OntoNeH - Reference ontology for hematologic neoplasms with a focus on Hodgkin Lymphoma, especially the classic type

Vanderléia da Silva França [1,†], Henrique Monteiro Cristovão[2,†], Silvia das Dores Rissino[1,†] and Maria das Graças da Silva Teixeira[1,*,†]

[1]*Federal University of Espírito Santo (UFES) – São Mateus, ES – Brazil*
[2]*Federal University of Espírito Santo (UFES) – Vitória, ES – Brazil*

Abstract
Cancer is a pathology formed by various diseases characterized by uncontrolled cell proliferation, including Hodgkin's Lymphoma. There's a remarkable advance in methods for diagnosis and treatment with better prognosis. It should be noted that the most appropriate diagnosis and treatment for cancer undergoes a better understanding of the disease itself. In this scenario, artifacts that facilitate and even expand knowledge of the domain are of paramount importance. Thus, the present work uses the SABIO approach and UFO foundational ontology to develop a reference ontology, called OntoNeH (Reference Ontology for Hematological Neoplasms), which has the domain of Cancer Healthcare and its subareas of study, focusing on Oncohematology. Extensive research has been carried out to obtain domain knowledge and Ontology Engineering was applied for the construction of the proposed ontology. As a result, there is a graphical representation of the domain, which can be beneficial for understanding and communicating the subject among those concerned, in addition to serving as a conceptual model for decision-making in a health information system.

Keywords
Hodgkin's Lymphoma; Conceptual Model; Oncohematology; Ontology; SABIO; UFO.

1. Introduction

The word cancer originates from the Greek *karkinos*, which means crab. It is the name given to a set of more than 100 diseases, having as a common point the disordered growth of cells, which tend to invade neighboring tissues and organs [1]. Cancer is one of the leading causes of death in the world, being one of the main public health problems worldwide. Among the hundreds of different types of cancers are those linked to blood cells, such as lymphomas and leukemias.

Estimates from the *Global Cancer Observatory* reveal that there were 19.3 million new cases of cancer in the world in 2020, with an estimated 544 thousand cases of Non-Hodgkin's Lymphoma, 83 thousand cases of Hodgkin's Lymphoma and 475 thousand cases of Leukemia. In Brazil, in the three-year period from 2023 to 2025, 704 thousand new cases of cancer are expected per year. Of these, the annual estimates for Non-Hodgkin Lymphoma, Hodgkin Lymphoma, and Leukemia are 12040, 3080, and 11540 new cases, respectively [2].

Considering the scenario, the interest in identifying, understanding and treating cancer is increasing. There is a remarkable advance in diagnostic methods, and also the emergence of innovative treatments and, consequently, better prognosis. It is worth noting that the correct diagnosis and the most appropriate treatment for cancer depend on a better understanding of the disease, and the chances of a positive result are directly linked to the rapid action of these activities.

Proceedings of the 17th Seminar on Ontology Research in Brazil (ONTOBRAS 2024) and 8th Doctoral and Masters Consortium on Ontologies (WTDO 2024), Vitória, Brazil, October 07-10, 2024.
* Corresponding author.
† These authors contributed equally.
vanderleiasf@hotmail.com (V. S. França); henrique.cristovao@ufes.br (H. M. Cristovão); silvia.rissino@ufes.br (S. D. Rissino); maria.teixeira@ufes.br (M. G. S. Teixeira)
0009-0007-6856-3116 (V. S. França); 0000-0003-2011-7022 (H. M. Cristovão); 0000-0002-1180-9142 (S. D. Rissino); 0000-0003-0398-8029 (M. G. S. Teixeira)

In this scenario, artifacts that make it possible to facilitate and even expand knowledge of the subject are of paramount importance. And this is where ontologies come in, because the support for information organization and retrieval, such as those based on ontological databases, becomes relevant to deal with the large mass of data produced in Biomedicine related to the topic of interest.

Thus, this article presents the Reference Ontology for Hematological Neoplasia (OntoNeH), a domain ontology with the purpose of representing a fragment of the complex Health domain, focusing on the Oncohematology sub-domain.

The text was designed to be accessible to both healthcare professionals and a technical Computing team. For this reason, it was decided not to delve deeply into design decisions but to include brief explanations of concepts such as UFO, SABiO, OntoUML and Cancer, making it easier for those who may be unfamiliar with one or both fields to understand.

The text is structured as follows: Section 1 has an introduction; Section 2 brings a brief theoretical framework of Ontology and Ontology Engineering, and the methodology used; Section 3 presents a theoretical review of the domain to be modeled, Oncohematology; Section 4 introduces OntoNeH; and finally, section 5 contains the final considerations.

2. Ontology and Ontology Engineering

The term ontology is used in several areas and with different definitions. For the purpose of this article, the definition is the one that meets the interests of the integration between Computing and Health, namely: ontology is a "formal and explicit specification of a shared conceptualization" [3]. In the taxonomy of classification of knowledge organization systems, according to Zeng [4] ontology appears prominently in the highest category as being capable of representing concept models with complex relationships between objects, including the rules and axioms.

There are also several ways of classifying ontologies, one of the best known being proposed by Guarino, in which the distinction is made by the level of generality of the artifact. The categories are: (1) top-level ontologies, which describe broader concepts, and are independent of a specific problem or particular domain (e.g., object and event); (2) and (3) domain ontologies and task ontologies, describe, respectively, vocabulary related to a specific domain (e.g., disease) or a generic task or activity (e.g., treatment), and contribute to domain-dependent problem solving; and (4) application ontologies, which describe concepts dependent on both a domain and a specific task (e.g., radiotherapy treatment for Hodgkin's Lymphoma) [5].

Considering the size and complexity that an ontology can reach, Ontology Engineering (EO) emerged to assist in the process of creating ontologies that satisfy the needs of representation and appropriate use of information [6]. Thus, EO is defined as a set of activities related to the development process, life cycle, methods, tools, and languages for the construction of ontologies. Basically, three types of EO activities are considered: management, development and support [7].

In other words, when working with ontology, there needs to be a development process. For the construction of OntoNeH, the *Systematic Approach for Building Ontologies (SABiO*) [8] was selected, together with a foundation ontology, the *Unified Foundational Ontology (UFO*) [9]. This combination was selected by SABiO to have a solid partnership with UFO, which has already been applied in a number of case studies. It also allows OntoNeH to be integrated in the future with other biomedical ontologies developed with the same foundation, according to the interest of future works.

There are several editors that support the construction of ontologies, and for the purpose of this article Visual Paradigm CE was chosen, which offers free access in its community version and allows integration with the OntoUML plugin [10], enabling the use of UFO stereotypes and meta properties. More details on the use of SABiO, UFO, and OntoUML are discussed in the next section.

2.1. SABiO

SABiO is a systematic approach composed of development and support activities, as shown in Figure 1.

