scope, which is diabetes mellitus.

Requirement specification

This phase determines the aspects that DDO must cover. Requirement specification can be defined informally by a set of competency questions that are defined by a diabetes diagnosis expert, or from the most recent diabetes CPGs. To correctly diagnose diabetes, physicians must know the answers to the following questions:

1. What are the patient's laboratory test results?
2. What are the patient's demographic data?
3. What are the patient's current complications?
4. What are the patient's symptoms?
5. What are the patient's physical examination results?

Fig. 2 Ontology engineering phases
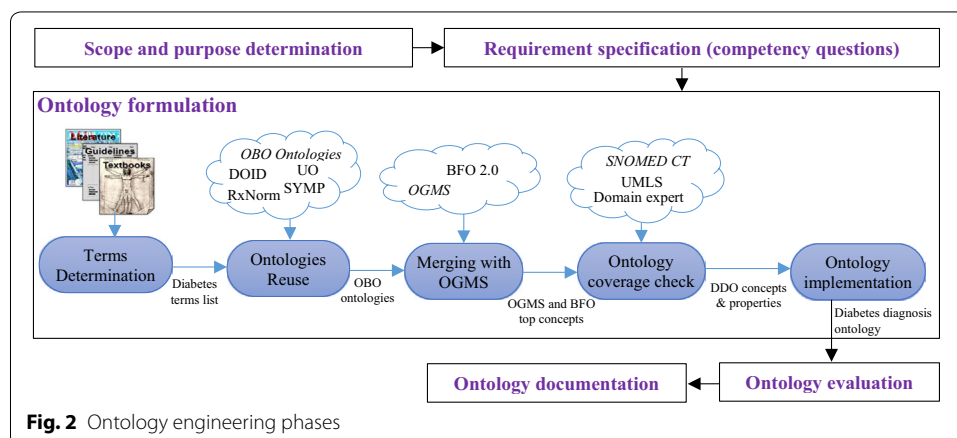
6. What drugs does the patient currently take that may affect glucose level or pancreas function, and what are the chemical ingredients in these drugs?
7. What are the possible diagnoses for diabetic patients?

DDO must have the ability to answer these questions. Moreover, these questions will guide the next steps, including the ontology evaluation.

Ontology formulation

This step consists of five sequential steps to collect related diabetes diagnosis terms and build the full ontology. The following sections discuss these steps in more detail.

Terms determination   The first step in the ontology formulation is the identification of an initial list of terms in the domain of diabetes diagnosis. We survey existing studies concerning diagnosis diabetes, examine EHRs databases, and deeply study the most recent CPGs and books (Ortiz-Lopez et al. 2012; Bos and Agyemang 2013; Blaslov et al. 2015; Buysschaert et al. 2015; U.S. Department of Health and Human Services 2016; American Diabetes Association 2016; Canadian Diabetes Association 2016; National Institute for Health and Care Excellence (NICE) 2016). Moreover, we review existing ontologies, including SYMP, DIAB, DOID, and OGMD, concerning their relations with DM. Meetings with domain experts are scheduled to extract their experience and practice in diagnosing diabetic patients. Our concentration is on factors related to diabetes diagnosis, including complications, symptoms, demographics, laboratory tests, physical examinations, and drugs. We search some databases, such as PubMed, link.Springer, IEEE.Xplore, and ScienceDirect, using keywords related to diabetes diagnosis such as "diabetes diagnosis," "diabetes mellitus laboratory tests," "diabetes mellitus symptoms," "diabetes-related diseases," "diabetes-induced drugs," "diabetes complications," "diabetic foot infection," "diabetes risk factors," "drugs increase blood glucose," "drugs cause hyperglycemia," and "drugs cause hypoglycemia," For example, Xu et al. (2014) recently proposed dRiskKB, a disease–disease and disease–drug risk relationships knowledge base. This study asserted that a total of 35 diseases are associated with diabetes. The SCT browser provided by UMLS Terminology Services (https://www.uts.nlm.nih.gov/snomedctBrowser.html) identifies 497 complications associated with DM. Moreover, searching PubMed for the phrase "is a risk factor for diabetes" returns 87,060 articles that relate diabetes with other diseases, symptoms, and drugs. A total of 28,920 articles assert that "hypertension is a risk factor for diabetes." The list of laboratory tests relevant to diabetes diagnosis is collected from the EHR of Mansura University Hospitals, Mansura, Egypt. Genetic aspects of diabetes mellitus, such as the SCT disorders "609568004|diabetes mellitus due to a genetic defect in beta cell function" or "609569007|diabetes mellitus due to a genetic defect in insulin action," will be handled in future releases of DDO. We have identified eight main knowledge elements that are fundamental to problem-solving in diabetes diagnosis. Table 1 presents examples of information provided for some of these knowledge types.

Reuse of ontologies   In building an ontology, the reuse of ontologies is better than creating one from scratch. Existing medical literature already contains many high-quality ontologies. We searched existing ontology repositories, including NCBO BioPortal for

Table 1 Examples of knowledge types represented in DDO ontology

Knowledge type	Definition	Examples
Laboratory test	A set of tests to measure the level of blood glucose, and others	Blood glucose tests: HbA1C, 2 h OGTT, FPG, RPG
		Hematological profile: prothrombin INR, red cell count, Hbg, hematocrit (PCV), MCV, MCH, MCHC, platelet count, white cell count, basophils, lymphocytes, monocytes, eosinophils
		Kidney function laboratory tests: serum potassium, serum urea, serum uric acid, serum creatinine, serum sodium
		Lipid profile: LDL cholesterol, total cholesterol, triglycerides, HDL cholesterol
		Tumor markers: FERRITIN, AFP serum, CA-125
		urine analysis: chemical examination (protein, blood, bilirubin, glucose, ketones, urobilinogen)
		Microscopic examination (pus, red blood cells (RBcs), crystals)
		Liver function tests: serum albumin, total bilirubin, direct bilirubin, SGOT (AST), SGPT (ALT), alk phosphatase, γ GT, total protein
		Females history: amenorrhea, birth, dysmenorrhea
Disease	Diseases that are often present with diabetes (diabetes complications)	Acute: cardiomyopathy, bacterial infection, ketoacidosis, obesity, ovarian disease, neoplastic disease, etc
		Chronic: nonvascular such as dental disease, kidney disease, skin disease, sexual disease, and gastropathy; vascular such as macrovascular and microvascular
Symptom	The set of common diabetes symptoms	Weight loss, polyuria (increased urination), polydipsia (increased thirst), and polyphagia (increased hunger), weight gain, blurred vision, etc
Drug	The drugs that increase blood glucose, and can induce diabetes either by interfering with insulin production and secretion, reducing the effectiveness of insulin in regulating metabolism, or others	Cardiovascular drugs, CNS drugs, an enzyme inhibitor, a hematologic drug, an immunotherapeutic agent, retinoid, antiallergenic drugs, and others

suitable ontologies, for our eight dimensions (e.g., disease, symptom, drug). DDO fully imports the whole of BFO and OGMS. We implement DDO by using Protégé 5.0.0. Using the OntoFox (Xiang et al. 2010) software program, DDO imports external terms from some other existing ontologies and resources, as illustrated in Table 2. As a principal in OBO Foundry ontologies, an identifier is always bipartite, in the form of *ID-space*: *Local-ID*. The *ID-space* entries represent the identifiers of ontologies that are used, i.e., DDO. The *Local-ID* represents the unique ID of a concept. In our ontology, each concept or property has a unique identifier in the form of DDO_0000000.

