



## Management of diabetic patient profiles using ontologies

### *Administración de perfiles de pacientes diabéticos usando ontologías*

Maricela Bravo\*, Darinel González, José A. Reyes-Ortiz,  
Leonardo Sánchez

*Systems Department, Autonomous Metropolitan University, Mexico*

*Received September 14, 2020; accepted November 26, 2020  
Available online November 27, 2020*

#### **Abstract**

In 2019 the International Diabetes Federation estimated that 12.8 million Mexicans had diabetes. The diabetes epidemic ranked second in causes of death in Mexico, a situation that was severely complicated during the second quarter of 2020 with the COVID-19 pandemic. Studies carried out by the Ministry of Public Health showed that the comorbidity of diabetes with COVID-19 has become a risk factor for serious complications, increasing the mortality rate. For this reason, it is necessary to develop personalized information management systems to support medical decision-making considering the specific characteristics of patients in Mexico. Information management of the diabetic patient profile begins with the investigation and registration of the relevant clinical data, data used by the physician to make the diagnosis and determine a personalized treatment. This article reports the development and integration of an ontology model for the management of diabetic patient profiles, incorporating medical ontologies. The results of the evaluation show the feasibility of using this integrated ontology for the management of diabetic patient presenting comorbidities. Likewise, a consistent ontological model is achieved, which complies with extensibility and reusability quality characteristics.

*JEL code: I18, D83, C88, C63, C55*

*Keywords: Medical information management; Ontology modeling; Diabetic patient profile; Decision making for healthcare*

---

\* Corresponding author.

*E-mail address: [mcbc@azc.uam.mx](mailto:mcbc@azc.uam.mx) (M. Bravo).*

*Peer Review under the responsibility of Universidad Nacional Autónoma de México.*

<http://dx.doi.org/10.22201/fca.24488410e.2020.3050>

0186- 1042/©2019 Universidad Nacional Autónoma de México, Facultad de Contaduría y Administración. This is an open access article under the CC BY-NC-SA (<https://creativecommons.org/licenses/by-nc-sa/4.0/>)

## Resumen

En 2019, la Federación Internacional de Diabetes estimó que 12.8 millones de mexicanos tenían diabetes. La epidemia de diabetes ocupó el segundo lugar en causas de muerte en México, una situación que se complicó gravemente durante el segundo trimestre de 2020 con la pandemia de COVID-19. Estudios realizados por la Secretaría de Salud Pública mostraron que la comorbilidad de la diabetes con COVID-19 se ha convertido en un factor de riesgo de graves complicaciones, aumentando la tasa de mortalidad. Por ello, es necesario desarrollar sistemas personalizados de administración de la información que apoyen la toma de decisiones médicas considerando las características específicas de los pacientes en México. La administración de la información del perfil del paciente diabético comienza con la investigación y registro de los datos clínicos relevantes, datos que utiliza el médico para realizar el diagnóstico y determinar un tratamiento personalizado. Este artículo reporta el desarrollo e integración de un modelo ontológico para la administración de información de perfiles de pacientes diabéticos, incorporando ontologías médicas. Los resultados de la evaluación muestran la viabilidad de utilizar esta ontología integrada para el manejo del paciente diabético con comorbilidades. Asimismo, se logra un modelo ontológico consistente, que cumple con las características de calidad de extensibilidad y reutilización.

Código JEL: I18, D83, C88, C63, C55

Palabras clave: Administración de Información Médica; Modelado de ontologías; Perfil de paciente diabético; Toma de decisiones para el cuidado de la salud

## Introduction

The first case of COVID-19 pandemic was reported in Wuhan China on December 19<sup>th</sup>, 2019. COVID-19 is a disease caused by the SARS-CoV-2 virus. During the first weeks of the pandemic, *Diabetes Mellitus* rapidly emerged as the main comorbidity of patients who developed serious complications or death. According with the World Health Organization (WHO), diabetes is a chronic disease that appears when the pancreas does not produce enough insulin or when the body does not effectively use the insulin it produces. The Standards of Medical Care in Diabetes (2019) state that the effect of uncontrolled diabetes is hyperglycemia. Type 2 diabetes has its origin in the body's inability to effectively use insulin, which is often the result of excess weight or physical inactivity. Regarding the data about diabetes in Mexico, in 2019 the International Diabetes Federation<sup>1</sup> estimated that 12.8 million Mexicans had diabetes (IDF, 2019). Regarding causes of death, data reported by INEGI in 2018 showed that diabetes mellitus represents the second cause of mortality in Mexico with 106,525 cases. Similarly, the National Institute of Public Health (INSP) notes that more than 74,000 people die every year due to diabetes and its complications in Mexico. Another important risk factor is *obesity*. INSP indicates that overweight and obesity are health problems precursors of diabetes; in Mexico, 68% of the population over 20 years of age is overweight and obese. According to (Denova-Gutiérrez, 2020), 17.4% of Mexicans diagnosed with COVID-19 had obesity, 14.5% diabetes, 18.9% hypertension and 2.8% had cardiovascular disease.

This research project focuses on Type 2 Diabetes Mellitus (T2DM), which has several therapeutic options. The purpose of this project is to develop an ontological model for the management of patient profiles with T2DM. This ontological model will serve as a broad knowledge base of the clinical information related with the diagnosis and medical treatment of diabetes mellitus and other diseases. The most relevant information considered for the diagnosis and personalized medical treatment is the patient data, physical exercise, body mass index, basal metabolism rate, symptoms, laboratory tests, medication and diseases. The risks faced in the treatment of diabetes derived from the selection of medications are: hyperglycemia, hypoglycemia, kidney disease, retinopathy, gastroparesis, and sexual dysfunction.

With regard to treatment, not all patients with T2DM will respond equally to the same treatment, the reason could be that the doctor does not consider all the specific characteristics of the patient. The treatment that works best for most patients is selected first, even if this treatment is not effective for "some" patients. According to (Lasierra, 2013) there are no identical patients, each requires their own treatment according to their chronic conditions. In fact, it is very

---

<sup>1</sup> <https://idf.org/>

common that patients who have a chronic disease suffer from another chronic condition (comorbidity). Personalized medical treatment is a relevant and emerging approach that involves the design of strategies to prevent, detect, treat and monitor patients individually according to their complete medical profile (Inzucchi, 2012) (Hempo, 2015) (El-Sappagh, 2018).

The use of ontologies for modelling profiles of patients with T2DM will allow the detailed characterization of all the factors involved in diagnosis and management, achieving a high degree of semantic expressiveness by using description logics, and the incorporation of inference rules based on medical knowledge. A revision of existing ontologies and related sources of information was conducted to define the most relevant patient profile characteristics, attributes and semantic relations that need to be included in the ontological model. Additionally, data required for diagnosis and treatment will be included: patient height, weight, age, BMI, Basal Metabolic Rate, daily caloric intake, physical exercise, symptoms, laboratory tests, medications, and disease. In order to evaluate the usefulness and applicability of the ontology model two study cases are presented and the quality design criteria are explained.

The main contribution of this research is the diabetic patient profile ontology, which incorporates knowledge bases from valid medical resources and quality design criteria. The semantic approach derived from ontologies can be effectively exploited in personalized management services according to the patient, since it confers an inference capacity on the information, making use of logic-based languages. The rest of the paper is organized as follows: Section 2 presents a revision of related ontologies with diabetes mellitus, Section 3 describes the ontology development methodology, Section 4 presents the global ontology integration, Section 5 describes the evaluation, and finally in Section 6 conclusions are presented.

## **Revision of Related Ontologies**

The use of ontologies to support diagnosis and treatment of diabetes mellitus is not a new issue, many researchers have investigated the benefits of logic-based reasoning approaches to automate diagnosis and support Clinical Decision Support Systems (CDSS). In this section we present a revision of ontologies that address the representation of health care support for diabetes mellitus.