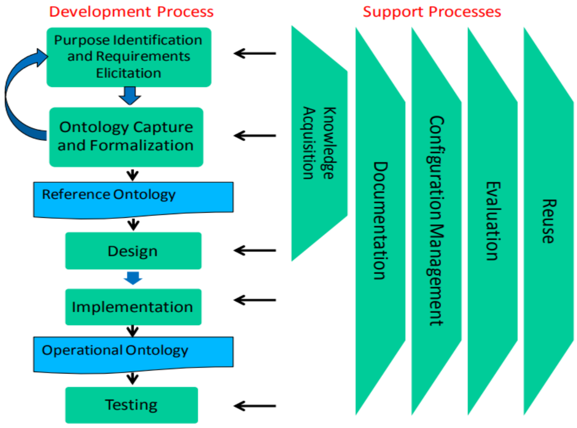

Figure 1: SABiO development process (FALBO, 2014).

The activities carried out in the construction of OntoNeH are described below.

The first activity was the Purpose Identification and Requirements Elicitation, where initially the identification of uses and purposes of ontology was carried out, followed by the elicitation of the requirements. Functional requirements refer to the content to be represented by the ontology and should be written in the form of Competency Questions (CQs). These refer to the questions that the ontology must be able to answer, through which it is possible to achieve an effective way to determine its scope, as well as a way to evaluate it [8]. Table 1 presents some of OntoNeH's CQs. In the non-functional requirements, there are those of interoperability, where OntoNeH was based on a specific foundational ontology, the UFO, which will allow and facilitate its integration with other existing ontologies.

Table 1
OntoNeH's Competency Questions

Code[2]	Question	Justification of Interest
DS_CQ1	What is Staging?	Cancer staging is extremely important as it provides crucial information about the extent and severity of the disease, guiding diagnosis, treatment, and prognosis.
NeS_CQ2	What is Cancer?	Understanding what cancer is, is essential for comprehending one of the main causes of death in the world and for promoting prevention, early diagnosis, and appropriate treatment.
OnS_CQ3	What are the types of Hematological Neoplasm?	Knowing the different types of blood cancer is crucial for diagnosis, treatment, and proper prognosis of the disease.
DiS_CQ4	How is Hodgkin's Lymphoma and its types diagnosed?	Knowing how cancer is diagnosed is essential to ensuring quick and effective treatment, in addition to improving the chances of cure or disease control.

Based on the CQs, in the Capture and Formalization activity, the domain conceptualization was captured, resulting in conceptual modeling, where the main concepts and relevant relationships of the domain were identified, as well as the glossary of terms[3]. As suggested by SABiO, the OntoUML modeling language was used.

In the support process, the Knowledge Acquisition was highlighted, in which various documents were used to study the domain, as well as the exchange of knowledge with a specialist in the area. This stage was very time-consuming, having been the most costly in the construction of OntoNeH; as for Reuse, the reuse of OntoSaúde was made - whose objective is to present a conceptual model that describes a fragment of the Collective Health domain and characterizes the relationships of a

[2] The meaning of the acronyms are in Figure 2.
[3] The glossary of terms can be found on GitHub at https://github.com/VanderleiaFranca/OntoNeH-ONTOBRAS_2024.

person and their possible health problems [11], which allows standardizing and integrating the ontology developed here with future works, as well as taking advantage of existing terms such as Person, Health Problem, Disease, Symptom and Treatment; Documentation, in which the results of the development of the ontology were documented, occurred in parallel with the other processes; and, finally, an Evaluation was carried out that checked whether the model has the appropriate quality. Its focus was on the technical review, in which the authors discussed the developed model at length. Unfortunately, there was no appropriate opportunity to work with experts in the health area, only a brief exchange of ideas and review of the theoretical framework of the domain.

2.2. UFO

The *Unified Foundational Ontology* (UFO) provides guidelines for the development of other ontologies, being a high-quality ontology applied in the development of domain ontologies [12].

UFO is divided into three aspects: UFO-A (*Ontology of Endurants*), UFO-B (*Ontology of Perdurants*) and UFO-C (*Ontology of Social and Intentional Entities*) [13]. In OntoNeH, UFO-A was used, which is based on endurantes, or continuantes, which are entities that do not change their essence over time, that is, they preserve their identity even with the changes that occur during the period of their existence. It deals with issues such as: types and their instances; objects and their properties; relationship between entity and classification; distinctions between types and their relations [9].

Some of the UFO stereotypes used to construct OntoNeH are: `kind`, which provides a principle of identity for its instances, being responsible for structuring the taxonomy representing the domain - Person, Disease and Cell; `subkind`, a specialization of `kind` - Female, Hodgkin's Lymphoma, and Stem Cell are specializations of Person, Disease, and Cell, respectively; `phase` is temporal, that is, it is characterized by a passing phase - Adult, Elderly, Healthy and Unhealthy; and `category`, which does not provide a principle of identity and encompasses entities of different species, with essential characteristics in common - Symptom B which encompasses Fever > 38° (`quality`), Weight Loss >10% in 6 Months (`quality`) and Night sweats (`mode`).

2.3. OntoUML

OntoUML, the Unified Ontology Modeling Language, was conceived as a well-grounded ontologically based version of a UML 2.0 fragment [14]. The language was constructed following a process in which the UML 2.0 metamodel is repaired to ensure an isomorphism in its mapping to the structure defined by the UFO-A foundation ontology; Through formal constraints incorporated into the meta-model of language, the axiomatization of ontology is transferred to it. Thus, OntoUML serves as an engineering tool that enables the use of formal ontological theory in the construction of conceptual and domain models, since it is a language with formal semantics of the real world explicitly defined [9].

The modeling language editor used to build the models was Visual Paradigm, which provides both more formal modeling and casual drawing capabilities, due to the set of design, analysis, and management tools that are provided. Its community version allows free license use [15]. It also allows the installation of the OntoUML plug-in, which enables the use of stereotypes and metaproperties of the language in class diagrams and even the verification and transformation of models [16].

3. The Health Domain

In 1946, the World Health Organization (WHO) defined health as "a state of complete physical, mental and social well-being and not just the absence of disease or infirmity" [17], and its maintenance is fundamental for a full life. However, numerous diseases pose challenges to global health, and among them, cancer emerges as a leading cause of morbidity and mortality. Among the various types of cancer, those related to blood, known as hematological cancer, which encompasses diseases such as leukemia, lymphoma, and myeloma, stand out. A deep understanding of these conditions is crucial for advances in prevention, early detection, and treatment, thus contributing to the promotion of health and quality of life.

3.1. Neoplasm

Most normal cells of a living being grow, multiply and die in an orderly manner, in a continuous and natural process - cell proliferation - however, not all are the same. Some never divide, such as neurons; others, such as the cells of epithelial tissue, divide rapidly and continuously. Cell proliferation is an inherent characteristic of cells and does not necessarily imply the presence of malignancy (some health problem), and may simply respond to the specific needs of the different tissues of the body [1].

Cell proliferation can be controlled or uncontrolled. In controlled growth, there is a localized and self-limited increase in the number of cells of normal tissues that form an organism, caused by physiological or pathological stimuli. In it, the cells are normal or with small changes in their shape and function, and may be the same or different from the tissue in which they are installed. The effect is reversible after the end of the stimuli that caused it [1].

Disordered, or uncontrolled, and autonomous growth of cells, on the other hand, results in a progressive accumulation of neoplastic or tumor cells, which produces a more or less voluminous mass, which is called tumor. Tumors deform, compress, or, depending on the speed of their growth, destroy the organ in which they originate. In addition, tumor cells can infiltrate through the interstices of the organs, reaching the blood and lymphatic circulation and being transported to other organs, where they will grow and form new tumor masses, called metastases [18].