The ontology is a global or abstract representation of a domain. Therefore, it does not contain instances or individuals in most cases. In our design, the ontology contains only concepts and properties. When we use the ontology to build a CDSS connected with EHR, ontology instantiation will be performed according to each set of customized patient conditions and characteristics. This customization facilitates personalized

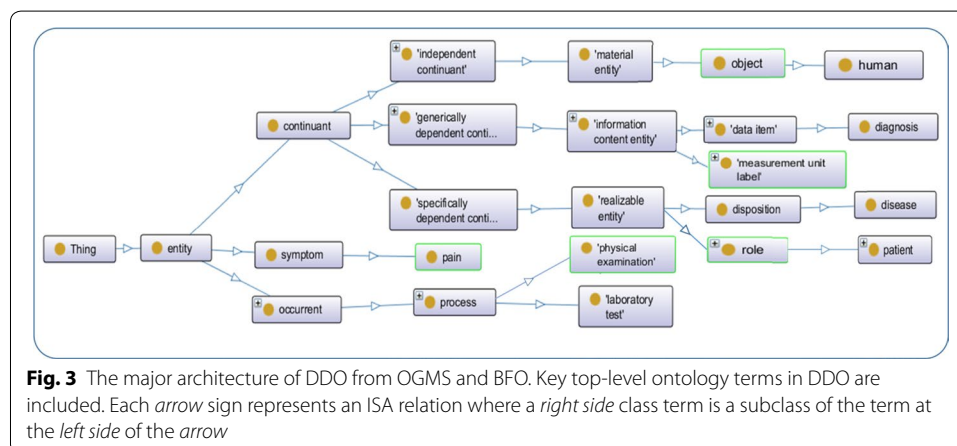**Table 2 External ontologies cross-referenced in DDO**

#	Ontology	Use in DDO	Classes	Object properties	Data properties	Total
Imported ontologies as whole						
1	BFO (Basic Formal Ontology)	Top-level reference ontology	35	0	0	35
2	OGMS (Ontology for General Medical Science)	Mid-level reference ontology	120	0	0	120
Imported from other external ontologies						
1	DOID (Human Disease Ontology)	Externally referenced disease classes	1760	0	0	1760
2	SYMP (Symptoms Ontology)	Select classes for symptoms	80	0	0	80
4	UO (Units of Measurement Ontology)	Laboratory tests units of measurements	37	1	0	38
5	OBO RO (Relation ontology)	Object properties design	0	19	0	19
6	RxNorm ontology	Externally referenced drugs and chemical substances	1042 drug + 1069 substance	0	0	2111
Imported from other external terminologies						
1	SNOMED CT	Select classes to enhance granularity of diseases, drugs, substances, symptoms, etc.	1987 diseases, 72 symptoms, 83 laboratory tests, 1 disease course, 23 physical examinations, 5 diagnoses, 107 disorders, 23 living organisms (*classes*)			2301
2	UMLS	Selection of concepts UMLS CUIs for all concepts				
Total			6444	20	0	6464

diagnosis and treatment. Moreover, the numerical values of laboratory tests have not been added, as they vary from one CPG to another.

Merging with OGMS   The merging of the collected ontologies requires the presence of a unified upper-level ontology. We have selected OGMS to act as the feeder ontology. Building an ontology based on this universal ontology has significant benefits (Button et al. 2013; Suzuki et al. 2015; Mugzach et al. 2015). We must define the top-level concepts in this ontology to insert our specific concepts under them. In Fig. 3, we present the asserted upper-level hierarchy of DDO, which shows how the top-level domain-specific classes are classified under the OGMS classes. The figure illustrates the major architecture of DDO. As shown, all diabetes-related terms are subclasses of terms from higher-level ontologies, including OGMS and BFO 2. Our ontology is distributed in six upper-level classes. The topmost class is *BFO: entity*, which has four main subclasses—*BFO: continuants*, *BFO: occurrence*, *OGMS: symptom*, and *OGMS: sign*, as shown in Fig. 3. Continuants are further classified as *BFO: independent continuant*, *BFO: generally dependent continuant*, and *BFO: specifically dependent continuant*. An independent continuant has two subclasses—*BFO: immaterial entity* and *BFO: material entity*.

Specifically, a dependent continuant has two sub-concepts—*BFO: realizable entity* and *BFO: quality*. *BFO: disposition* is a subclass of *BFO: quality*, which has *OGMS: disease* as a subclass. This disease concept represents a certain disposition to undergo pathological processes that exist in an organism as a result of one or more disorders in that organism. A disease is a dependent continuant consisting of one or more causal chains of clinical disorders appearing in a human body and initiated by at least one disorder (Richard et al. 2009). For a detailed description of all of the BFO classes, the reader is referred to the BFO manual (Spear 2015). The *OGMS: disease* concept is the parent of the *DDO: diabetic complication* concept, which subsumes all sub-tree hierarchies representing diabetes complications. The *OGMS: symptom* concept is the parent of the *DDO: diabetes symptom* concept, which is the patient of all diabetes symptoms.

The *BFO: material entity* concept is the parent of the *DDO: drug* concept, which is the parent of all drugs affecting the level of glucose, as constructed in the drug ontology. A symptom is a bodily feature that a patient observes (2009). On the other hand, the *BFO: process* concept is a subclass of the *BFO: occurrent* concept in OGMS, and the

Fig. 3 The major architecture of DDO from OGMS and BFO. Key top-level ontology terms in DDO are included. Each *arrow* sign represents an ISA relation where a *right side* class term is a subclass of the term at the *left side* of the *arrow*

OGMS: laboratory test concept is a subclass of the *BFO: process* concept. This *OGMS: laboratory test* concept is the parent concept of all sub-tree hierarchies representing diabetes-related laboratory tests, and *OGMS: physical examination* is the parent of all of a patient's physical examinations. Finally, the *OGMS: diagnosis concept* in the path *thing→ entity→ continuant → generally dependent continuant→ information content entity→ data item* is the parent of *DDO: diabetes diagnosis concept* and is used to represent the concepts relating to patient diagnosis. Diagnosis is defined as "a conclusion of an interpretive process that has as input a clinical picture of a given patient and as output an assertion." These selections of top-level concepts are performed based on the definitions of disease, diagnosis, and symptom given by Richard et al. (2009).

In this version of the ontology, only relevant OGMS concepts are used. However, other top-level OGMS terms may be added if necessary in subsequent refinements. Every OWL class has a DDO-specific CUI code, SCT concept ID, preferred name, textual definition, synonyms, and external reference ID defined using the OWL annotation properties *UMLS CUI*, *SCT CID*, *preferred name*, *definition*, *synonym*, and *dBXrefID*, respectively.