In (Paganelli, 2007) authors described an ontology-based context model, and a context management middleware to monitor and assist patients at home. The model consists of four ontologies: Patient Personal Domain Ontology, Home Domain Ontology, Alarm Domain Ontology, and Social Context Ontology. Additionally, authors presented a Web application for manual input of biomedical parameters: heart rate frequency, pulse oximetry, systolic and diastolic blood pressure, body temperature, and glycemia.

In (Buranarach, 2011) presented an ontology for Type II diabetes and a clinical support system. The clinical support system consists of an information system and a reminding system. The ontology concepts included are: patient card, person data, family history, signs, symptoms, and health status (diabetes and diabetes complications).

In (Chen, 2012) authors presented a recommendation system for diabetes medication; this recommendation system was built using WSRL rules and Jess Rule engine. The kind of recommendations are related with specific medication treatment depending of various input data: HbA1c levels, safety, tolerance, among others. They also included information about patient tests. However, this recommendation system lacks of other important information for drug recommendation, for instance, disease history, familiar disease history, physical activity, and diet.

(Lasier, 2013) authors described an ontology-based solution to provide personalized care to chronic patients at home. They implemented a method in three stages: ontology design and implementation, ontology application study, and software prototype implementation. Of particular interest is the incorporation of physician's knowledge and clinical guidelines to represent patient profiles, which is a similar approach to the work reported in this paper. The ontology

incorporated biological measurements: weight, height, glucose level, peripheral capillary oxygen saturation (SpO<sub>2</sub>), amount of air forced from lungs per second (FEV<sub>1</sub>), body water, body fat, blood pressure, and pulse rate.

In (Rahimi, 2014) authors presented a validation of the Diabetes Mellitus Ontology (DMO) using real world Electronic Health Records (EHR) data. They evaluated the accuracy of the ontology to execute inferences regarding the detection of the DM2. The sources of information utilized for the experimentation were taken from a literature review, the Australian National Guidelines for Type 2 Diabetes Mellitus, and 908 real world Electronic Health Records (HER).

(Hempo, 2015) authors presented a personalized care recommendation approach for diabetic patients, using an ontology model that consists of three ontologies: one that describes the particularities of patient profiles, a second ontology for the description of complications of diabetes, and a third ontology that describes self-care practices focusing on food and exercise recommendations.

In (El-Sappagh, 2016) authors described the Diabetes mellitus Diagnosis Ontology (DDO), which was constructed following the principles of the Open Biomedical Ontologies (OBO) Foundry; authors also specialized the Basic Formal Ontology (BFO) upper level ontology; and reused the Ontology for General Medical Science (OGMS). The main classes included in DDO are: diabetic complication (disease), drug, laboratory test, physical examination, and diabetes symptom. This ontology model does not consider treatment plans or medicaments, and family history.

Sherimon and Krishnan (2016) describe OntoDiabetic, an ontology-based decision support system to assess the risk factors and generate treatment suggestions for diabetic patients. Authors define a set of inference rules to derive information about the health status of the patient. This ontology model lacks up-to-date information on drugs, diseases, and important relationships between them.

In (El-Sappagh, 2018) authors described the development of Diabetes Mellitus Treatment Ontology (DMTO) following the design principles of the OBO Foundry, reusing the OGMS ontology, and incorporating the DDO previously defined concepts. The novel inclusions in DMTO are: interactions between drugs, interactions between drugs and complications, interactions between drugs and food, and interactions between drugs and exercise. The main goal of DMTO is to support the automation of diabetes treatment process and provide an intelligent and distributed clinical decision support system to be integrated in an Electronic Health Record (EHR):

In (Chen, 2019) authors describe an Ontology-based Model for Diabetic Patients (OMDP). In OMDP authors specialized the concepts from the BFO and reused the OGMS ontology concepts. OMDP consists of the following main definitions: Gene, Laboratory test, Diagnosis, Treatment, Diet, Exercise, Drug, Complications, Physical examination, Symptoms, and Patient information.

The review of related ontologies is done with the main objective of reusing one or more models, either complete or partial. Therefore, the breadth of concepts they cover regarding the diagnosis and treatment of diabetes mellitus must be observed first. From this perspective, we consider that the ontologies DMTO and OMDP are the most complete because they include treatment plans. However, the reuse of a complete model must be carefully decided. DMTO and OMDP are complex models that extend and specialize from BFO and OGMS ontologies. Therefore, these ontologies import several models and handle many naming conventions that require the ontology designer to become familiar with these rules and regulatory principles of the Open Biomedical Ontology (OBO) Foundry.

The main difficulties found in reusing of DMTO and OMDP ontologies is that there are no reports of current applications that use them. This is an indicator of the following design problems: lack of intelligibility, very large and complex models, performance problems related to memory resources for reasoning and inference execution when dealing with hundreds of patient profiles. Literature reporting these ontologies mention nothing about performance measures regarding large volumes of patient data. Therefore, in this work we have decided to carry out an hybrid ontology design which consists of two phases: design and develop from scratch the patient profile model; and select, modularize and

integrate existing medical ontologies that are usable, clear and valid. The integration procedure of existing resources consists of extracting modules in the form of vocabularies and integrating them as lighter modules, thus avoiding performance issues.

## **Ontology Design Method**

Taking as a reference the methodology reported in (Bravo, 2019) we designed the ontology model that is reported in this section. The process of designing an ontology begins by establishing the domain and competence of the ontology.

### *Ontology Requirements Specification*

This ontology model aims at representing the knowledge base for DMT2 diagnosis, and treatment. DMT2 diagnosis and treatment depend mainly on the particular characteristics of the patient, therefore this ontology will consider the acquisition of the patient data. In order to specify the initial requirements of the ontology model, it is necessary to define a list of competency questions that the ontology model should be capable of answering.

Regarding the patient profile, the following competency questions should be answered:

- a) *What is the basic data that characterizes the diabetic patient profile?*
- b) *What is the basal metabolism of the patient?*
- c) *What is the body mass index?*
- d) *How many calories does the patient consume daily?*
- e) *What is the BMI classification of the patient?*
- f) *What type of physical exercise does the patient practice?*
- g) *How many calories does the patient burns per day?*

Regarding diabetes diagnosis, the following competency questions should be answered

- a) *What are the symptoms associated to T2DM?*
- b) *What are the laboratory tests required to confirm a T2DM diagnosis?*
- c) *What are the T2DM complications?*

Regarding diabetes *Clinical Diagnostic*, the following questions should be answered:

- a) *What are the drugs utilized in T2DM treatment?*
- b) *What are the main elements of a treatment plan?*
- c) *What type of exercise can a diabetic treatment include considering the patient characteristics?*
- d) *What type of diet, food and meals can a diabetic treatment include considering the patient preferences?*

There are other important questions that an ontology model of DMT2 should answer; however, in this article we concentrate on the representation of the diabetic patient profile, since this represents a crucial entry point for the entire process of diagnosis and subsequent treatment.

### *Ontology Design*

The process of designing an ontology consists of making decisions about the concepts to be included into the model and the relationships between them. The type of relationships that are to be implemented are hierarchical relations (is-a) and semantic relations (object and data properties). Based on the competency questions specified above, and the revision of related ontologies, the following terms were identified as the core concepts involved in diagnosis and treatment of DMT2:

Patient Profile, Demographic Data, Physical Activity, BMI Classification, Disease, Drug (Medication), Lab Test, and Clinical Diagnosis among others. Figure 1 shows the main concepts included in the ontology model.