Neoplasms (tumors) can be benign or malignant and their proper classification considers as main characteristics: cell differentiation and anaplasia; tumor growth speed; local invasion and metastasis. This differentiation refers to how similar the tumor parenchyma cell[4] is to the normal parenchymal ancestor cell. Cases in which the cells of the tumor parenchyma are very differentiated and look very similar to their normal counterparts, making it difficult to distinguish normal cells from tumor cells. Thus, although there is accelerated multiplication of tumor cells, reproduction is limited and occurs in such a way that the cells that derive from it have enough time to differentiate to present a functional configuration similar to normal. Cases in which the process of cell multiplication is so fast and aggressive that it does not allow proper cell differentiation, characterize malignant tumors. This cellular indifferentiation of tumor cells is called anaplasia, and tumors that have this characteristic are called anaplastic tumors [18].

Cancer is a malignant neoplasm and its formation process is called carcinogenesis or oncogenesis. The cumulative effects of different carcinogens or carcinogens are responsible for tumor initiation, promotion, progression, and inhibition. However, the individual characteristics of each person should not be left aside, which facilitate or hinder the installation of cell damage [1].

3.2. The Blood

Considered a special type of connective tissue, blood is divided into two parts: plasma and the figurative elements of blood, or blood tissue cells. In it, the cells of the blood tissue are separated by a large amount of extracellular matrix, the plasma [19], which is basically composed of water, carbohydrates, proteins, lipids, electrolytes, and salts [20].

All blood cells originate from a single bone marrow's cell type, pluripotent stem cells. The fates of the cells originated by the stem cells are: self-renewal - they remain as stem cells; and differentiation, in which they diversify into cell types with characteristic specificities, that is, they transform into other cell types [19].

When they multiply, pluripotent stem cells originate: lymphoid cells, which give rise to lymphocytes, which include the so-called B cells and T cells; and myeloid cells, which form erythrocytes (red blood cells), leukocytes (white blood cells), and platelets, which are the smallest blood cells [19]. Produced in the bone marrow and released daily into the circulation, blood cells are specialized cells - cells that have undergone differentiation [20]. And being cells, they are also subject to the development of neoplasms.

---

[4] The parenchyma is represented by the set of "functional" cells and tissues, that is, those that perform the specific functions that characterize a certain organ. In the case of tumors, the parenchyma is made up of neoplastic cells [18].

3.3. Hematologic malignant neoplasms

Hematological malignant neoplasms, or blood cancer, are a category of onco-hematological diseases that affect the bone marrow, blood, and lymphatic system[5]. In general, the types of blood cancer are related to the type of cell in which the malignant neoplasm emerged. They can be classified into: leukemia, lymphoma and myeloma, or multiple myeloma.

Leukemia is a clonal disease that involves bone marrow progenitor cells. Depending on the duplication rate of the neoplastic clone, they are classified as acute or chronic, which in turn are classified as myeloid or lymphoid leukemias, depending on the clonal involvement in cells of the myeloid or lymphoid lineage. Thus, the subtypes of leukemia are determined: Acute Myeloid Leukemia (AML), Acute Lymphocytic Leukemia (ALL), Chronic Myeloid Leukemia (CML) and Chronic Lymphocytic Leukemia (CLL) [21].

Multiple Myeloma (MM) is a type of blood cancer that develops due to genetic alterations of a neoplastic clonal B lymphocyte, which lead to the accumulation of malignant plasma cells [22]. These accumulate in the bone marrow forming plasmacytomas, which can grow inside or outside the bone, that is, intramedullary or extramedullary, respectively. When several plasmacytomas are found inside and outside the bone, MM is characterized [23].

Lymphomas are a group of diseases that are characterized by the appearance of malignant lymphocytes that accumulate in the lymph nodes, producing a clinical condition called adenopathy. They are subdivided into Hodgkin's Lymphoma (HL) and Non-Hodgkin's Lymphoma (NHL) [24].

Non-Hodgkin's Lymphomas comprise a wide and heterogeneous group of lymphoid neoplasms [22], and are due to genetic damage during lymphoid cell maturation [21]. They originate from abnormalities that occur at any stage of the development of B or T lymphoid cells, which give rise to neoplastic clones with distinct molecular, genotypic, and phenotypic characteristics, leading to distinct histological subtypes [22].

Hodgkin's Lymphoma, on the other hand, is a clonal lymphoid malignant neoplasm, resulting from a malignant clone of cells called Reed-Sternberg cells [21], restricted mainly to lymph nodes and lymphoid organs [25]. In 2008 and 2014, respectively, the WHO classified and reviewed Hodgkin's Lymphoma into two groups: Classical Hodgkin Lymphoma (CHL) and Nodular Lymphocyte-predominant Hodgkin Lymphoma (NLPHL). The LHC is subdivided into four histological subtypes: nodular sclerosis, mixed cellularity, lymphocyte predominance, and lymphocyte depletion [21]., which are defined based on the relative proportions of Reed-Sternberg cells, lymphocytes, and fibrosis. NLPHL differs from CHL by the absence of Reed-Sternberg cells and the predominant presence of lymphocyte cells, which are morphological variants of Reed-Sternberg cells with expression of B-lineage markers [26].

The diagnosis of HL is made by histopathological evaluation of the affected lymph node, via excisional biopsy. An entire lymph node is preferably removed for evaluation, and biopsy of inguinal lymph nodes should be avoided if other peripheral lymph nodes are affected. If there are only organs or deep lymph nodes, it is recommended to perform thoracoscopy, laparoscopy, laparotomy, or image-guided biopsy [21].

It is important to confirm the diagnosis with immunophenotype study by immunohistochemistry, as it allows the subclassification of Lymphoma into classic type or nodular lymphocytic predominance [27], in which the presence of Hodgkin and Reed-Sternberg cells is defining of CLH, while the detection of lymphocyte predominant cells (LP) is necessary for the diagnosis of NLPHL [28]. This distinction is important, as NLPHL and CLH have different natural histories, prognoses, and treatments [29].

Cancer staging is done by clinical, laboratory and imaging tests [24], and is performed after diagnosis, and is used to define the extent of the disease, as well as the presence or absence of factors that affect prognosis. Thus, it is possible to attribute a stage to the disease [25].

---

[5]It is a system of specialized vessels and organs that, like the venous system, transports fluids throughout the body and regulates their quantity. It is also a vital part of the immune system, which helps fight infections and diseases [30].

4. OntoNeH

4.1. OntoNeH Graphical Representation

A graphical representation is essential to illustrate the relationships between entities and their characteristics in a clear and concise way, and has been applied in OntoNeH to facilitate the visualization and understanding of the complex structures of the Oncohematology domain.

4.1.1. Overview

The Purpose of OntoNeH is to capture, represent and organize knowledge related to Oncohematology; facilitate the integration of information and data from different sources and systems related to hematological neoplasms; and to promote a consistent and standardized structure to describe concepts and relationships, making it possible to integrate and expand them to other ontologies in the Healthcare domain.