Ontology coverage check   The final step is to check the content coverage of the resulting ontology. The collected diabetes-related concepts must be represented in the optimum detailed form. There are two types of enrichment, those where a term cannot be represented in the imported ontology and those where the term can be represented in an abstract form. We refine the hierarchies by adding the absent concepts and comparing the level of granularity of existing concepts in the resulting ontology with SCT. If there are more details present in the latter, we enrich our ontology using these concepts. Concept-based mapping is used to map the remaining terms to the corresponding SCT concepts. To maximize consistency in this mapping, mapping is limited in this study to pre-coordinated concepts only (Kim et al. 2014). In other words, post-coordination is not permitted. LePendu et al. (2011) asserted that SCT concepts are suitable for DOID and symptoms enrichment. Phutthachan et al. (2014) used SCT as a source for a Thai drug ontology. SCT is required for enhancing RxNorm concepts, because RxNorm lacks a classification hierarchy (Saitwal et al. 2012). For example, to enhance the drug concept hierarchy, we use the SCT concept: *373873005|pharmaceutical/biologic product.* To enhance the diabetes complication concept, we use the concept: *64572001|disease.* For symptoms, we use the concept *404684003|clinical finding,* and for laboratory tests we use the concept: *71388002|procedure,* and in particular its sub-concept *122869004|measurements.* Using some SCT browsers, such as cliniClue (http://www.cliniclue.com/cliniclue_xplore), and the UMLS metathesaurus browser (https://www.uts.nlm.nih.gov/metathesaurus.html), we collect all concepts that do not exist in the existing reused ontologies and enhance the granularity of the hierarchies of the reused concepts. We collect all of the parents and the tree of sub-concepts for each existing concept. For each non-existing concept, we use exact matching with SCT terms to determine the equivalent SCT concept. The determined concept is then added to its parents and sub-tree concepts.

Regarding the DOID ontology, DDO imports 1760 disease concepts from the 6598 DOID diseases, which constitutes 26.67 % of the total DOID concepts. DDO contains 3780 disease concepts. Disease concepts imported from DOID represent 46.56 % of the

DDO disease concepts. The remaining 2020 (53.44 %) disease concepts are imported from SCT, either as new concepts that did not exist in DOID, such as *retinopathy* and *coma,* or as sub-concepts of DOID diseases, such as *respiratory system disease* or *acute pancreatitis,* where we add 17 children concepts from SCT.

Regarding the OGMD ontology, the OGMD ontology was originally designed without any standard top-level ontology. Moreover, all concepts of OGMD are directly imported from the SCT concepts *disorder of glucose metabolism|126877002* and *diabetic complication|74627003.* As a result, we decided to ignore this ontology and work directly with SCT, because many of its concepts are represented in an abstract form than in SCT. For example, the concept *hyperglycemia|80394007* in OGMS has one sub-concept, but SCT has 35 sub-concepts.

Regarding the RxNorm Ontology, all drug and chemical substance concepts are imported from RxNorm. DDO imports 1042 drug and 1069 chemical substance concepts from RxNorm. In addition, DDO imports some concepts from SCT to organize RxNorm into hierarchies.

Regarding the SYMP ontology, DDO imports 80 symptoms from the 840 SYMP concepts (9.52 %) that are related to diabetes diagnosis. DDO has 152 symptoms. It enriches the list of symptoms by 72 concepts, either by extending the imported SYMP concepts, such as *hyperesthesia* where we add six other concepts, including *abdominal hyperesthesia*, *allodynia*, and *hyperalgesia*, or by adding new symptom concepts such as *cholestasis*, *foot symptoms*, and *neuroglycopenia*. The concepts imported from SYMP represent 52.63 % of the total number of symptoms.

Ontology implementation   For the implementation of an ontology, ontology development tools and languages are employed. The OWL 2 language is chosen to describe our ontology model. The ontology was implemented using the Protégé 5.0.0 ontology editor, and its consistency was checked using a set of reasoners, including FACT++, Racer, and Pellet. OWL 2 was chosen to formulate the ontology because it provides the maximum expressivity capacity that can be offered, while guaranteeing the total computational capacity. The main components of DDO are presented in Fig. 4. According to the ontology, the task of patient diagnosis involves checking many conditions. The data identified by experts, literature, and guidelines as being of interest for the diagnosis of diabetes are classified into eight groups, described in the ontology as eight general classes: disease, laboratory test, physical examination, demographic, symptom, disorder, drug, and chemical substance. See Fig. 4 for details. The ontology is available at BioPortal for Protégé 5 (http://www.bioportal.bioontology.org/ontologies/DDO/).

Ontology evaluation

Evaluation is an emerging field, and involves several inherent problems (Brank et al. 2005). The ultimate goal of the ontology is to answer the defined competency questions in a complete and accurate way. The proposed competency questions guide the process of collecting diabetes diagnosis terms, and these terms guide the formulation of ontology concepts, properties, and axioms. Domain experts have tested the content coverage of DDO and it fully satisfies the competency questions. In a future study, we will use the ontology to build an interoperable CDSS project. In "Conclusion", we provide an

Fig. 4 Portions of the DDO ontology hierarchies
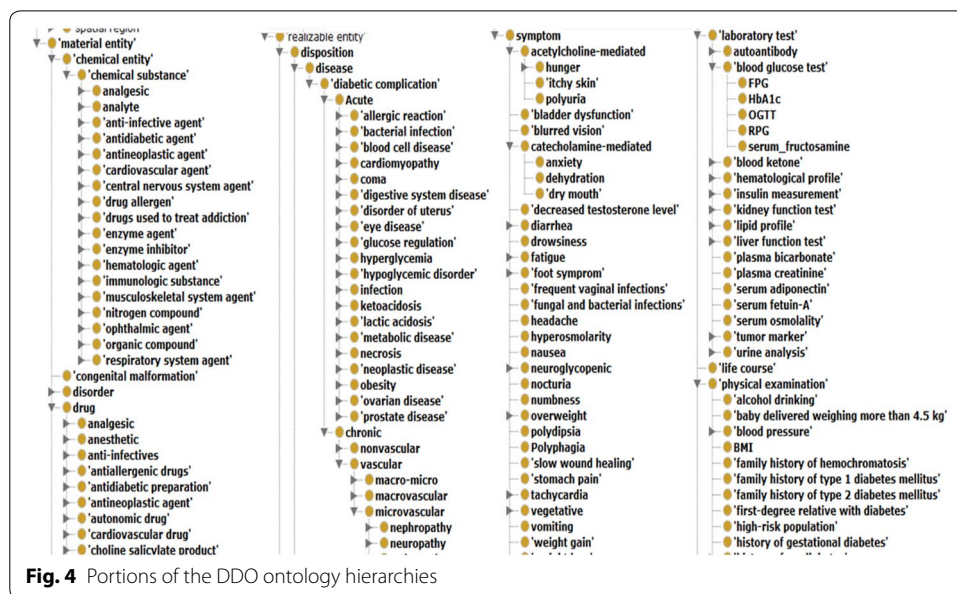
example of the capabilities of DDO for enhancing the semantic similarity between medical concepts and facilitating the interoperability between CDSS and EHR. There are no existing gold standards with which to compare DDO. However, DDO is built based on clinical terms collected mostly from standard CPG documents and concepts from standard ontologies. Moreover, concept properties are defined based on existing standard relationships from the RO ontology, and the taxonomic hierarchy of concepts is defined based on SCT. Finally, our focus will be on the following measures.

Consistency checking This checking describes the *syntactic-level* evaluation. The HermiT (version 1.3.8), Pellet, and FaCT++ reasoners are used with the Protégé 5 editor to check that DDO is consistent and free of errors. They do not reveal any discrepancies regarding this version of the ontology. Moreover, the Ontology Pitfall Scanner! (OOPS!)[17] online tool can help to detect some of the most common pitfalls occurring in the development of ontologies. We run OOPS! on the DDO ontology to ensure that the ontology is free of such pitfalls.