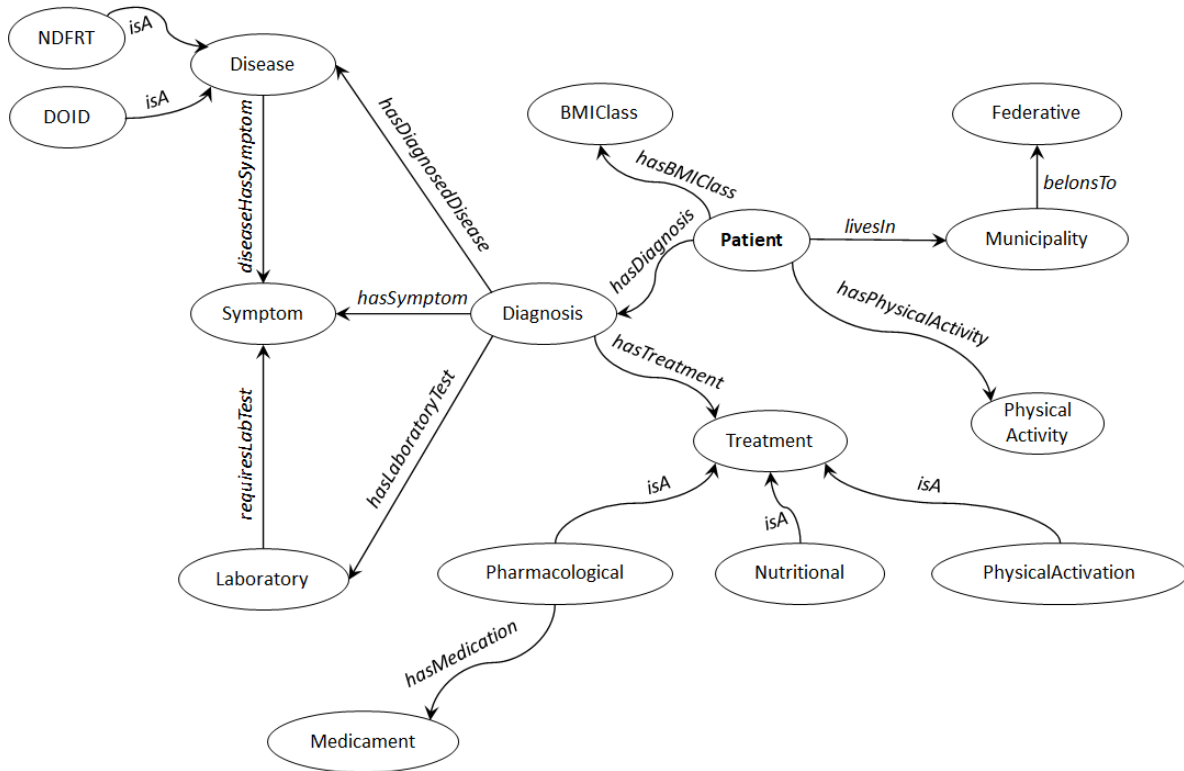


Figure 1. Ontology model to represent *Patient* profiles.  
 Source: Author's own

**Patient** concept was designed considering the important data (without private personal data such as name, address or phone number) to integrate a basic clinical record of the patient. The personal data of the patients is not stored, only the data necessary for the establishment of the state of health are obtained with the express authorization of the participants, the identity of the patients is never known. However, it is necessary to have a way to identify the patient and his / her medical and biomedical data, therefore a username and password are requested as identifiers to access to the system. During the process of deciding which were the most important characteristics to represent a patient, the need to include the data that allowed calculating the metabolism data was observed. It is relevant for the development of this project to calculate the basal metabolic rate and the body mass index.

**Patient** concept consists of the following data type properties: *hasWeight*, *hasHeight*, *hasGender*, *hasBirthDate*, *hasNumberOfChildren*, *hasBloodType*, *bodyMassIndex*, *numberCigarettesPerDay*, *basalMetabolism*, where *basalMetabolism* and *bodyMassIndex* are two data properties which are calculated using inference rules. Figure 2 and Table 1 details the data type properties of the *Patient* class.

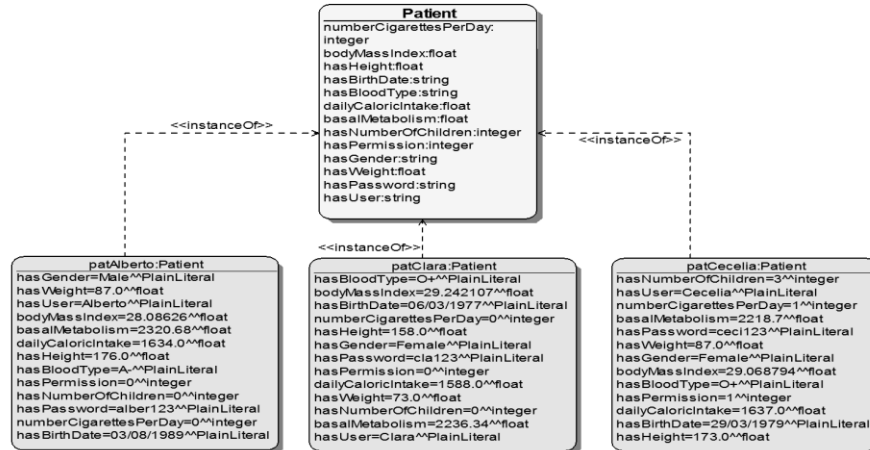


Figure 2. Ontology model to represent *Patient* profiles.  
 Source: Author's own

Table 1  
 Details of the data properties of the *Patient* concept

Data property	Domain	Range	Description
<i>hasUser</i>	Patient	xsd:string	Represents a user identification to access the system. This data attribute is generated when the patient registers for the first time.
<i>hasPassword</i>	Patient	xsd:string	This is an access password created by the user when registered for first time.
<i>hasWeight</i>	Patient	xsd:float	This data type is used to specify the weight in kilograms of the patient during registration.
<i>hasHeight</i>	Patient	xsd:float	This data type is used to specify the height in meters of the patient during registration.
<i>hasGender</i>	Patient	xsd:string	This data type is used to specify the gender of the patient.
<i>hasBirthDate</i>	Patient	xsd:string	This data type is used to specify the date of birth of the patient.
<i>hasBloodType</i>	Patient	xsd:string	This data type is used to specify the type of blood of the patient.
<i>hasNumberOfChildren</i>	Patient	xsd:integer	This data type is used to specify the number of children of the patient.
<i>numberCigarettesPerDay</i>	Patient	xsd:integer	This data type is used to specify the number of cigarettes smoked per day of the patient.
<i>dailyCaloricIntake</i>	Patient	xsd:float	This type of data is used to establish the number of calories that the patient ingests in total per day.
<i>bodyMassIndex</i>	Patient	xsd:float	This data type is calculated using inference rules to represent the weight of a person in kilograms divided by the square of height in meters.
<i>basalMetabolism</i>	Patient	xsd:float	This data type represents the rate of energy required for the body.

Source: Author's own

According with the Center for Disease Control and Prevention (CDC)<sup>2</sup> the Body Mass Index (BMI) represents the weight of a person in kilograms divided by the square of height in meters. Likewise, the CDC indicates that a high BMI can be an indicator of high body fatness. Therefore, BMI calculation can be used to filter out categories of body weight that may lead to possible health problems. In (Pelley, 2012) the Basal Metabolic Rate (BMR) is defined as the rate of energy expenditure of a person at rest; it eliminates the variable effect of physical activity.

The **BMIClassification** concept was defined based on the World Health Organization (WHO)<sup>3</sup> definition, with the objective of identifying the patient BMI class of body weight. **BMIClassification** is defined with three data type attributes: *hasMinValue*, *hasMaxValue* and *hasNameBMI*. Table 2 shows the particular ranges of each individual. Figure 3 shows the semantic relationship that enables the correlation of each *Patient* individual with its weight classification.

Table 2  
 Classification of the weight status of patients according to the BMI

hasNameBMI	hasMinValue	hasMaxValue	WHO classification
<i>UnderWeight</i>		18.49f	Under weight
<i>Normal</i>	18.5f	24.9f	Normal weight
<i>OverWeight1</i>	25.0f	26.9f	Over weight
<i>OverWeight2</i>	27.0f	29.9f	Over weight
<i>Obesity1</i>	30.0f	34.9f	Class I obesity
<i>Obesity2</i>	35.0f	39.9f	Any BMI >= 35 is severe obesity
<i>Obesity3</i>	40.0f	49.9f	Any BMI >= 40 is morbid obesity
<i>Obesity4</i>	50.0f		Any BMI >= 50 is super obese

Source: Author's own

During *Patient* registration, the *basalMetabolism* and *bodyMassIndex* values are calculated, also a semantic relation is defined using the *hasBMIClassification* object property to associate every patient with his weight class.