As for Use, the proposed ontology will facilitate the dissemination and understanding of knowledge, serving as an educational tool for training and qualification of professionals, students and researchers in the field; it can be used in the implementation and development of computational systems and tools to support clinical decision-making that use ontology to offer evidence-based recommendations.

OntoNeH has six sub-ontologies, as identified in Figure 2, which presents a package diagram with the overview of the ontology, its sub-ontologies and relationships.

The packages were chosen in order to follow the natural process that goes from the formation of the disease to its treatment. In this article, the sub-ontologies Disease, Neoplasm, Oncohematology and Diagnosis are addressed.

A color legend was created to identify the origin of the terms in the different diagrams, since each package (a sub-ontology) is represented by a color and the terms from it are marked with this color.

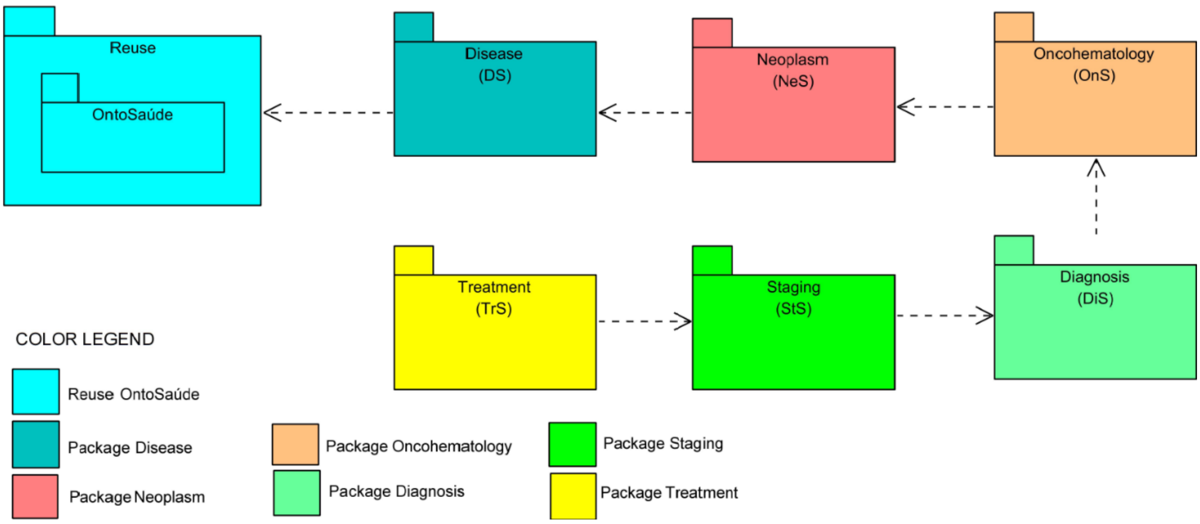

Figure 2: OntoNeH - Overview Diagram[6].

4.1.2. Disease Sub-ontology

The Disease sub-ontology, whose diagram is in Figure 3, reuses OntoSaúde through more generic terms of a Disease, in addition to preparing the integration with other ontologies that are also based on OntoSaúde. Since they are related works and ontological models with the same foundation, that is, UFO, the changes proposed in OntoSaúde are an evolution of it, visualized from its integration with other ontologies, without altering its essence or the commitments made.

[6] Images of OntoNeH's subontologies can be found on GitHub at https://github.com/VanderleiaFranca/OntoNeH-ONTOBRAS_2024.

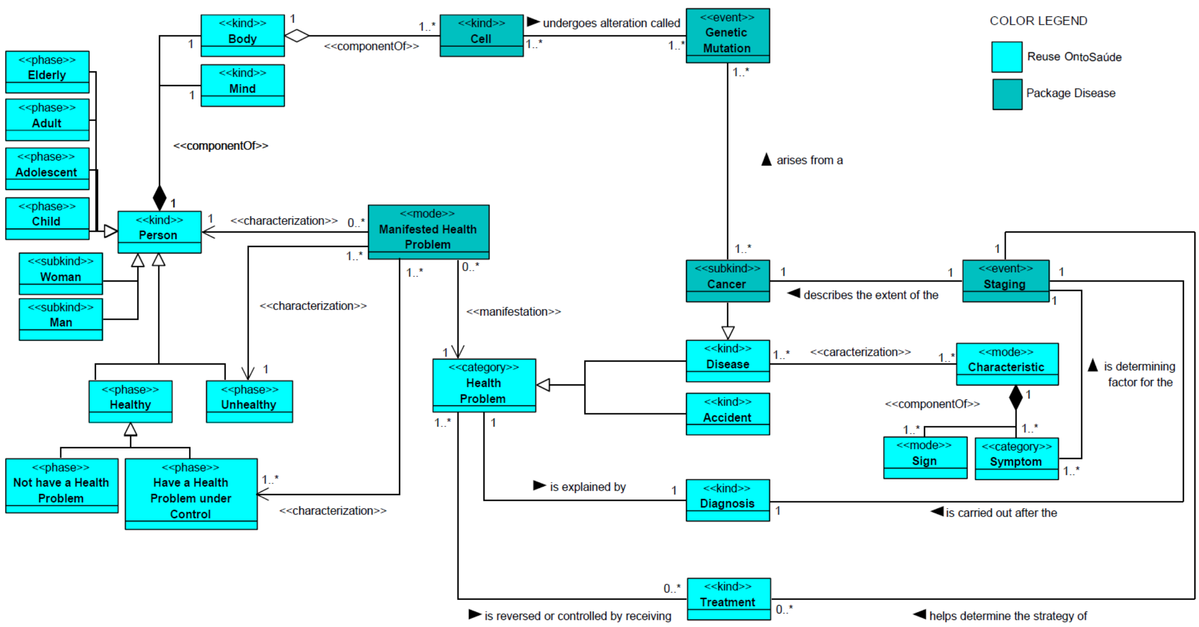

Figure 3: Diagram of the Disease sub-ontology.

A Person can be a Man or a Woman and is at a certain stage of life, whether they are a Child, Adolescent, Adult or Elderly. A Person is formed by Body and Mind. The Cell is a component of the Body and undergoes Genetic Mutation, which in turn causes Cancer, which is a type of Disease. Even more, a Person can be Healthy - Not have a Health Problem, or Have a Health Problem under Control, and these concepts are characterized by a Manifested Health Problem. A Person can also be Unhealthy when they have a Health Problem, be it an Accident or an Illness.

Answering DS_CQ1, Staging describes the extent of Cancer, being an event performed after Diagnosis, and which explains a Health Problem. It will also assist in defining the Treatment strategy. In turn, Symptom is a factor used to determine Staging, and together with the Sign are Characteristic of a Disease.

It is worth highlighting a change identified in OntoSaúde, based on the needs of OntoNeH: the addition of the concept of Manifested Health Problem, which is a characteristic `(characterization)` of a Person, and a manifestation `(manifestation)` of a Health Problem, which allowed a better understanding of the relationships between the concepts. Changes have also been made to the stereotypes of some concepts to better suit, such as Symptom, from `mode` to `category`, since the nature of them is different types of things.

4.1.3. Neoplasm Sub-ontology

Figure 4 presents the Neoplasm sub-ontology, which represents the process of formation of a neoplasm, and its distinction from cancer, which is a malignant neoplasm.