Comparison with existing ontologies DDO has achieved a 100 % content coverage. All existing ontologies related to the DM disease achieve only a limited coverage of the necessary concepts. All diabetes terms collected from a domain expert, CPGs, textbooks, and literature research are defined in DDO. Moreover, the usability of the ontology is high, because it uses standard labels for concepts. Each concept is annotated by a unique ID in the form of DDO_000000, a unique SCT concept ID, a unique UMLS CUID, a unique RxNorm RxCUI, a textual definition, and synonyms. DDO is 100 % annotated with SCT concept IDs. In future releases, the other IDs will be added. These IDs prevent redundancy.

---

17 http://www.oops.linkeddata.es/.

All DDO concepts and synonyms are collected from standard sources, either from BioPortal ontologies, such as RxNorm, DIAB, SYMP, and DOID, or from standard terminologies such as SCT. DDO achieves a 100 % coverage of DIAB ontology concepts. That is, all DIAB concepts are present in DDO. Furthermore, DDO classifies concepts in a meaningful way. For example, all symptoms are sub-concepts of the "*diabetes symptom*" concept, and diabetes complications occur under the "*disease*" concept. On the other hand, 366 of the 375 concepts in DIAB are direct subclasses of the *phenotype* concept and no concept hierarchies exist, which reduces the semantic and inference capabilities of the DIAB ontology. In comparison with DOID, DDO provides domain-specific knowledge about diabetes diagnosis, but DOID is a general ontology. All concepts related to diabetes complications in DOID are collected in DDO, to preserve consistency with existing BioPortal ontologies. The 134 concepts in OGMD are organized into five main concepts regarding diabetes complications such as *hyperglycemia*, *hyperinsulinism*, and *diabetes complication*. However, the ontology is not built using a standard mechanism. It concentrates on one dimension of diabetes (i.e., complications) and achieves a limited coverage, even in the complications dimension.

Class-level representation   DDO is constructed according to the OBO Foundry principles, as an extension of OGMS, which provides a set of general reference classes related to diseases, their patients, and diagnoses. OGMS follows the paradigm of BFO. There are many advantages of using BFO. First, it provides a formal structure for the classification of domain terms. Second, it offers a set of well-defined principles, being known for best ontology practices in the biomedical arena. Third, it helps make DDO interoperable with other ontologies that have the formal structure of BFO. Finally, the use of BFO ensures that the information represented in the ontology is clear, rigorous, and unambiguous as it is expanded through collaborative development. Table 3 outlines some of the high-level classes employed in DDO. The first column lists the formal names the high-level classes,

Table 3 The top-level classes table

Meta-level class	Description	Example subclasses	# of subclasses
DDO:diabetic complication	Subsumes all complications related to diabetes	Acute disease, chronic disease, nephropathy, foot disease, and cardiovascular	3781
DDO:drug	Subsumes all drugs affecting glucose level or pancreas function	Analgesic, antibacterial, antihistamine, autonomic, hematologic, retinoid, antiasthmatic, immunotherapeutic	1042
DDO: laboratory test	Subsumes all of the required laboratory tests to diagnose diabetes	HbA1c, serum ketone, insulin antibodies, basophils count, C-peptide, serum creatinine, total cholesterol, direct bilirubin	83
DDO: physical examination	Subsumes the physical signs that support diagnosis decision	Smoking, physical activity, BMI, blood pressure, thyroid function, waist circumference, alcohol drinking	23
DDO: diabetes symptom	Subsumes all of the diabetes-related symptoms	Cholestasis, diarrhea, drowsiness, edema, emphysema, fatigue, foot symptom, hyperesthesia, nausea, numbness	152

and the corresponding descriptions are listed in the second column. Examples of sub-classes and sample instances derived from each class are listed in the third column, and a count of the number of subclasses is presented in the rightmost column.

Based on SCT, the high-level classes of DDO utilize extensive sub-classing to cover specific abstractions of entities.

Ontology properties　　The *is_a* relationship was used to provide the main taxonomic structure. See Fig. 4 for details. The properties are divided into object properties, which link ontology concepts, and data properties, which link ontology concepts with primitive data types (e.g., integer, string.). Table 4 provides examples of properties of modeled objects, and Table 5 provides the properties of modeled data. To enhance the standardization of DDO, we import 19 properties from the OBO RO ontology, which form 42.2 %

Table 4　Examples of OWL object properties (class-level association) in the DDO

Property	Definition	Inverse of	Domain	Range
has_disposition	Defines the patient's current diseases	disposition_of	Patient	Disease
has_quality	Defines the patient's current symptoms	Quality_of	Patient	Diabetes symptom
has_lab_test	Defines the patient's laboratory tests		Patient	Diabetes laboratory test
has_physical_exami-nation	Defines diabetes associated physical examinations		Patient	Diabetes physical examination
is_about	Links the patient's type of diagnosis with the diabetes disease type		Diabetes diagnosis	Diabetes mellitus
leads_to	Associates a drug or chemical substance with a disorder	result_from	Drug, chemical sub-stance	Disorder
Realizes	Associates a diabetes disease course with a diabetes disease	realized_in	Diabetes disease course	Diabetes mellitus
takes_drug	Determines the drugs taken by the patient		Drug	Patient
has_role	Defines a relationship between human and patient	role_of	Human	Patient
Contains	Defines the chemical entities and drugs	contained_in	Drug	Chemical entity
has_measurement_unit	Defines the measure-ment units of laboratory tests		Diabetes laboratory test	Measurement unit label
has_demographic	Defines patient demographics		Patient	Demographic
has_diagnosis	Defines the patient diagnosis		Patient	Diabetes diagnosis
has_symptom	Is the relationship between a disease and its symptoms	symptom_of	Disease	Symptom
finding_site	Specifies the body site affected by a condition		Disease or disorder	Organ
Total number of object properties				45

Table 5 Data properties of DDO

Property	Definition	Domain	Range
has_boolean_value	Defines the Boolean value of the patient's physical examinations	6 concepts	Boolean
has_qualitative_value	Determines the string value of some physical examination features and assigns a qualitative representation for the patient laboratory tests	4 concepts	String
has_quantitative_Value	Assigns a value for diabetes laboratory tests and some other features	6 concepts	Float
has_risk_level	Summarizes the risk level of having diabetes for a patient by using a linguistic value such as low_risk and high_risk	Patient	5 values
has_risk_percent	Summarizes the risk level of having diabetes for a patient by using a numerical value such as 85 %	Patient	Float
has_symptom	Defines the patient's symptoms	5	String
Total number of data properties			6

of the 45 total object properties. Data properties are mainly used to represent patient laboratory test results.

Ontology documentation

Saitwal et al. (2012) asserted that medical concepts could be mapped to SCT and UMLS concepts. Concept enrichment and the addition of annotations from SCT, UMLS, and Rxnorm can be performed automatically, semi-automatically, or manually. Automatic methods can employ the available files and database for each terminology such as RxNorm's RXCONSO, the UMLS Metathesaurus' MRCONSO table, and the three SCT data files. All of these datasets can be accessed from the UMLS website (https://www.nlm.nih.gov/research/umls/). However, automatic methods require reviewing by domain experts, and this action affects the ontology quality. As a result, this study adopts a manual process for ontology enrichment and annotation. The RxNav browser for RxNorm searching (https://www.rxnav.nlm.nih.gov/) is used to annotate drugs and chemical substances with RxNorm RxCUI. The CliniClue software package is used for SCT browsing (http://www.cliniclue.com/cliniclue_xplore) to collect SCT concept IDs, and the UMLS metathesaurus browser is used for UMLS searching, to obtain the UMLS CUID (https://uts.nlm.nih.gov/metathesaurus.html).