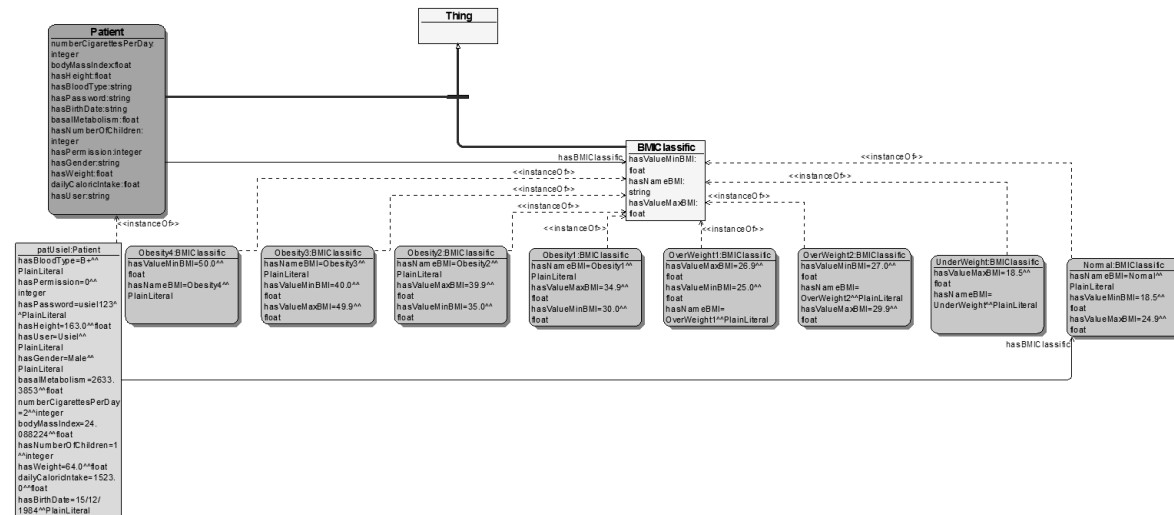


Figure 3. Semantic relationships between *Patient* and *BMIClassification*  
 Source: Author's own

According with a position statement of the American Diabetes Association (Colberg et al., 2016) adoption and maintenance of physical activity is a relevant aspect for the prevention and treatment of diabetes mellitus. Therefore, it is important to obtain this information from the patient during registration. For the definition of the concept **PhysicalActivity** the following data were considered: the frequency and type of exercise that is performed. For the type of exercise we

<sup>2</sup> <https://www.cdc.gov/healthyweight/assessing/bmi/index.html>

<sup>3</sup> [https://www.who.int/health-topics/obesity#tab=tab\\_1](https://www.who.int/health-topics/obesity#tab=tab_1)



defined a specific list of physical activities that the patient can perform routinely, and the approximate number of calories per hour associated, so that it is possible to make an approximate calculation of the number of calories burned. Figure 4 shows the semantic relationship used to define the type of physical activity that the patient performs on a periodical basis. The physical activity also defines a numerical value to represent the number of calories that are burn per hour.

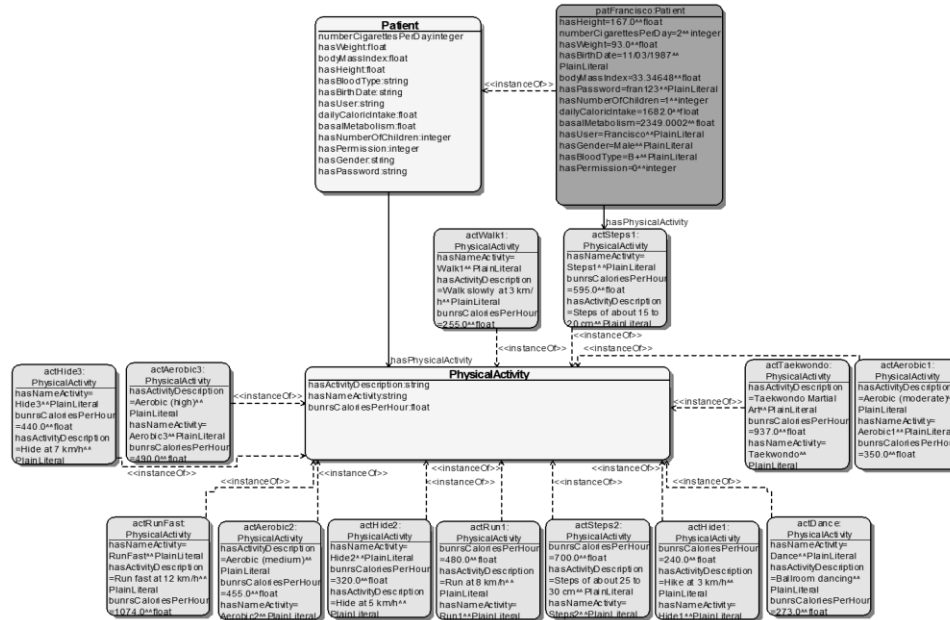


Figure 4. Semantic relationships between *Patient* and *PhysicalActivity*  
 Source: Author's own

The purpose of this research project is to develop a comprehensive infrastructure to facilitate the integration and development of applications for remote monitoring of diabetic patients. In this sense, it is important to study other factors that affect DMT2. In (Ansari, 2019) authors presented a conceptual model of the aspects for the self-management of diabetes, which includes: socio-demographic characteristics, such as age, gender, education; behavioral and psychological characteristics: monitoring of blood sugar, adherence to diet, and physical activity; barriers to self-management: lack of knowledge, self-confidence, financial, and family support; and cultural characteristics: cultural beliefs, and dietary preferences. This article describes a prototype that does not cover exhaustively all the aforementioned aspects; however, the model is possible to be augmented easily. For now, the information regarding the state, province or municipality where the patient was born and lives are included. Figure 5 shows the semantic relationship between the Patient and the Municipality.

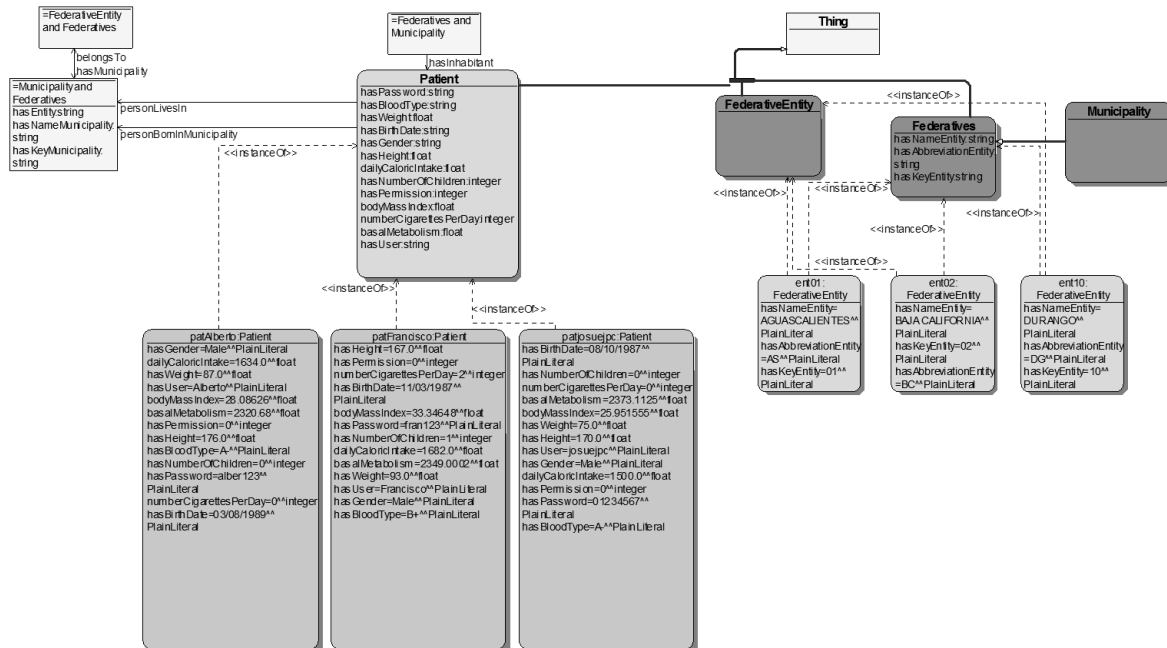


Figure 5. Semantic relationships between *Patient* and *Municipality*  
 Source: Author's own

The semantic relationship of *personBornInMunicipality* is used to describe the specific municipality where the patient was born, also *personLivesIn* is used to indicate the current address of the patient. This information will be used to correlate the patient profile with his social context.