Bone Marrow is composed of Stem Cells, which can undergo Self-Correction or Differentiation. One of its subtypes is the Pluripotent Stem Cell, which when it undergoes Differentiation is said to be a Differentiated Pluripotent Stem Cell, giving rise to the Myeloid Cell and Lymphoid Cell. In turn, these can give rise to the Leukocyte, Erythrocyte, Platelet, B Lymphocyte and T Lymphocyte subtypes. These subtypes are part of the Figured Element of Blood, which is produced in the Bone Marrow. Another component of Blood is Plasma, composed of Electrolyte, Lipid, Carbohydrate, Water, Protein and Salts.

A Cell undergoes Cell Multiplication. When this is a Controlled Multiplication it is said to be a Normal Cell. However, when Uncontrolled Multiplication occurs, it is classified as a Cell with Mutation, which is part of a Neoplasm. This can specialize in Benign Neoplasm, which occurs when there is Negative Capacity, that is, when the Cell with Mutation does not have Ability to Invade Neighboring Tissue; and in Malignant Neoplasm, also called Cancer, a type of Neoplasm (or Tumor), formed by a Cell with Mutation that results from Cell Multiplication in an uncontrolled way (Uncontrolled Multiplication) and is characterized by the Positive Ability to invade neighboring tissue (Ability to Invade Neighboring Tissue), and that's the answer to NeS_QC2. When Malignant Neoplasm occurs in the Blood it is called Hematological Neoplasm.

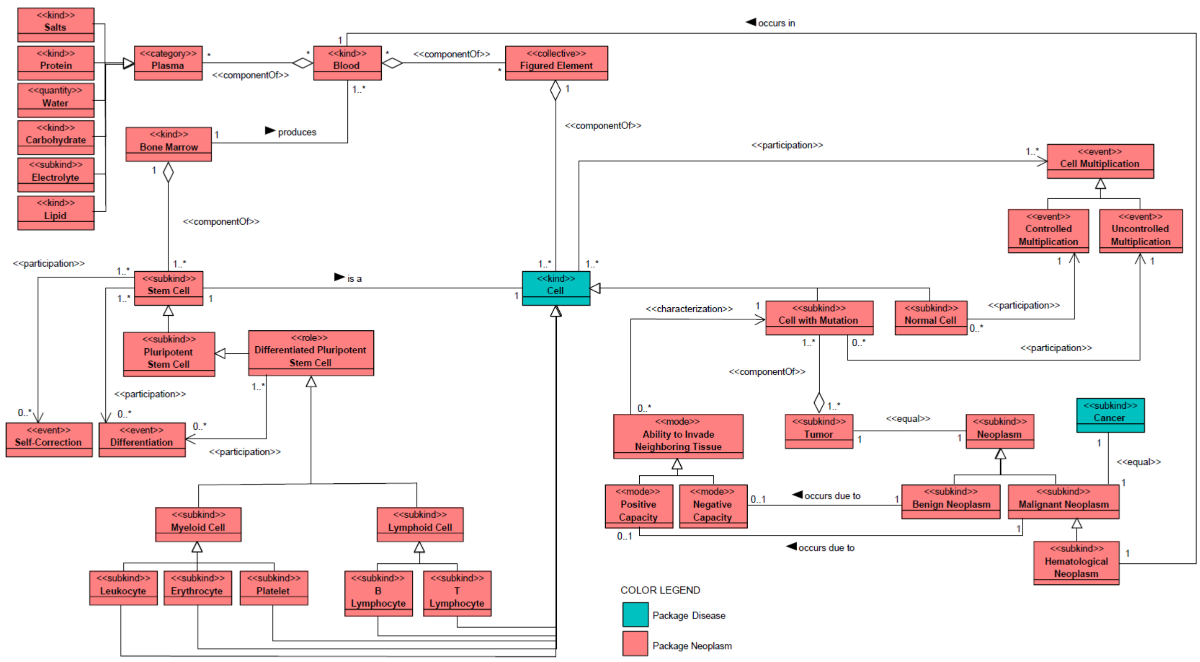

Figure 4: Diagram of the Neoplasm sub-ontology.

Several changes occurred in the sub-ontology during its development. Concepts were eliminated, such as the division of Bone Marrow into red marrow and yellow marrow, being replaced by Stem Cell, in a `componentOf` relation of the marrow - which allowed greater fluidity and compatibility with the concept of Pluripotent Stem Cell.

One point that generated intense debate was the choice of the UFO stereotypes of Positive Ability and Negative Ability. At first, the instinct was to classify them as `phase`, as their concepts may instinctively suggest, in the same way that occurred with the concepts of Not Having a Health Problem and Having a Health Problem under Control, Healthy and Unhealthy in the Disease sub-ontology. However, a benign nodule will remain benign, that is, it does not become malignant [31], and so there is no possibility of change between the Positive Capacity and Negative Capacity, as a `phase` would ask – even if the change does not occur, the possibility should exist. It was then decided to use the `mode` stereotype, as they are specializations of Ability to Invade Neighboring Tissue, a characteristic of the Cell with Mutation.

4.1.4. Oncohematology Sub-ontology

Figure 5 represents the Oncohematology sub-ontology, which brings the different types of cancer in the blood.

A Hematological Neoplasm can specialize in Leukemia, Lymphoma or Myeloma, which is equivalent to Multiple Myeloma, answering OnS_QC3. In turn, Lymphoma specializes in Hodgkin's Lymphoma, whose subtypes are Nodular Lymphocyte-predominant Hodgkin Lymphoma and Classical Hodgkin Lymphoma, with the subtypes Nodular Sclerosis, Mixed Cellularity, Lymphocyte Depletion and Lymphocyte-Rich; and Non-Hodgkin Lymphoma.

A Leukemia can be Acute or Chronic. The first can be classified into Acute Lymphocytic Leukemia and Acute Myeloid Leukemia. The second has associated the subtypes Chronic Lymphocytic Leukemia and Chronic Myeloid Leukemia.

Because it is a diagram that deals with the taxonomy of Hematological Neoplasms, that is, classify the types of cancers in the blood, this was the diagram that generated the least doubts related to modeling decisions.