Results and discussion

In this section, we describe the key features of DDO. This ontology is serialized in the OWL 2 format with the Protégé 5.0.0 tool. DDO imports classes from other ontologies by using OntoFox (Xiang et al. 2010), and the enrichment of the ontology from SCT is performed using the CliniClue Xplore[18] SCT browser. The metric data collected using Protégé is presented in Table 6. The first column displays the item, and the second column displays the associated value. The current DDO version incorporates a class count of 6444, organized hierarchically by using the *is_a* relationship, as well as 42 object

---

[18] http://www.cliniclue.com/cliniclue_xplore.

Table 6 DDO ontology metrics

Metric	Value
Class count	6444
Object property count	42
Data property count	6
Axioms count	27,127
Subclass axiom count	6821
Annotation assertion axiom count	13,551
Logical axiom count	6975

properties, six data properties, 13,551 annotations, and 27,127 axioms. We followed the design principles of W3C (http://www.w3.org/TR/swbp-specified-values).

In an attempt to standardize the terminology used to refer to DM concepts and integrate this with other terminological sources, all DDO concepts are annotated with standard concept identifiers, synonyms, and definitions collected from SCT, UMLS, and RxNorm, where such information is available. As a result, DDO concepts are annotated with many types of additional information such as SCT concept IDs, UMLS UCIs, RxNorm CUIs, textual definitions, and alternative terms (synonyms). Synonym annotation is used to specify alternative names of concepts in DDO. This is applied to indicate alternative spelling variants and commonly used acronyms. For example, the term "*obesity*" is equivalent with "*adiposity*" and "*adiposis*." As another example, the term "*angina pectoris*" is equivalent to "*stenocardia*" and "*ischemic heart disease—angina*." At the same time, these concepts have unique identifiers. These representations support an interoperability and integration with CDSS and EHR systems. Lasierra et al. (2013) added physical concepts to represent SCT concept IDs, but this method doubles the number of concepts in the ontology through duplication. In the current version of DDO, every concept is associated with a specific SCT concept ID. However, additional definitions will be added in subsequent versions.

Following the principles of the OBO Foundry, DDO has reused or adapted external OBO Foundry ontologies and candidate ontologies. Figure 5 illustrates the major architecture of DDO, which includes key top-level terms in DDO from BFO and OGMS. As shown in this figure, all diabetes-specific terms are subclasses of terms from higher-level ontologies. All data properties and only selected object properties are displayed in Fig. 5, to enhance the readability of the figure.

All diabetes-related complications are modeled under the *diabetic complication* concept in DDO. This hierarchy subsumes all diseases that are correlated with and go hand in hand with diabetes. This hierarchy contains 3781 diseases, organized under the two main concepts *DDO: acute* and *DDO: chronic*. We have built a large hierarchy of diabetes-related diseases. This large number of concepts is further subdivided into the additional classification of 1274 acute diseases and 2506 chronic diseases. The chronic diseases consist of those that are vascular (1822 concepts) and nonvascular (752 concepts). These diseases are organized into a tree, which facilitates the calculation of the

Fig. 5 A portion of DDO in relation with OGMS and BFO
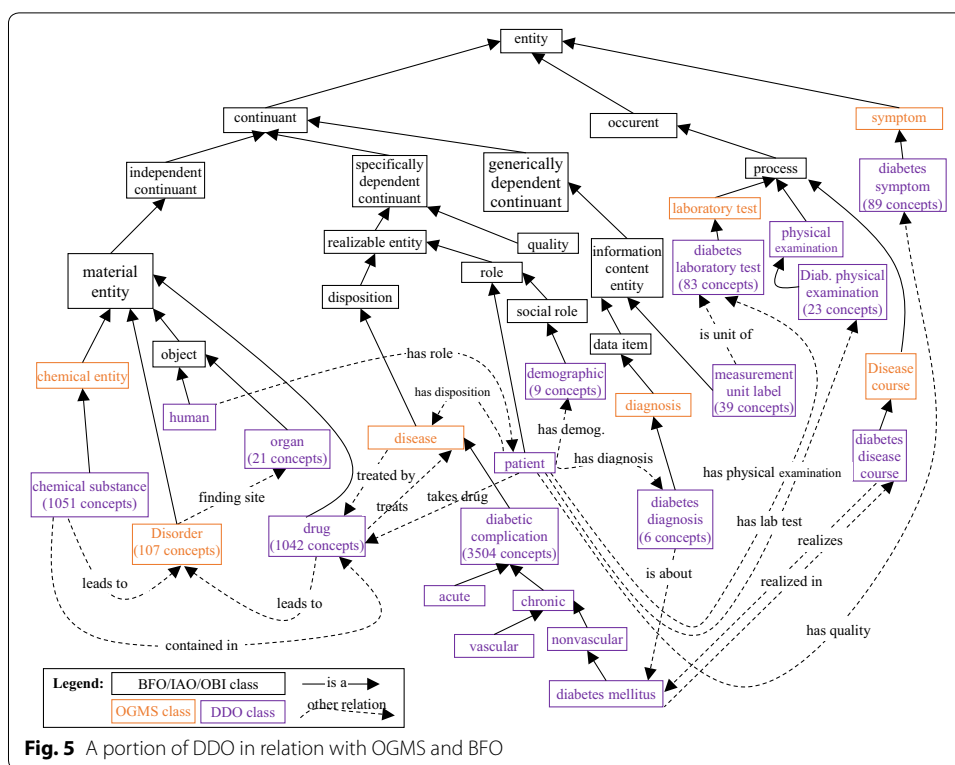
semantic similarity between diseases. In this version of DDO, we concentrate mainly on the *is-a* relationship. Other relationships will be added in subsequent updates such as *after*, *due to*, *associated with,* and *indicates.* All diabetes symptoms are organized under the *DDO: diabetes symptom* concept. There are 152 concepts related to diabetes symptoms. There are 83 concepts for diabetes laboratory tests subsumed by the *DDO: diabetes laboratory test* concept. We add some data properties to model the quantitative and qualitative values of these tests. The physical examinations required for diabetes patients (23 concepts) were collected under the *DDO: diabetes physical examination* concept. All drugs (1042 concepts) that can affect sugar levels or pancreas function are grouped under the *DDO: drug* concept. Moreover, chemical substances (1069) that affect sugar levels are collected in the *DDO: chemical substance* concept. The method of representing drugs and substances is similar to that in Wang et al. (2013). The modeled drugs and chemicals are arranged in a hierarchy, which also enhances the calculation of semantic similarities between the conditions for the current patient and the rules modeled for existing knowledge. Moreover, this will enhance the subsequent implementation of diabetes medication, by facilitating the implementation of drug–drug, drug–disease, and drug–food interactions.