### Integrate Medical Ontologies

Based on the list of required concepts related with the diagnosis and treatment of diabetes, a search was made for existing ontological resources to be reused in this model. Specifically, the following concepts were searched: disease, drug or medication, symptom, and laboratory tests. A good repository of biomedical ontologies is *BioPortal*<sup>4</sup>, which provides search mechanisms to find and explore among a large collection of biomedical ontologies. We executed the search of the afore-mentioned concepts, Table 3 shows the list of ontologies found for each concept.

<sup>4</sup> <http://bioportal.bioontology.org/>

Table 3  
 List of ontologies found using the *BioPortal*

Ontology	General Description	URL	Institutions involved
SNOMEDCT	Consist of a large collection of cross referred medical vocabularies aiming at supporting the clinical recording of patient data.	<a href="https://bioportal.bioontology.org/ontologies/SNOMEDCT">https://bioportal.bioontology.org/ontologies/SNOMEDCT</a>	Published by the College of American Pathologists (CAP). Currently the intellectual property of SNOMED belongs to International Health Terminology Standards Development Organisation (IHTSDO).
DOID	Is a standardized ontology which provides descriptions of human disease terms.	<a href="https://bioportal.bioontology.org/ontologies/DOID">https://bioportal.bioontology.org/ontologies/DOID</a>	Developed at the Institute for Genome Sciences and the University of Maryland School of Medicine.
NDF-RT	Is the National Drug file reference terminology	<a href="https://bioportal.bioontology.org/ontologies/NDFRT">https://bioportal.bioontology.org/ontologies/NDFRT</a>	Produced by the U.S. Department of Veterans Affairs, Veterans Health Administration (VHA).
DRON	The Drug Ontology.	<a href="https://bioportal.bioontology.org/ontologies/DRON">https://bioportal.bioontology.org/ontologies/DRON</a>	Division of Biomedical Informatics, University of Arkansas for Medical Sciences.
SYMP	The Symptom ontology.	<a href="https://bioportal.bioontology.org/ontologies/SYMP">https://bioportal.bioontology.org/ontologies/SYMP</a>	Institute for Genome Sciences (IGS) at the University of Maryland.
CSSO	An ontology for describing clinical signs and symptoms.	<a href="https://bioportal.bioontology.org/ontologies/CSSO">https://bioportal.bioontology.org/ontologies/CSSO</a>	The National Bioscience Database Center (NBDC) is the organization for life sciences databases in Japan. To promote the creation of new knowledge, we conduct research and development and provide services that contribute to data sharing and integration.
LOINC	The international standard for identifying health measurements, observations, and documents.	<a href="https://bioportal.bioontology.org/ontologies/LOINC">https://bioportal.bioontology.org/ontologies/LOINC</a>	The LOINC vocabulary is developed and maintained by the Regenstrief Institute. The LOINC project was initiated in 1994.
LABO	Is an ontology that describes laboratory tests prescriptions and reporting documents.	<a href="https://bioportal.bioontology.org/ontologies/LABO">https://bioportal.bioontology.org/ontologies/LABO</a>	Institut de Recherche en Informatique de Toulouse (IRIT), CNRS, France. Groupe de Recherche Interdisciplinaire en Informatique de la Santé (GRIIS), Université de Sherbrooke, Quebec, Canada.

Source: Author's own

Search results returned a large list of ontologies that matched the concepts; therefore the selection of the ontologies was made based on the intelligibility and usability criteria. That is, that the ontologies that were clear, legible and usable were selected. In this section we describe the specific ontologies selected and integrated into the Patient Profile ontology model.

a) The National Drug File – Reference Terminology (**NDF-RT**) is a controlled medical terminology developed by the Department of Veterans Affairs Veterans Health Administration (VHA). The NDF-RT provides a formal model to describe *medications*, includes the FDA Established Pharmacologic Class(EPC), a set of Mechanisms of Action (MoA), Physiological Effects, Therapeutic Categories, and cross references to other important vocabularies such as the RxNorm<sup>5</sup> provided by the National Library of Medicine (see Figure 6).

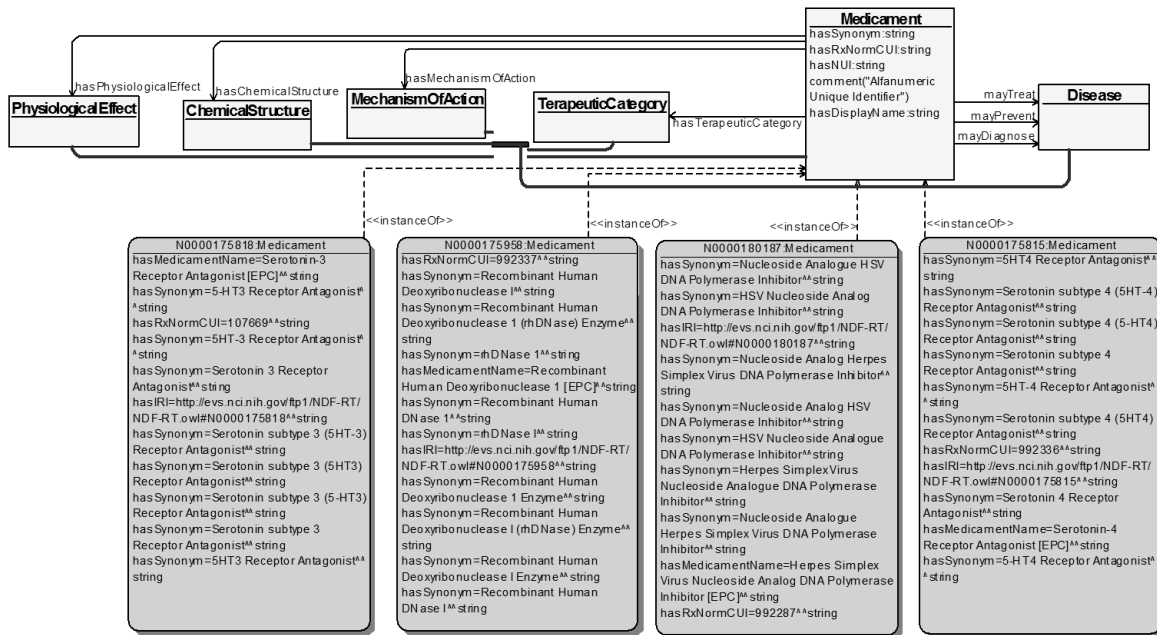


Figure 6. The National Drug File – Reference Terminology (NDF-RT) ontology reference.  
 Source: Author's own

NDF-RT ontology was selected because it provides a revised and valid vocabulary of medicaments with relevant information such as the physiological effect, the chemical ingredients, the mechanism of action, the diseases that each medication may prevent, may diagnose or may treat. This latter information is crucial to correlate pharmacological treatments with diseases.