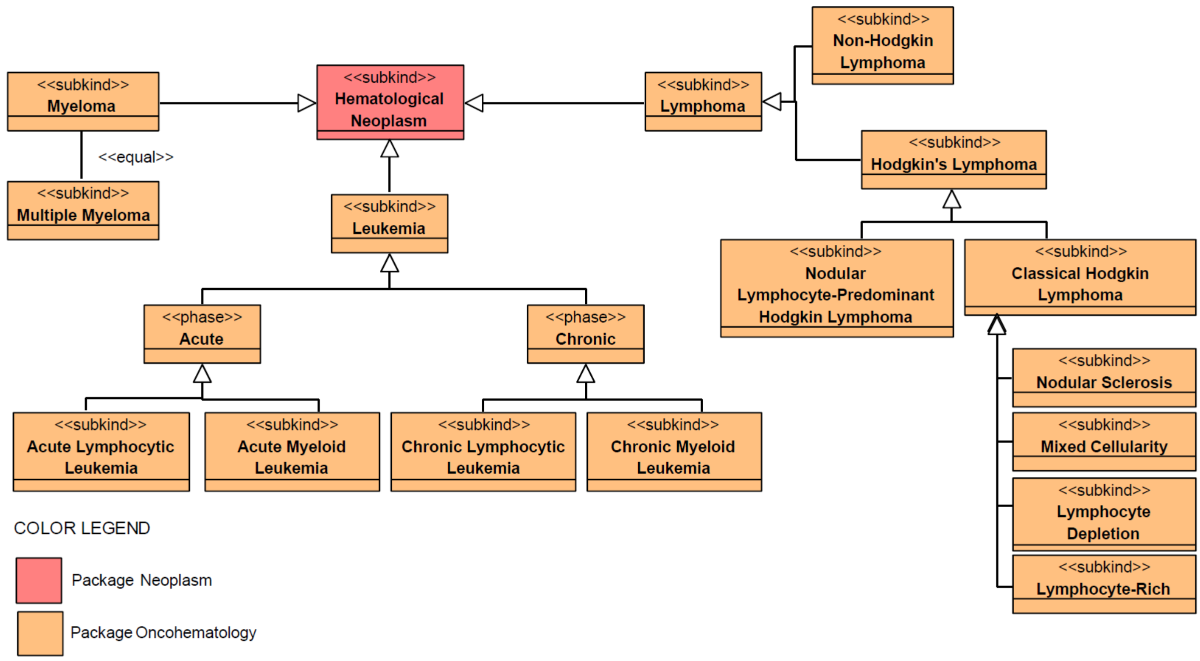

Figure 5: Diagram of the Oncohematology sub-ontology.

4.1.5. Diagnostic Sub-ontology

Figure 6 presents the Diagnosis sub-ontology, which shows the characteristics of Lymphoma, as well as the process for diagnosing Hodgkin's Lymphoma and the aspects considered for staging the disease.

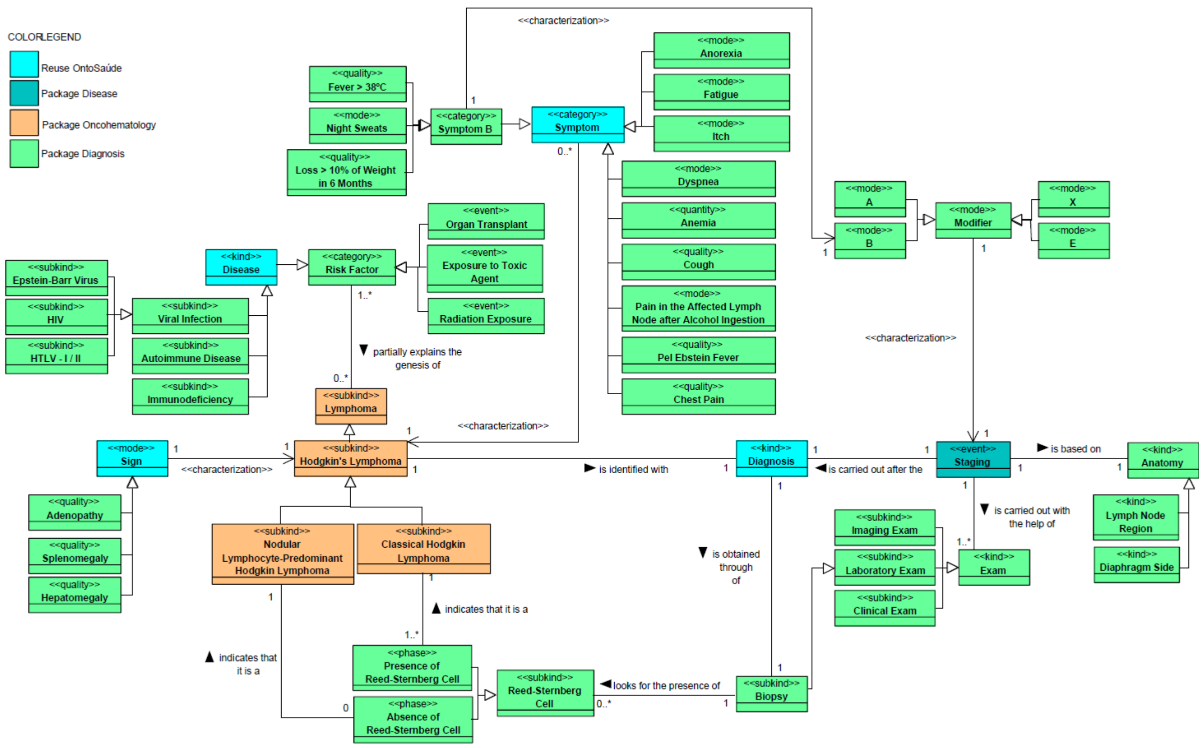

Figure 6: Diagram of the Diagnostics sub-ontology.

The Risk Factor - factors whose exposure increases the probability of occurrence of the disease - partially explains the genesis of Lymphoma, and among them are: Exposure to Toxic Agent; Radiation Exposure; Organ Transplant; and certain diseases such as Autoimmune Disease, Immunodeficiency, and Viral Infection, which can be HIV, HTLV-I/II, and Epstein-Barr Virus.

Hodgkin's Lymphoma has some characteristics as Sign, including Adenopathy, Splenomegaly and Hepatomegaly. As Symptom, the following are identified, among others: Chest Pain, Cough, Anemia, Dyspnea and Loss > 10% of Weight in 6 Months. Hodgkin's Lymphoma is identified through a Diagnosis, which is obtained by performing a Biopsy, that looks for the Presence of Reed-Sternberg Cell, and if it's positive, indicate Classical Hodgkin Lymphoma and if it's negative, Absence of Reed-Sternberg Cell, indicate Nodular Lymphocyte-predominant Hodgkin Lymphoma, as an answer for DiS_CQ4.

After the Diagnosis, Staging occurs, which is based on Anatomy as a means of classification, as well as on Modifier, which include: A; B, which gives its name to Symptom B; X; and E. Anatomy includes the Diaphragm Side and the Lymph Node Region. Staging is performed with the help of Exam, which can be Clinical Exam, Imaging Exam and Laboratory Exam, in which a Biopsy is a specialization of the latter.

This sub-ontology initially included the details of staging and treatment, but for better visualization and representation of the relationships they were separated into different sub-ontologies. One issue highlighted, and much discussed, was about the concept of Staging - in the beginning the staging classifications (stages I, II, III and IV) were presented as specializations of the staging itself, but after questioning who has stage, it was seen that the classifications are related to the disease, as it is the disease that specializes them.

5. Conclusion

OntoNeH is a reference ontology in the field of Health, focusing on the oncology sub-area of blood malignancies, i.e., blood neoplasms. It proposes to establish a basis for concepts about cancers that affect the blood such as myeloma, leukemia and lymphoma, the latter being the most detailed fragment in this text, especially Hodgkin's Lymphoma and its Classic subtype.