Diabetes diagnosis consists of four main types: *diagnosis of diabetes mellitus type-I, diagnosis of diabetes mellitus type-II, diagnosis of gestational diabetes,* and *diagnosis of pre-diabetes.* These concepts are connected to diabetes mellitus as a disease under the *DDO: diabetes mellitus* concept, which has 104 sub-concepts. Moreover, data properties are modeled to define the risk level associated with a diabetes diagnosis (i.e., *DDO:has_ risk_level* and *DDO_ has_risk_percent*). DDO imports 17 relations from the OBO RO

ontology. RO provides consistent and unambiguous formal definitions of the relations employed in biomedical ontologies (Smith et al. 2005). An individual patient is defined in terms of most of the previously modeled concepts as:

```
Patient ≡ role ∧ (has_level_of_education exactly 1 'level of
education') ∧ (has_demographic  some  demographic) ∧ (has_lab_
test some 'diabetes laboratory test') ∧ (has_quality some 'dia-
betes symptom') ∧ (has_disposition some disease) ∧ (takes_drug
some  drug) ∧ (has_diagnosis  exactly  1  disease) ∧ (has_physi-
cal_examination some 'physical examination') ∧ (has_risk_level
exactly 1 string) ∧ (has_risk_percent exactly 1 float)
```

DDO-based CDSS supports integration with an EHR distributed environment. As shown in Fig. 6, the EHR of a patient can be distributed to multiple hospitals. Different hospitals can use different EHR data models and coding systems. A unified design and coding methodology is required to solve the interoperability problem. These two issues (i.e., syntax and semantic interoperability) have been discussed in many branches of research. For example, we have previously proposed a unified data model based on the HL7 RIM v3 information model (Shaker et al. 2015). Moreover, Marcos et al. (2013) proposed a methodology for modeling EHR by using archetypes.

In addition, different EHR systems may use different coding systems. As a result, a global encoding terminology must be used to unify the clinical meaning of applied concepts. SNOMED CT is the most acceptable terminology for coding EHR data. We have previously proposed an encoding methodology for medical data by using SCT (Shaker and Elmogy 2015). Diabetes mellitus is a chronic disease, and its diagnosis requires personalization according to each individual patient's conditions. At the same time, there are many features involved in making a diagnosis decision, which is a strain on a physician's time. As a result, the integration between CDSS systems and EHR facilitates the collection of patient features from distributed EHRs. The physician must enter the patient's current state only. The additional necessary features are collected from the patient's EHR. DDO supports the semantic collection of such data. For example, the concept of renal disease can be represented differently in different hospitals, i.e., *renal disease*, *nephropathy*, and *nephrosis* (see Fig. 6). These concepts may be coded with different terminologies. SCT supports the mapping to many different terminologies such as

Fig. 6 Integration of DDO-based CDSS in the distributed EHR environment
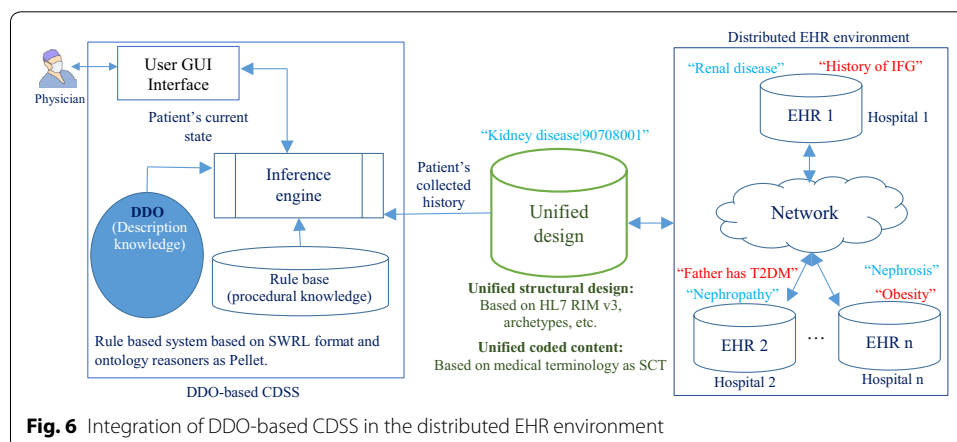
ICD (Unified Medical Language System (UMLS) 2016) and LOINC (Bodenreider 2008). Moreover, SCT supports the mapping to archetypes (García et al. 2012) and HL7 RIM (Rico-Diez and 2013). In addition, concepts in ADDO are furnished with annotations noting the unique UMLS identifiers (UMLS_CUI) and RxNorm identifiers (RxCUI) of the concepts. In this way, a unified representation can be obtained. For example, see *kidney disease|90708001* in Fig. 6. A unified concept (such as *kidney disease|90708001*) can be found directly in the ontology, or a semantic similarity algorithm can be used to find the most similar concept in DDO (Harispe et al. 2014). A physician's knowledge implemented in the form of rules can be coded into DDO by using SWRL rules.[19] For example, consider the following rule:

```
(FPG > 7.0 mmol/L and/or 6.0 < HbA1c < 6.4 % and
OGTT > 11.0 mmol/L and age > 45 and disease = "kidney dis-
ease" and drug = "Phenylephrine" and symptom = symp)→ "Type
2 diabetes".
```

This is formulated by the American Diabetes Association 2015 (American Diabetes Association 2016), and can be represented in SWRL format as follows:

Patient(x), **HbA1c** (l1), **FPG** (l2), **OGTT** (l3), **age** (ag), **has_lab_test** (x, l1), **has_lab_test** (x, l2), **has_lab_test** (x, l3), **has_quantitative_value** (l1, v1), **has_quantitative_value** (l2, v2), **has_quantitative_value** (l3, v3), **has_UoM** (l1, "percent"), **has_UoM** (l2, "mmol/l"), **has_UoM** (l3, "mmol/l"), **has_disposition** (x, d), has_quantitative_Value (ag, av), **Kidney_disease** (d), **takes_drug** (x, dr), **Phenylephrine** (dr), **has_quality** (x, symp), **symptom** (symp), **swrlb:graterthan** (v2, 7.0), **swrlb:graterthan** (v1, 6.0), **swrlb:lessthan** (v1, 6.4), **swrlb:graterthan (av, 45), swrlb:graterthan** (v3, 11.0) -> **has_diagnosis** (x, diag), **type_2_diabetes_mellitus** (diag).

To make a decision regarding a patient, the patient is first instantiated in the ontology (i.e., the ontology ABOX). In this example, the patient data states that patient has a kidney disease, so an exact match is found between the patient feature and the rule. Expert and CPG knowledge are often represented by aggregated concepts such as cardiovascular disease or kidney disease. However, in most cases, when instantiating the ontology concepts from EHR, a patient can be described by other more specific concepts. For example, regarding kidney disease, patients can have *nephritis*, *glomerulonephritis*, or *acute tubular necrosis* diseases. DDO is able to discover these relationships. For example, it contains 157 concepts for kidney disease only. Semantic similarity between the collected patient features and DDO concepts can determine the level of similarity or applicability of a diagnosis rule. We have proposed a semantic similarity measure to calculate the clinical distances between medical concepts (Shaker et al. 2015). DDO supports the calculation of semantic similarity for diseases, drugs, chemical substances, and symptoms. As a result, the resulting decision is semantically intelligent, and any DDO-based CDSS will mimic the thinking of an expert physician. Moreover, the physician is

---

[19] http://www.w3.org/Submission/SWRL/.

not required to enter all of the patient's features, because DDO ontology supports interoperability with EHR systems, and this facilitates the collection of a patient's history. For example, in Fig. 6, the patient's features "*history of IFG*," "*obesity,*" and "*father with type-II diabetes mellitus*" are collected from the distributed EHR systems.