b) The Human Disease Ontology (**DOID**)<sup>6</sup> is a standardized reference that provides human disease terms, phenotype characteristics and medical vocabulary. DOID integrates mappings with important medical vocabularies: MeSH, ICD, NCI's thesaurus, SNOMED and OMIM. The DOID ontology was selected because it represents disease terms correlated with several important vocabularies.

c) **SYMP**<sup>7</sup>. The symptom ontology (SYMP, 2020) is a collaborative effort between researchers from Colorado State University and Northwestern University, Chicago, IL. This ontology represents general classes of symptoms according with the human anatomy systems. The Symptom Ontology was developed as part of the Gemina project starting in 2005. Currently this ontology is maintained on the project at the Institute for Genome Sciences (IGS) at the University of Maryland.

d) The Logical Observation Identifiers Names and Codes (**LOINC**)<sup>8</sup> system. The LOINC ontology provides standardized vocabularies for laboratory terms. LOINC is currently maintained by the Regenstrief Institute,

<sup>5</sup> <https://www.nlm.nih.gov/research/umls/rxnorm/index.html>

<sup>6</sup> <https://disease-ontology.org/>

<sup>7</sup> <https://bioportal.bioontology.org/ontologies/SYMP/>

<sup>8</sup> <https://bioportal.bioontology.org/ontologies/LOINC>

which in 1994 initiated the LOINC project to address the lack of standardization with multiple laboratories that used different codes for different test observations. LOINC provides a standard way of identifying observations using approximately 41,000 observation terms. 31,000 of these terms are used for laboratory testing.

To determine the ontologies that could be reused and imported into the model, we considered that the importance of each ontology or vocabulary lies mainly in the validity of the knowledge represented. Therefore, we have selected the vocabularies from institutions and groups of researchers that meet the requirements of the OBO Foundry to publish their ontologies in the BioPortal. Additionally, the intelligibility and ease of use of each vocabulary were considered for reutilization and integration.

## Global Ontology Integration

Once the design and implementation of the core concepts have been completed, and having modularized the ontologies for reuse, the integration process consisted of the following steps: integrate core concepts of the Patient Profile in a single ontology file named *PatientElectronicRecord.owl*; select and import modularized ontologies and concepts to fulfill the requirements for the Patient Profile modeling; define the global semantic relationships (object properties) between the core concepts and the imported concepts that are necessary to meet the requirements of the ontology; and evaluate the resulting integrated ontology by representing case studies, and executing the logical consistency.

### Definition of Global Semantic Relations

The list of global semantic relations is shown in Table 4, specifying the domain and range for each.

Table 4  
 Details of the object properties defined for the Patient profile ontology

Object property	Domain	Range	Description
<i>personLivesIn</i>	Patient	Federative Municipality	This relationship was created to indicate the particular municipality and federative entity where the patient lives.
<i>hasBMIClassification</i>	Patient	BMIClassification	This relationship was established to classify the patient's health status with respect to his weight.
<i>hasChemicalStructure</i>	Medicament	ChemicalStructure	This relationship is used to indicate the chemical structure of each drug.
<i>hasMechanismOfAction</i>	Medicament	MechanismOfAction	This relationship is used to describe how a medicament produces an effect in the body. For example, a drug's mechanism of action could be how it affects a specific target in a cell, such as an enzyme, or a cell function, such as cell growth.
<i>hasPhysiologicalEffect</i>	Medicament	PhysiologicalEffect	It is used to establish the biochemical and physiological effects of drugs and their mechanisms of action and the relationship between the concentration of the drug and its effect on an organism.
<i>diseaseHasSymptom</i>	Disease	Symptom	This relationship was established to indicate the symptoms associated to each disease.
<i>requiresLaboratoryTest</i>	Disease	LaboratoryTest	This relationship is used to specify the laboratory tests required for disease confirmation.
<i>hasLaboratoryTest</i>	ClinicalDiagnosis	LaboratoryTest	This relationship is used to indicate the set of laboratory tests necessary for each diagnosis.
<i>hasSymptom</i>	ClinicalDiagnosis	Symptom	It is used to establish the symptoms associated with the clinical diagnosis.
<i>hasDiagnosedDisease</i>	ClinicalDiagnosis	DOID_Disease or NDFRT_Disease	This relationship is used to indicate the diagnosed disease.
<i>hasDiagnosis</i>	Patient	ClinicalDiagnosis	This relationship is used to establish the correlation between the patient and the clinical diagnosis.
<i>hasTreatment</i>	ClinicalDiagnosis	Treatment	This semantic relationship is used to indicate that every diagnosis has an associated treatment.
<i>hasMedication</i>	PharmacologicalTreatment	Medicament	This semantic relationship is used to indicate that all pharmacological treatment consists of the prescription of drugs.

Source: Author's own

## Clinical Diagnosis

According with the National Cancer Institute (NCI) Clinical Diagnosis is defined as the process of identifying a disease, condition, or injury based on the signs and symptoms of a patient, and the patient's health history and physical exam. In order to represent this concept, we included a class Clinical Diagnosis and its semantic relationships with the patient's profile (shown in Figure 7).

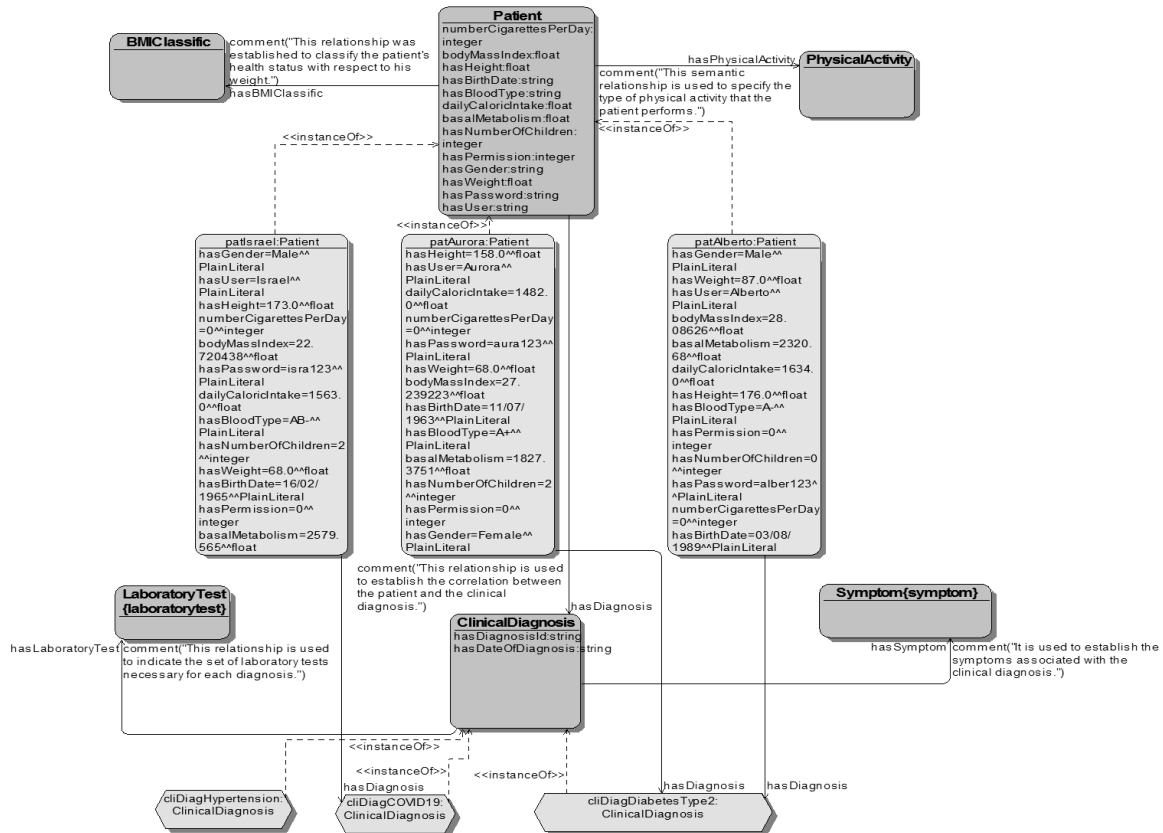


Figure 7. Patient profile and clinical diagnosis integration.  
 Source: Author's own

Through clinical diagnosis, symptoms, and laboratory tests can be established to determine what disease it is. An important relation of the Clinical Diagnosis is the corresponding Treatment. In this ontology model, Treatment is sub-classified into Pharmacological Treatment, Psychological Treatment, and Nutritional Treatment.