General concepts of the domain and subsequent specification for the characteristics of Hodgkin's Lymphoma were addressed and OntoSaúde concepts were reused, which contributes to the standardization of OntoNeH and integration with other ontologies, such as OntoTB [32] and OntoHI [33], which are related works, making use of the same methodology and rationale, SABiO and UFO, respectively. This allows the model to be prepared to address other diseases. Thus, it is expected that future works can expand the study of the disease and create ontologies for other hematological neoplasms or even solid cancers; to publish expanded OntoNeH, adding the sub-ontologies that were not exposed here; to map the reference ontology to operational ontology, making it machine-processable; conduct a search for ontological models in the cancer domain that use foundations other than UFO, seeking potential conceptual integrations. It is recommended that in the future the studies would be made with the assistance of a professional in the field of the domain, hematologist or oncologist, which will allow a more consistent work in accordance with reality in the verification of the information collected and a final work validated in a broader and more complete way, since here this was restricted to a discussion between the authors, who conducted a technical review.

The use of ontologies in the field of oncology can provide significant benefits from a humanitarian point of view. Since they allow for the semantic and structured representation of medical knowledge, ontologies can be used to integrate and organize information coming from different sources, ensuring that healthcare professionals have access to structured data in an appropriate way. They can facilitate collaboration between health professionals and researchers, allowing an efficient exchange of knowledge and data, leading to the acceleration of scientific research in the field of cancer. This can result in the development of diagnostic support systems, enabling faster and more accurate diagnoses, which allows for appropriate medical intervention in the early stages of the disease. They can also be used in the creation of structured educational resources that facilitate the understanding of concepts related to the disease, promoting awareness among health professionals, patients and the general public, contributing to its prevention and early detection.

The benefits related to the use of ontologies can also be economical. Cost reduction combined with better treatment results due to faster and more accurate decision-making by health professionals with the use of better tools to support clinical decision-making, which would be created with the represented and shared knowledge provided by ontologies. The optimization of resources invested in research and consequent discovery of new therapeutic approaches, since ontologies facilitate the organization and categorization of information, which accelerates research processes. A more in-depth understanding of cancer, facilitated by ontologies, can contribute to a more efficient management of resources such as better allocation of equipment and personnel, as well as lead to an improvement in the quality of care, resulting in improved quality of life for patients, potentially reducing costs associated with complications and hospital readmissions. On the other hand, the standardization in the representation of concepts, provided by ontologies, facilitates communication between health systems and institutions, leading to a reduction in the redundancy of information, minimization of errors and improvement in the efficiency of processes, which can have positive implications for operating costs.

The large amount of domain information available requires tools that speed up analysis and improve its accuracy and understanding. In this scenario, ontologies play a fundamental role. Even more so when fast, integrated, detailed and clear access to information is fundamental for a professional's decision-making, such as in the treatment of a person. OntoNeH can be applied in a variety of areas, from clinical practice to research and technological development. It can provide a clear and structured model of knowledge capable of assisting in the teaching of health professionals and students. In the field of research, ontology offers a standardization of the knowledge obtained in clinical and laboratory studies, which would facilitate the comparison of data and the results of different researches, in addition to allowing the creation of new computational models for the analysis of large volumes of data. Integrating OntoNeH with electronic medical record systems, decision-making support tools, and other medical software, coupled with a standardized database, can lead to better interoperability between different hospital systems, ensuring that all data is interpreted consistently. A database with structured information from OntoNeH can be used to build a tool that helps to reduce the waiting lines for triage and improve its flow, especially if integrated with OntoSaúde and expanded to other domains of neoplasms. In addition, it can be used to train artificial intelligence (AI) and machine learning algorithms in assisted diagnosis systems, allowing them to identify patterns and make predictions about the development of hematological malignancies based on structured clinical data, which can lead to the creation of predictive models and comparative studies between different types of cancer, making it possible to identify new patterns and relationships that can be exploited for the development of more effective therapies.

References

[1] INCA. ABC do câncer: abordagens básicas para o controle do câncer. José Alencar Gomes da Silva National Cancer Institute. 6. ed. Rio de Janeiro, 2020. Available at: < https://www.inca.gov.br/sites/ufu.sti.inca.local/files/media/document/livro-abc-6-edicao-2020. pdf>. Accessed on: March 30, 2023.

[2] INCA. Estimativa 2023: incidência de câncer no Brasil. National Cancer Institute / Ministry of Health. Rio de Janeiro, 2022. Available at: < https://www.inca.gov.br/sites/ufu.sti.inca.local/files//media/document//estimativa-2023.pdf>. Accessed on: June 02, 2023.

[3] BORST, W. N. Construction of Engineering Ontologies. Thesis, University of Tweenty – Centre for Telematics and Information Technology. Enschede - Nederland, 1997.

[4] ZENG, Marcia Lei. Knowledge organization systems (KOS). Knowledge Organization, [S. l.], v. 35, n. 2–3, p. 160–182, 2008. DOI: 10.5771/0943-7444-2008-2-3-160. Available at: <https://www.nomos-elibrary.de/10.5771/0943-7444-2008-2-3-160.pdf>. Accessed on: July 22, 2024.

[5] GUARINO, Nicola. Formal Ontologies and Information Systems. Conference: FOIS'98 Conference At: Trento, Italy. June 1998. Available at: < https://www.researchgate.net/publication/272169039_Formal_Ontologies_and_Information_Systems>. Accessed on: March 29, 2023.

[6] ISOTANI, Seiji; BITTENCOURT, Ig Ibert. Dados Abertos Conectados. Center for Studies on Web Technology-Web.br. Novatec Publishing, 2015. Available at: https://ceweb.br/livros/dados-abertos-conectados//. Accessed on: March 29, 2023.

[7] GÓMEZ-PÉREZ, A.; FERNANDEZ-LOPEZ, M.; CORCHO, O. Ontological Engineering. Springer-Verlag Berlin, Heidelberg, 2004. ISBN:1852335513.