Limitations of the current study

The version of DDO described in this paper has some limitations. First, the ontology concentrates on the diagnosis of diabetes, and no treatments have been discussed. Further studies will add treatment aspects, including medications, foods, education, diet, physical exercise programs, drug–drug interactions, and drug-and-disease interactions. In addition, this study has only focused on the creation of the ontology. Future research could implement a complete CDSS system connected with the EHR environment. The system could use a rule-based reasoning mechanism implemented using SWRL rules supported by the ontology, and use inference engines such as JESS or Pellet. Moreover, the ontology concepts have only been annotated with SCT concept IDs. Future studies could annotate the concepts with UMLS CUIDs, RxNorm RxCUI, synonyms, and textual definitions. Many axioms can be added to the ontology to model the relationships between drugs and ingredients, diseases, diseases and drugs, and diseases and disorders. As a result, future studies will enhance the logic of the ontology with such axioms. Finally, this study only focused on best practices regarding otology development and domain knowledge content. Thus, our future work will also include the design of evaluation studies to assess how successfully DDO supports CDSS of diabetes diagnosis.

Conclusion

The purpose of this study was to develop a theoretically sound and semantically intelligent knowledge base for solving problems related to the diagnosis of diabetes. Such knowledge can enable a new class of patient-centric CDSS that can help physicians to diagnose diabetics quickly and accurately. DDO provides a standard ontology that can support the interoperability between CDSS and healthcare systems. Moreover, it can be used in combination with a rule base to build a rule-based diabetes diagnosis system. The ontology is comprehensive, as it contains all diabetes-related complications, laboratory tests, symptoms, physical exams, demographics, and diagnoses. DDO is the first reported diabetes disease ontology developed to represent different disease aspects in a formal logical format. Future work will concentrate on using DDO to build a CDSS system. Moreover, the ontology annotations for UMLS and RxNorm ids will be completed, and the ontology will be upgraded to cover diabetes treatment, by including concepts and axioms related to diabetes medication and follow-up actions.

Author details
[1] Information Systems Department, Faculty of Computers and Information, Minia University, El-Minia, Egypt, Egypt.
[2] Department of Information and Communication Engineering, Inha University, Incheon, South Korea.

Competing interests
The authors declare that they have no competing interests.

References
Agrawal A, Elhanan G (2014) Contrasting lexical similarity and formal definitions in SNOMED CT: consistency and implications. J Biomed Inform 47:192–198
Ahmadian L, Cornet R, de Keizer N (2010) Facilitating pre-operative assessment guidelines representation using SNOMED CT. J Biomed Inform 43:883–890
Ahmed A (2011) Towards an online diabetes type ii self management system: ontology framework. In: IEEE third international conference on computational intelligence, communication systems and networks. p 37–41
American Diabetes Association (2016) http://www.diabetes.org/. Accessed 10 Jan 2016
Anyanwagu U, Idris I, Donnelly R (2015) Drug-induced diabetes mellitus: evidence for statins and other drugs affecting glucose metabolism. Clin Pharmacol Ther
Arp R, Smith B, Spear A (2015) Building ontologies with basic formal ontology. MIT Press, Cambridge
Bickley L, Szilagyi P (2012) Bates' guide to physical examination and history taking, 8th edn. Lippincott Williams & Wilkins, Philadelphia
BioPortal repository, NCBO (2016) http://www.b.ioportal.bioontology.org/. Accessed 15 Jan 2016
Blaslov K, Bulum T, Knezevic-Cuca J, Duvnjak L (2015) Relationship between autoantibodies combination, metabolic syndrome components and diabetic complications in autoimmune diabetes in adults. Endocrine 48(2):551–556
Bodenreider O (2008) Issues in mapping LOINC laboratory tests to SNOMED CT. AMIA Annu Symp Proc 2008:51–55
Bos M, Agyemang C (2013) Prevalence and complications of diabetes mellitus in Northern Africa, a systematic review. BMC Publ Health 13(1):387
Brank J, Grobelnik M, Mladenic D (2005) A survey of ontology evaluation techniques. In: Proceedings of the conference on data mining and data warehouses (SiKDD 2005). p 166–170
Brown J et al (2000) The global diabetes model user friendly version 3.0. Diab Res Clin Pract 50(3):15–46
Button K, van Deursen R, Soldatova L, Spasic I (2013) TRAK ontology: defining standard care for the rehabilitation of knee conditions. J Biomed Inform 46:615–625
Buysschaert M, Medina J, Bergman M, Shah A, Lonier J (2015) Prediabetes and associated disorders. Endocrine 48(2):371–393
Canadian Diabetes Association (2016) https://www.diabetes.ca/. Accessed 10 Jan 2016
Chalortham N, Buranarach M, Supnithi T (2009) Ontology development for type ii diabetes mellitus clinical support system. In: proceedings 4th international conference on knowledge information and creativity support systems
Chen R, Huang Y, Bau C, Chen S (2012) A recommendation system based on domain ontology and SWRL for anti-diabetic drugs selection. Expert Syst Appl 39:3995–4006
Doan A, Madhavan J, Dhamankar R, Domingos P, Halevy A (2003) Learning to match ontologies on the semantic web. VLDB J Int J Very Large Data Bases Arch 12(4):303–319
El-Sappagh S, Elmogy M (2015) An encoding methodology for medical knowledge using SNOMED CT ontology. J King Saud Univ Comp Inf Sci 28(3):311–329
García M, Allones J, Hernández D, Iglesias M (2012) Semantic similarity-based alignment between clinical archetypes and SNOMED CT: an application to observations. Int J Med Inform 81(8):566–578
Gruber T (1995) Toward principles for the design of ontologies used for knowledge sharing. Int J Hum-Comput Stud 43(5–6):907–928
Harispe S, Sanchez D, Ranwez S, Janaqi S, Montmain J (2014) A framework for unifying ontology-based semantic similarity measures: a study in the biomedical domain. J Biomed Inform 48:38–53
Hayuhardhika W, et al. (2014) Weighted ontology and weighted tree similarity algorithm for diagnosing diabetes mellitus. In: IEEE international conference on computer, control, informatics and its applications. p 267–272
Hempo B, Arch-int N, Arch-int S, Pattarapongsin C (2015) Personalized care recommendation approach for diabetes patients using ontology and SWRL. Inf Sci and Appl. Springer, Berlin, p 959–966
Homola M, Serafini L (2010) Towards formal comparison of ontology linking, mapping and importing. In: Proceeding of 23rd int. workshop on description logics (DL2010), vol 10. p 291–302
Hsu W et al (2015) An integrated, ontology-driven approach to constructing observational databases for research. J Biomed Inform 55:132–142
Kim T, Hardiker N, Coenen A (2014) Inter-terminology mapping of nursing problems. J Biomed Inform 49:213–220

La-Ongsri S, Roddick J (2015) Incorporating ontology-based semantics into conceptual modelling. Inf Syst 52:1–20

Lasierra N, Alesanco A, Guillén S, Garcia J (2013) A three stage ontology-driven solution to provide personalized care to chronic patients at home. J Biomed Inform 46:516–529