## Evaluation

In this section we present the general metrics of the resulting ontology model; three study cases to evaluate the management patient profiles, and Coherence of the entire ontology model through consistency checking.

### Ontology Metrics

The resulting integrated ontology consists of 281097 axioms, with a total of 92 classes (or concepts), 20 object properties, 45 data properties, and 65501 individuals. Figure 8 shows the general metrics of the integrated ontology model, and the set of external imported ontology modules.



Figure 8. Metrics of the ontology model.  
 Source: Author's own

### Study Cases

Three study cases were implemented using the ontology model (Figure 9 shows the details). The first case study corresponds to a 57-year-old female patient, who was born and lives in Irapuato, Guanajuato. She is 1.58 meters tall, weighs 68 kilos, doesn't smoke, has A + blood type. Her body mass index is 27.23, which is why her BMI rating is overweight. Her physical activity is walking, her base metabolism is 1827.37 and her daily caloric intake is 1482. He has a diagnosis of Type 2 Diabetes Mellitus.

The second case study corresponds to a 55-year-old male patient, who was born in Cozumel, Quintana Roo and currently lives in Santiago Yaveo, Oaxaca. The patient weighs 68 kilos, measures 1.73 meters, his blood type is AB-, he does not smoke. His body mass index is 22.72 and his basal metabolic rate is 1579. This patient has a diagnosis of COVID.

The third case study corresponds to a 44-year-old male patient, who was born in Tuxpan, Jalisco and currently lives in Xalisco, Nayarit. The patient weighs 75 kilos, measures 1.71 meters, his blood type is B+, he smokes two cigarettes per day. His body mass index is 25.64 and his basal metabolic rate is 2856. This patient has a diagnosis of Hypertension.

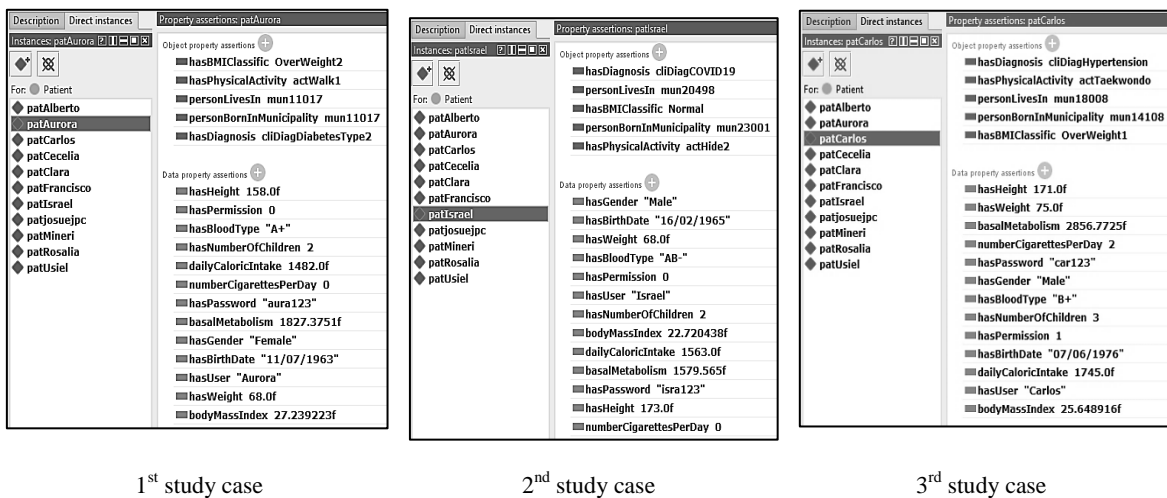


Figure 9. Representation of three patient profiles using the integrated ontology model.  
 Source: Author's own

The three cases presented show the feasibility of representing profiles of patients with chronic diseases such as COVID, diabetes and hypertension. It is worth mentioning that this ontology imports a long vocabulary of diseases, which enables the representation of clinical records of patients with varied and multiple diagnoses and histories of suffered diseases.

### *Ontology Consistency Checking*

Ontology development and evaluation require the verification of Consistency of the model. Reasoning with inconsistent ontologies may lead to erroneous conclusions. The verification of consistency of an ontology is the task carried out by the reasoner program that contrasts what is formally defined in the T-Box with respect to what is asserted in the A-Box. If inconsistencies are found, the reasoner will generate explanations of the inconsistencies. In the case of the ontological model presented in this article, the verification of logical consistency is of particular interest, since modules from external ontologies and vocabularies are imported, which can unintentionally cause the occurrence of logical inconsistencies. Ontology consistency is an important evaluation criteria, because it states that the entire integrated ontology is consistent and has the possibility to execute inference rules to find new concepts and execute queries over the concepts.

Figure 10 shows the result of executing the reasoner over the integrated model. During the execution of the reasoner, Pellet executes a tableaux reasoner that has only one functionality: checking the consistency of an ontology. According with (Sirin, 2007) an ontology is consistent if there is an interpretation that satisfies all the facts and axioms in the ontology. The tableaux algorithm constructs a graph-based representation of the A-Box, where each node is associated with its corresponding type, all property-value assertions are represented as directed edges between nodes. The reasoner repeatedly applies expansion rules until a contradiction is detected. All other reasoning tasks are defined in terms of consistency checking. Therefore, inference about the class hierarchy, the object property hierarchy, and the data property hierarchy derives all non-explicit superclass-subclass relationships which exist in these hierarchies. Regarding the inference about assertions of classes, object properties, and individuals, the reasoner determines if there are no inconsistencies between what is established as class axioms and what is instantiated in the A-Box. From a system management perspective, the results of the reasoner shown in Figure 10 means that user applications or software agents will be able to ask questions and obtain logically correct answers, as well as produce inferences about the facts established in the ontology, facilitating decision making process.

```
INFO 08:04:36 ----- Running Reasoner -----
INFO 08:04:39 Pre-computing inferences:
INFO 08:04:39   - class hierarchy
INFO 08:04:39   - object property hierarchy
INFO 08:04:39   - data property hierarchy
INFO 08:04:39   - class assertions
INFO 08:04:39   - object property assertions
INFO 08:04:39   - same individuals
INFO 08:06:53 Ontologies processed in 136965 ms by Pellet
INFO 08:06:53
```

Figure 10. Ontology Consistency Checking.  
Source: Author's own

Gruber (1983) introduced the definition of Coherence design principle as follows: “an ontology should sanction inferences that are consistent with the definitions”. Based on the results obtained by the reasoner, after integration of ontological models, and some case scenarios it is possible to state the resulting ontology is *Coherent*.



## Conclusions

In this work we have presented an ontological model that integrates two fundamental concepts: the patient's profile and the clinical diagnosis. The model was designed and built considering relevant information requirements for medical decision making. With this ontological model, it was possible to integrate a broad knowledge base that includes medical information from existing valid resources such as DOID, NDF-RT, SYMP, and LOINC ontology references.