[8] FALBO, Ricardo de Almeida. SABiO: Systematic approach for building ontologies. CEUR Workshop Proceedings, 1301. 2014. Available at: <https://ceur-ws.org/Vol-1301/ontocomodise2014_2.pdf>. Accessed on: March 29, 2023.
[9] GUIZZARDI, Giancarlo. Ontological Foundations for Structural Conceptual Models. ISBN 90-75176-81-3, Universal Press. Enschede, The Netherlands, 2005. Available at: < https://nemo.inf.ufes.br/wp-content/papercite-data/pdf/ontological_foundations_for_structura l_conceptual_models_2005.pdf>. Accessed on: October 10, 2023.
[10] SUCHÁNEK, Marek. OntoUML specification Documentation. 2022. Available at: <https://ontouml.readthedocs.io/_/downloads/en/latest/pdf/>. Accessed on: July 09, 2023.
[11] COSTA, Danielli do Reis; TEIXEIRA, Maria das Graças da S.; RISSINO, Silvia das Dores; GUARNIER, Thayza S. O Uso da Abordagem SABiO na Construção do Overview de OntoSaúde em Engineering Ontologies and Ontologies for Engineering. Celebrating Ricardo FALBO's Career. Editors: ALMEIDA, João Paulo A.; GUIZZARDI, Giancarlo. Nemo, Vitória - ES, Brazil. 2020. Available at: https://nemo.inf.ufes.br/wp-content/papercite-data/pdf/engineering_ontologies_ and_ontologies_for_engineering_celebrating_ricardo_falbos_career_2020.pdf>. Accessed on: June 03, 2023.
[12] GUIMARÃES, Rachel Cristina Mello; GUIZZARDI, Giancarlo; GOTTSCHALG-DUQUE, Cláudio. Nomeação de Elementos Ontológicos para Criação de Ontologias: uma Proposta Metodológica. University of Brasilia. Brasília, 2015. Available at: <http://www.realp.unb.br/jspui/handle/10482/22049?locale=es>. Accessed on: April 12, 2023.
[13] GUIZZARDI, Giancarlo; ALMEIDA, João Paulo Andrade; GUIZZARDI, Renata S.S.; FALBO, Ricardo. Ontologias de Fundamentação e Modelagem Conceitual Núcleo de Estudos em Modelagem Conceitual e Ontologias (NEMO). Federal University of Espírito Santo (UFES). Vitória - ES, 2009. Available at: https://nemo.inf.ufes.br/wp-content/papercite-data/pdf/ontologias_de_fundamentacao_e_modelagem_conceitual_2009.pdf>. Accessed on: October 10, 2023.
[14] OMG® SDO. About The Unified Modeling Language Specification Version 2.0. Object Management Group® (OMG®) Standards Development Organization (SDO), 2005. Available at: <https://www.omg.org/spec/UML/2.0/>. Accessed on: November 20, 2023.
[15] Visual Paradigm. Latest version: 17.1 of May 20, 2023. Available at: <https://www.visual-paradigm.com/editions/>. Accessed on: November 20, 2023.
[16] FONSECA, Claudenir M.; SALES, Tiago Príncipe; VIOLA, Vitor; BASSETTI, Lucas. Plug-in OntoUML para diagrama visual. GitHub, 2021. Available at: <https://github.com/OntoUML/ontouml-vp-plugin>. Accessed on: November 20, 2023.
[17] WHO. Constitution. World Health Organization (WHO). 2023. Available at: <https://www.who.int/about/governance/constitution>. Accessed on: November 22, 2023.
[18] MEDRADO, Leandro. Carcinogênese: Desenvolvimento, Diagnóstico e Tratamento das Neoplasias. 1st ed. São Paulo, SP: Editora Érica, 2015. Electronic book. ISBN 9788536520827. Available at: <https://integrada.minhabiblioteca.com.br/#/books/9788536520827/>. Accessed on: March 30, 2023.
[19] OLIVEIRA, Leiny Paula. Tecido sanguíneo e hematopoiético. Animal Anatomy Laboratory (LAAN) - Federal University of Goiás, 2015. Available at: <https://files.cercomp.ufg.br/weby/up/707/o/sangue.pdf?1448365978>. Accessed on: June 02, 2023.
[20] AZEVEDO, Maria Regina Andrade de. Hematologia Básica: Fisiopatologia e Diagnóstico Laboratorial. 6th ed. Rio de Janeiro, RJ: Thieme Brasil, 2019. Electronic book. ISBN 9788554651381. Available at: <https://integrada.minhabiblioteca.com.br/#/books/978855465138 1/>. Accessed on: October 03, 2023.

[21] VIEIRA, Sabas Carlos. Oncologia básica para profissionais de saúde. EDUFPI Publishing House. Teresina, 2016.

[22] HOFF, Paulo M. G.; KATZ, Artur; CHAMMAS, Roger; FILHO, Vicente O; NOVIS, Yana S. Tratado de Oncologia. Editora Atheneu. São Paulo, 2013.

[23] Scientific Medical Committee of ABRALE. Manual - O que você precisa saber sobre o Mieloma Múltiplo. Brazilian Association of Lymphoma and Leukemia (ABRALE), 2023. Available at: https://www.abrale.org.br/wp-content/uploads/2023/03/Manual-Mieloma-Multiplo-2023.pdf>. Accessed on: July 09, 2023.

[24] HOFFBRAND, A. Victor; MOSS, Paul A. H. Fundamentos em hematologia de HOFFbrand. Editora Artmed. 7. ed. Porto Alegre, 2018.

[25] CHABNER, Bruce A.; LONGO, Dan L. Manual de oncologia de Harrison. Editora AMGH - 2nd ed. Porto Alegre, 2015.

[26] GAMBONI, Mercedes; MIZIARA, Elias F. Manual de Citopatologia Diagnóstica. Barueri, SP: Editora Manole, 2013. Electronic book. ISBN 9788520436066. Available at: <https://integrada.minhabiblioteca.com.br/#/books/9788520436066/>. Accessed on: October 02, 2023.

[27] SANDES, Alex F. Diagnósticos em hematologia. 2nd ed. Barueri, SP: Editora Manole, 2021. Electronic book. ISBN 9786555760019. Available at: <https://integrada.minhabiblioteca.com.br/#/books/9786555760019/>. Accessed on: October 02, 2023.

[28] CONITEC. Diretrizes Diagnósticas e Terapêuticas do Linfoma de Hodgkin. National Commission for the Incorporation of Technologies in the Unified Health System, MINISTRY OF HEALTH. Brasília - DF, 2020. Available at: https://www.gov.br/conitec/pt-br/midias/consultas/relatorios/2020/ddt_linforma_de_hodgkin_cp_11_2020.pdf>. Accessed on: October 5, 2023.

[29] GOVINDAN, Ramaswamy; MORGENSZTERN, Daniel. Oncologia. (Washington Manual™). 3rd ed. Rio de Janeiro, RJ: Thieme Brazil, 2017. Electronic book. ISBN 9788567661940. Available at: <https://integrada.minhabiblioteca.com.br/#/books/9788567661940/>. Accessed on: October 02, 2023.

[30] CARMO, Lívia Lourenço do. Sistema Linfático. Kenhub GmbH. 2023. Available at: <https://www.kenhub.com/pt/library/anatomia/sistema-linftico>. Accessed on: July 23, 2023.

[31] FILHO, Dr. José d'OLIVEIRA Couto. Um nódulo benigno pode se transformar em câncer no futuro? Dr. Couto Clinic, Londrina - PR, 2020. Available at: <http://clinicadrcouto.com.br/d%C3%BAvidas-frequentes.html#:~:text=Um%20n%C3%B3dulo%20benigno%20pode%20se,ou%20n%C3%A3o%20se%20torna%20malignant.>. Accessed on: November 14, 2023.

[32] GUARNIER, Thayza S.; TEIXEIRA, Maria das Graças da S.; COSTA, Danielli dos Reis; SALES, Carolina M. M.; RISSINO, Silvia das Dores. Um Modelo Conceitual Baseado em Ontologia para Doenças Infecciosas com Ênfase em Tuberculose (An Ontology-Based Conceptual Model for Infectious Diseases Emphasiing Tuberculosis) ONTOBRAS (2020). Available at: https://www.semanticscholar.org/paper/Um-Modelo-Conceitual-Baseado-em-Ontologia-para-com-Guarnier-Teixeira/68834d4fc4e0cc34c756fb666c570367aca17b17>. Accessed on: June 03, 2023.

[33] COSTA, Danielli dos Reis. Aplicação da Metodologia SABiO na Construção de OntoHI. Monograph, Advisor: Profa. Dr. Maria das Graças da Silva Teixeira. Federal University of Espírito Santo – UFES, Collegiate Technological Center of the Computer Science Course. São Mateus - ES, 2020.