LePendu P, Musen M, Shah N (2011) Enabling enrichment analysis with the human disease ontology. J Biomed Inform 44:S31–S38

Lin Y, Sakamoto N (2009) Ontology driven modeling for the knowledge of genetic susceptibility to disease. Kobe J Med Sci 55(3):E53–E66

Liu L, Tang J, Cheng Y, Agrawal A, Liao W, Choudhary A (2013) Mining diabetes complication and treatment patterns for clinical decision support. In: proceedings of the 22nd ACM international conference on conference on information and knowledge management. p 279–288

López-García P, Lependu P, Musen M, Illarramendi A (2014) Cross-domain targeted ontology subsets for annotation: the case of SNOMED CORE and RxNorm. J Biomed Inform 47:105–111

Malhotra A, Younesi E, Gündel M, Müller B, Heneka M, Hofmann-Apitius M (2014) ADO: a disease ontology representing the domain knowledge specific to Alzheimer's disease. Alzheimers Dement 10(2):238–246

Marcos M, Maldonado J, Martínez-Salvador B, Boscá D, Robles M (2013) Interoperability of clinical decision-support systems and electronic health records using archetypes: a case study in clinical trial eligibility. J Biomed Inform 46(4):676–689

Miller A, Moon B, Anders S, Walden R, Brown S, Montella D (2015) Integrating computerized clinical decision support systems into clinical work: a meta-synthesis of qualitative research. Int J Med Inf 84(12):1009–1018

Mohammed O, Benlamri R (2014) Developing a semantic web model for medical differential diagnosis recommendation. J Med Syst 38:79

Mugzach O, Peleg M, Bagley S, Guter S, Cook E, Altman R (2015) An ontology for autism spectrum disorder (ASD) to infer ASD phenotypes from autism diagnostic interview-revised data. J Biomed Inform 56:333–347

National Institute for Health and Care Excellence (NICE) (2016). https://www.nice.org.uk/. Accessed 17 Jan 2016

OBO Foundry repository (2016) http://www.obofoundry.org/. Accessed 12 Jan 2016

Ortiz-Lopez C, Lomonaco R, Orsak B, Finch J, Chang Z, Kochunov V, Hardies J, Cusi K (2012) Prevalence of prediabetes and diabetes and metabolic profile of patients with nonalcoholic fatty liver disease (NAFLD). Diab Care 35(4):873–878

Phutthachan S, Suntisrivaraporn B, Surangsrirat D (2014) A framework for mapping Thai drugs using a pharmaceutical ontology extension of Snomed CT. In: IEEE 11th international joint conference on computer science and software engineering (JCSSE). p 313–318

Rahimi A, Liaw S, Taggart J, Ray P, Yu H (2014) Validating an ontology-based algorithm to identify patients with type 2 diabetes mellitus in electronic health records. Int J Med Inf 83:768–778

Richard et al (2009) Toward an ontological treatment of disease and diagnosis. In: Proceedings of the 2009 AMIA Summit on translational bioinformatics. San Francisco, p 116–120

Rico-Diez A et al (2013) SNOMED CT normal form and HL7 RIM binding to normalize clinical data from cancer trials. In: IEEE 13th international conference on bioinformatics and bioengineering (BIBE). p 1–4

Saitwal H et al (2012) Cross-terminology mapping challenges: a demonstration using medication terminological systems. J Biomed Inform 45:613–625

Sanchez E, Toro C, Artetxe A, Grana M, Sanin C, Szczerbicki E, Carrasco E, Guijarro F (2013) Bridging challenges of clinical decision support systems with a semantic approach, a case study on breast cancer. Pattern Recognit Lett 34:1758–1768

Schreiber G (2000) Knowledge engineering, and management: the CommonKADS methodology. MIT Press, Cambridge

Schrim L, Mitraka E (2015) The disease ontology: fostering interoperability between biological and clinical human disease-related data. Mamm Genome 26:584–589

Shaker El-Sappagh, Elmogy M, Riad A (2015a) A CBR system for diabetes mellitus diagnosis: case-base standard data model. Int J Med Eng Inform 7(3):191–208

Shaker S, Elmogy M, Riad A (2015b) A fuzzy-ontology-oriented case-based reasoning framework for semantic diabetes diagnosis. Artif Intell Med 65(3):179–208

Shankaracharya D et al (2010) Computational intelligence in early diabetes diagnosis: a review. Rev Diab Stud 7(4):252–262

Sherimon P, Vinu P, Krishnan R, Takroni Y, AlKaabi Y, AlFars Y (2014) Adaptive questionnaire ontology in gathering patient medical history in diabetes domain. Proc First Int Conf Adv Data Inf Eng 285:453–460

Smith B et al (2005) Relations in biomedical ontologies. Genome Biol 6:R46

Smith B et al (2007) The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. Nat Biotechnol 25:1251–1255

Spear A (2015) Ontology for the twenty first century: an introduction with recommendations, basic formal ontology (BFO). Institute for Formal Ontology and Medical Information Science (IFOMIS). http://www.ifomis.uni-saarland.de/bfo/documents/manual.pdf

Suzuki A, Takai-Igarashi T, Nakaya J, Tanaka H (2015) Development of an ontology for periodontitis. J Biomed Semant 6:30

Tripathi B, Srivastava A (2006) Diabetes mellitus: complications and therapeutics. Med Sci Monit 12(7):RA130–RA147

U.S. Department of Health & Human Services (2016) National guideline clearinghouse. http://www.guideline.gov/. Accessed 5 Jan 2016

Unified Medical Language System (UMLS) (2016) ICD-9-CM diagnostic codes to SNOMED CT map. https://www.nlm.nih.gov/research/umls/mapping_projects/icd9cm_to_snomedct.html. Accessed 15 Jan 2016

Wang Y, Lin Z, Liu Z, Harris S, Kelly R, Zhang J, Ge W, Chen M, Borlak J, Tong W (2013) A unifying ontology to integrate histological and clinical observations for drug-induced liver injury. Am J Pathol 182(4):1180–1187

Wu T et al (2015) Generating a focused view of disease ontology cancer terms for pan-cancer data integration and analysis. Database (Oxf), vol 32

Xiang Z, Courtot M, Brinkman R, Ruttenberg A, He Y (2010) OntoFox: web-based support for ontology reuse. BMC Res Notes 3:175

Xu R, Li L, Wang Q (2014) dRiskKB: a large-scale disease-disease risk relationship knowledge base constructed from biomedical text. BMC Bioinform 15:105

Yao W, Kumar A (2013) CONFlexFlow: integrating flexible clinical pathways into clinical decision support systems using context and rules. Decis Support Syst 55:499–515

Zarkogianni K, Litsa E, Mitsis K, Wu P, Kaddi C, Cheng C, Wang M, Nikita K (2015) A review of emerging technologies for the management of diabetes mellitus. IEEE Trans Biomed Eng 62(12):2735–2749

Zhang X, Hu B, Ma X, Moore P, Chen J (2014) Ontology driven decision support for the diagnosis of mild cognitive impairment. Comput Methods Prog Biomed 113(3):781–791

Zhang Y, Tian Y, Zhou T, Araki K, Li J (2016) Integrating HL7 RIM and ontology for unified knowledge and data representation in clinical decision support systems. Comput Methods Prog Biomed 123:94–108