The ontology model has been designed following important quality criteria, such as extendibility and reusability. Extensibility was achieved by designing several separate ontologies. For example, the disease ontology is integrated in a single separate owl file, within this ontology the keys and references of the diseases obtained from the original DOID ontology are incorporated. The same criteria were followed for the ontologies of drugs, symptoms and laboratory tests. Thus achieving a modular design, which allows each ontology to be updated separately without affecting the main design of the ontology that integrates the patient's profile and the clinical diagnosis. The reusability of the patient profile is achieved as a consequence of the modular and lightweight design. The patient profile can be reused in different scenarios and applications. Likewise, patient data can be used to perform data mining and find patterns that allow recognizing risk groups by categories, be it gender, age, social context, etc.

Study cases showed the feasibility of the model to be used as a broad knowledge base to support the management of patient profiles. The resulting ontology will be augmented by including more aspects such as management of diabetic patient based on physical activation and treatment plans based on an analysis of the mental and psychological state of the patient. Likewise, an extension will be made of the patients' data on their family history. The long-range purpose is to develop remote patient follow-up and recommendation applications.

## References

- Ansari, R. M., Hosseinzadeh, H., Harris, M., & Zwar, N. (2019). Self-management experiences among middle-aged population of rural area of Pakistan with type 2 diabetes: a qualitative analysis. *Clinical Epidemiology and Global Health*, 7(2), 177-183. DOI: <https://doi.org/10.1016/j.cegh.2018.04.003>
- Chen, Yanshu; Tan, Dixin; Xu, Yihua; Wang, Bin; Li, Xuhui; Cai, Xiaonan; Li, Mengyu; Tang, Cong; Wu, Yun; Shu, Wen; Zhang, Guoliang; Huang, Jiayu; Zhang, Ying; Yan, Yaqiong; Liang, Xunchang; Yu, Songlin. (2020). In Patient Education and Counseling. 103(2):328-335 Language: English. DOI: <https://doi.org/10.1016/j.pec.2019.09.007>.
- Bravo, M., Hoyos Reyes, L. F., & Reyes Ortiz, J. A. (2019). Methodology for ontology design and construction. *Contaduría y Administración*, 64(4), 1–24. DOI: <https://doi.org/10.22201/fca.24488410e.2020.2368>
- Buranarach, M., Chalortham, N., Thein, Y. M., & Supnithi, T. (2011). Design and implementation of an ontology-based clinical reminder system to support chronic disease healthcare. *IEICE transactions on information and systems*, 94(3), 432-439. DOI: <https://doi.org/10.1587/transinf.e94.d.432>
- Chen, R. C., Huang, Y. H., Bau, C. T., & Chen, S. M. (2012). A recommendation system based on domain ontology and SWRL for anti-diabetic drugs selection. *Expert Systems with Applications*, 39(4), 3995-4006. DOI: <https://doi.org/10.1016/j.eswa.2011.09.061>
- Colberg, S. R., Sigal, R. J., Yardley, J. E., Riddell, M. C., Dunstan, D. W., Dempsey, P. C., ... & Tate, D. F. (2016). Physical activity/exercise and diabetes: a position statement of the American Diabetes Association. *Diabetes care*, 39(11), 2065-2079. DOI: <https://doi.org/10.2337/dc16-1728>
- Denova-Gutiérrez, E., Lopez-Gatell, H., Alomia-Zegarra, J. L., López-Ridaura, R., Zaragoza-Jimenez, C. A., Dyer-Leal, D. D., Cortés-Alcala, R., Villa-Reyes, T., Gutiérrez-Vargas, R., Rodríguez-González, K., Escondrillas-Maya, C., Barrientos-Gutiérrez, T., Rivera, J. A., & Barquera, S. (2020). The Association of Obesity, Type 2 Diabetes, and Hypertension with Severe Coronavirus Disease 2019 on Admission Among Mexican Patients. *Obesity* (Silver Spring, Md.). DOI: <https://doi.org/10.1002/oby.22946>
- El-Sappagh, S., & Ali, F. (2016, December). DDO: a diabetes mellitus diagnosis ontology. In *Applied Informatics* (Vol. 3, No. 1, p. 5). Springer Berlin Heidelberg. DOI: <https://doi.org/10.1186/s40535-016-0021-2>
- Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge acquisition*, 5(2), 199-220. DOI: <https://doi.org/10.1006/knac.1993.1008>
- Hempo, B., Arch-int, N., Arch-int, S., & Pattarapongsin, C. (2015). Personalized care recommendation approach for diabetes patients using ontology and SWRL. In *Information science and applications* (pp. 959-966). Springer, Berlin, Heidelberg. DOI: [https://doi.org/10.1007/978-3-662-46578-3\\_114](https://doi.org/10.1007/978-3-662-46578-3_114)

- International Diabetes Federation. IDF Diabetes Atlas, 9th ed. Brussels, Belgium: 2019. Available at: <https://www.diabetesatlas.org>, and last Accessed: 03/12/2020.
- Inzucchi, S. E., Bergenstal, R. M., Buse, J. B., Diamant, M., Ferrannini, E., Nauck, M., Peters, A. L., Tsapas, A., Wender, R., & Matthews, D. R. (2012). Management of hyperglycemia in type 2 diabetes: a patient-centered approach: position statement of the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). *Diabetes Care*, 35(6), 1364–1379. DOI: <https://doi.org/10.2337/dc12-0413>.
- John W. Pelley, in Elsevier's Integrated Review Biochemistry (Second Edition), 2012.
- Lasierra, N., Alesanco, A., Guillén, S., & García, J. (2013). A three stage ontology-driven solution to provide personalized care to chronic patients at home. *Journal of biomedical informatics*, 46(3), 516-529. DOI: <https://doi.org/10.1016/j.jbi.2013.03.006>.
- Li Chen, Dongxin Lu, Menghao Zhu, Muhammad Muzammal, Oluwarotimi Williams Samuel, Guixin Huang, Weinan Li, & Hongyan Wu. (2019). OMDP: An ontology-based model for diagnosis and treatment of diabetes patients in remote healthcare systems. *International Journal of Distributed Sensor Networks*, 15. DOI: <https://doi.org/10.1177/1550147719847112>
- Paganelli, F., & Giuli, D. (2007, May). An ontology-based context model for home health monitoring and alerting in chronic patient care networks. In 21st International Conference on Advanced Information Networking and Applications Workshops (AINAW'07) (Vol. 2, pp. 838-845). IEEE. DOI: <https://doi.org/10.1109/ainaw.2007.90>
- Sirin, E., Parsia, B., Grau, B. C., Kalyanpur, A., & Katz, Y. (2007). Pellet: A practical owl-dl reasoner. *Journal of Web Semantics*, 5(2), 51-53. DOI: <https://doi.org/10.1016/j.websem.2007.03.004>
- Rahimi, A., Liaw, S. T., Taggart, J., Ray, P., & Yu, H. (2014). Validating an ontology-based algorithm to identify patients with type 2 diabetes mellitus in electronic health records. *International journal of medical informatics*, 83(10), 768-778. DOI: <https://doi.org/10.1016/j.ijmedinf.2014.06.002>
- Shaker El-Sappagh, Daehan Kwak, Farman Ali, & Kyung-Sup Kwak. (2018). DMTO: a realistic ontology for standard diabetes mellitus treatment. *Journal of Biomedical Semantics*, 9(1), 1–30. DOI: <https://doi.org/10.1186/s13326-018-0176-y>.
- Sherimon, P. C., & Krishnan, R. (2016). OntoDiabetic: an ontology-based clinical decision support system for diabetic patients. *Arabian Journal for Science and Engineering*, 41(3), 1145-1160. DOI: <https://doi.org/10.1007/s13369-015-1959-4>
- American Diabetes Association. (2020). Standards of medical care in diabetes—2020 abridged for primary care providers. *Clinical diabetes: a publication of the American Diabetes Association*, 38(1), 10. DOI: <https://doi.org/10.2337/cd20-as01>
- SYMP; Symptom Ontology; DOI: <https://doi.org/10.25504/FAIRsharing.ay74mj>; Last edited: July 9, 2019, 10:34 a.m.; Available at: <https://bioportal.bioontology.org/ontologies/SYMP>, and last Accessed: 03/12/2020.